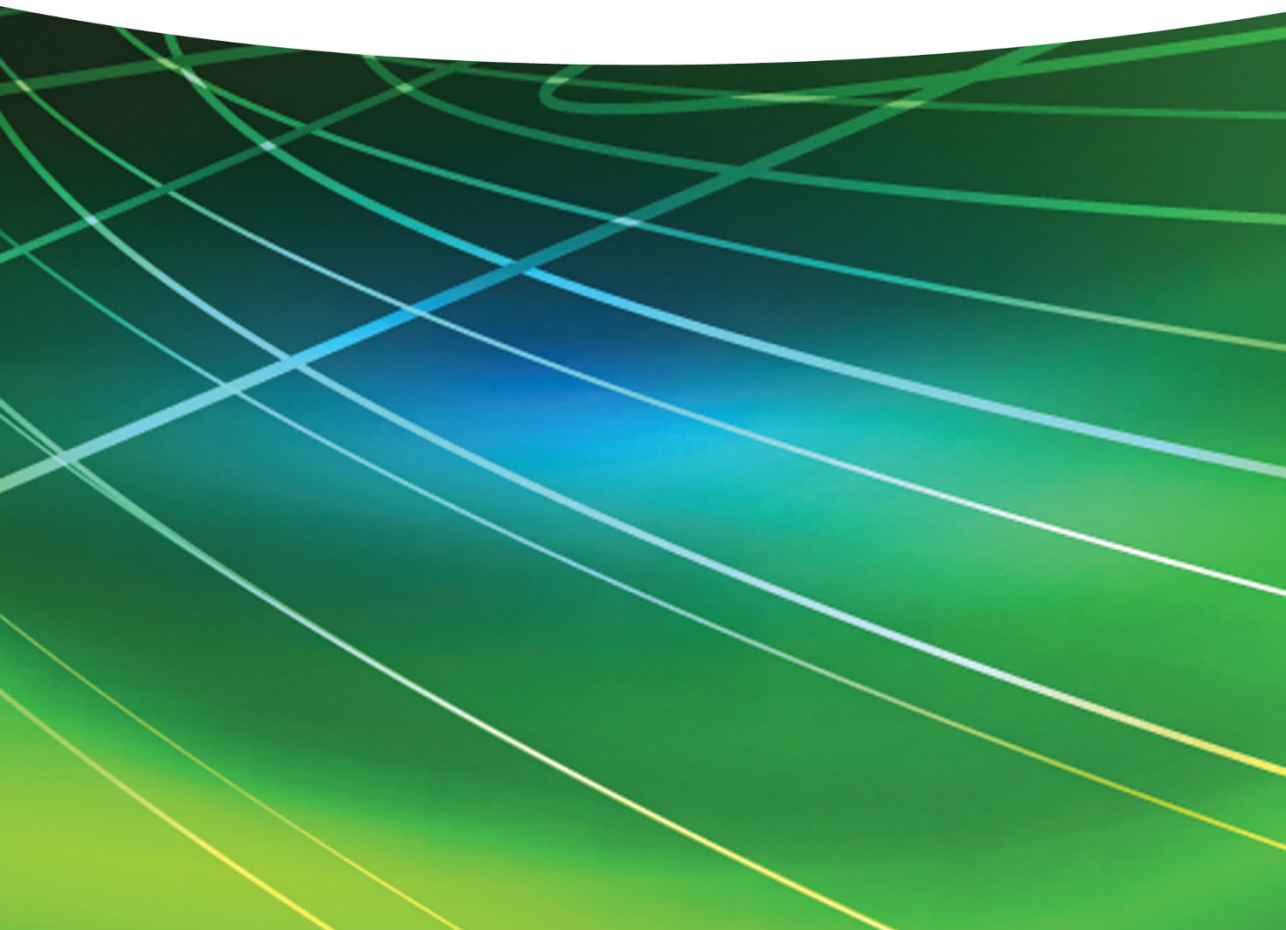
WILEY-VCH

Claude Cohen-Tannoudji, Bernard Diu, and Franck Laloë

Quantum Mechanics

Volume III: Fermions, Bosons, Photons, Correlations, and Entanglement



QUANTUM MECHANICS

Volume III

Fermions, Bosons, Photons, Correlations, and Entanglement

Claude Cohen-Tannoudji, Bernard Diu,
and Franck Laloë

Translated from the French by **Nicole Ostrowsky and Dan Ostrowsky**

Authors

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

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

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

Cover Image
© antishock/Getty Images

First Edition

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

Library of Congress Card No.:
applied for

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

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

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

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

Print ISBN 978-3-527-34555-7
ePDF ISBN 978-3-527-82274-4
ePub ISBN 978-3-527-82275-1

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

Printed on acid-free paper.

Directions for Use

This book is composed of chapters and their complements:

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

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

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

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

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

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

Foreword

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

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

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

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

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

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

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

Acknowledgments:

Volumes I and II:

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

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

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

Volume III:

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

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

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

Volume I

Table of contents vii

I WAVES AND PARTICLES. INTRODUCTION TO THE BASIC IDEAS OF QUANTUM MECHANICS 1

READER'S GUIDE FOR COMPLEMENTS 33

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

B_I Constraints imposed by the uncertainty relations 39

C_I Heisenberg relation and atomic parameters 41

D_I An experiment illustrating the Heisenberg relations 45

E_I A simple treatment of a two-dimensional wave packet 49

F_I The relationship between one- and three-dimensional problems 53

G_I One-dimensional Gaussian wave packet: spreading of the wave packet 57

H_I Stationary states of a particle in one-dimensional square potentials 63

J_I Behavior of a wave packet at a potential step 75

K_I Exercises 83

****

II THE MATHEMATICAL TOOLS OF QUANTUM MECHANICS 87

READER'S GUIDE FOR COMPLEMENTS 159

A_{II} The Schwarz inequality 161

B_{II} Review of some useful properties of linear operators 163

C_{II} Unitary operators 173

D_{II} A more detailed study of the r and p representations 181

E_{II} Some general properties of two observables, Q and P, whose commutator is equal to $i\hbar$ 187

F_{II} The parity operator 193

G_{II} An application of the properties of the tensor product: the two-dimensional infinite well **201**

H_{II} Exercises **205**

III THE POSTULATES OF QUANTUM MECHANICS **213**

READER'S GUIDE FOR COMPLEMENTS **267**

A_{III} Particle in an infinite one-dimensional potential well **271**

B_{III} Study of the probability current in some special cases **283**

C_{III} Root mean square deviations of two conjugate observables **289**

D_{III} Measurements bearing on only one part of a physical system **293**

E_{III} The density operator **299**

F_{III} The evolution operator **313**

G_{III} The Schrödinger and Heisenberg pictures **317**

H_{III} Gauge invariance **321**

J_{III} Propagator for the Schrödinger equation **335**

K_{III} Unstable states. Lifetime **343**

L_{III} Exercises **347**

M_{III} Bound states in a "potential well" of arbitrary shape **359**

N_{III} Unbound states of a particle in the presence of a potential well or barrier **367**

O_{III} Quantum properties of a particle in a one-dimensional periodic structure **375**

****

IV APPLICATIONS OF THE POSTULATES TO SIMPLE CASES: SPIN 1/2 AND TWO-LEVEL SYSTEMS **393**

READER'S GUIDE FOR COMPLEMENTS **423**

A_{IV} The Pauli matrices **425**

B_{IV} Diagonalization of a 2×2 Hermitian matrix **429**

C_{IV} Fictitious spin 1/2 associated with a two-level system **435**

D_{IV} System of two spin 1/2 particles **441**

E_{IV} Spin $1/2$ density matrix **449**

F_{IV} Spin 1/2 particle in a static and a rotating magnetic fields: magnetic resonance **455**

G_{IV} A simple model of the ammonia molecule **469**

H_{IV} Effects of a coupling between a stable state and an unstable state **485**

J_{IV} Exercises **491**

****

V THE ONE-DIMENSIONAL HARMONIC OSCILLATOR **497**

READER'S GUIDE FOR COMPLEMENTS **525**

A_V Some examples of harmonic oscillators **527**

B_V Study of the stationary states in the x representation. Hermite polynomials **547**

C_V Solving the eigenvalue equation of the harmonic oscillator by the polynomial method **555**

D_V Study of the stationary states in the momentum representation **563**

E_V The isotropic three-dimensional harmonic oscillator **569**

F_V A charged harmonic oscillator in a uniform electric field **575**

G_V Coherent "quasi-classical" states of the harmonic oscillator **583**

H_V Normal vibrational modes of two coupled harmonic oscillators **599**

J_V Vibrational modes of an infinite linear chain of coupled harmonic oscillators; phonons **611**

K_V Vibrational modes of a continuous physical system. Photons **631**

L_V One-dimensional harmonic oscillator in thermodynamic equilibrium at a temperature T **647**

M_V Exercises **661**

****

VI GENERAL PROPERTIES OF ANGULAR MOMENTUM IN QUANTUM MECHANICS **667**

READER'S GUIDE FOR COMPLEMENTS 703

A_{VI} Spherical harmonics 705

B_{VI} Angular momentum and rotations 717

C_{VI} Rotation of diatomic molecules 739

D_{VI} Angular momentum of stationary states of a two-dimensional harmonic oscillator 755

E_{VI} A charged particle in a magnetic field: Landau levels 771

F_{VI} Exercises 795

****

VII PARTICLE IN A CENTRAL POTENTIAL, HYDROGEN ATOM 803

READER'S GUIDE FOR COMPLEMENTS 831

A_{VII} Hydrogen-like systems 833

B_{VII} A soluble example of a central potential: The isotropic three-dimensional harmonic oscillator 841

C_{VII} Probability currents associated with the stationary states of the hydrogen atom 851

D_{VII} The hydrogen atom placed in a uniform magnetic field. Paramagnetism and diamagnetism. The Zeeman effect 855

E_{VII} Some atomic orbitals. Hybrid orbitals 869

F_{VII} Vibrational-rotational levels of diatomic molecules 885

G_{VII} Exercises 899

INDEX 901

****

Volume II

VOLUME II **923**

Table of contents **v**

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

READER'S GUIDE FOR COMPLEMENTS **957**

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

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

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

****

IX ELECTRON SPIN **985**

READER'S GUIDE FOR COMPLEMENTS **999**

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

B_{IX} Exercises **1009**

****

X ADDITION OF ANGULAR MOMENTA **1015**

READER'S GUIDE FOR COMPLEMENTS **1041**

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

B_X Clebsch-Gordan coefficients **1051**

C_X Addition of spherical harmonics **1059**

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

E_X Electric multipole moments **1077**

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

G_X Exercises **1107**

****

XI STATIONARY PERTURBATION THEORY 1115

READER'S GUIDE FOR COMPLEMENTS 1129

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

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

C_{XI} Van der Waals forces 1151

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

E_{XI} The variational method 1169

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

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

H_{XI} Exercises 1221

****

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

READER'S GUIDE FOR COMPLEMENTS 1265

A_{XII} The magnetic hyperfine Hamiltonian 1267

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

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

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

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

****

XIII APPROXIMATION METHODS FOR TIME-DEPENDENT PROBLEMS 1303

READER'S GUIDE FOR COMPLEMENTS 1337

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

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

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

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

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

F_{XIII} Exercises 1409

****

XIV SYSTEMS OF IDENTICAL PARTICLES 1419

READER'S GUIDE FOR COMPLEMENTS 1457

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

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

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

D_{XIV} Exercises 1496

****

APPENDICES 1505

I Fourier series and Fourier transforms 1505

II The Dirac δ-"function" 1515

III Lagrangian and Hamiltonian in classical mechanics 1527

BIBLIOGRAPHY OF VOLUMES I AND II 1545

INDEX 1569

****

Volume III

VOLUME III **1591**

Table of contents **v**

XV CREATION AND ANNIHILATION OPERATORS FOR IDENTICAL PARTICLES **1591**
A General formalism . . . 1592
B One-particle symmetric operators . . . 1603
C Two-particle operators . . . 1608

READER'S GUIDE FOR COMPLEMENTS **1617**

A_{XV} Particles and holes **1621**
1 Ground state of a non-interacting fermion gas . . . 1621
2 New definition for the creation and annihilation operators . . . 1622
3 Vacuum excitations . . . 1623

B_{XV} Ideal gas in thermal equilibrium; quantum distribution functions **1625**
1 Grand canonical description of a system without interactions . . . 1626
2 Average values of symmetric one-particle operators . . . 1628
3 Two-particle operators . . . 1631
4 Total number of particles . . . 1635
5 Equation of state, pressure . . . 1640

C_{XV} Condensed boson system, Gross-Pitaevskii equation **1643**
1 Notation, variational ket . . . 1643
2 First approach . . . 1645
3 Generalization, Dirac notation . . . 1648
4 Physical discussion . . . 1651

D_{XV} Time-dependent Gross-Pitaevskii equation **1657**
1 Time evolution . . . 1657
2 Hydrodynamic analogy . . . 1664
3 Metastable currents, superfluidity . . . 1667

E_{XV} Fermion system, Hartree-Fock approximation **1677**
1 Foundation of the method . . . 1678
2 Generalization: operator method . . . 1688

F_{XV} Fermions, time-dependent Hartree-Fock approximation **1701**
1 Variational ket and notation . . . 1701
2 Variational method . . . 1702
3 Computing the optimizer . . . 1705
4 Equations of motion . . . 1707

G_{XV} Fermions or Bosons: Mean field thermal equilibrium **1711**
1 Variational principle . . . 1712
2 Approximation for the equilibrium density operator . . . 1716
3 Temperature dependent mean field equations . . . 1725

H_{XV} Applications of the mean field method for non-zero temperature **1733**
1 Hartree-Fock for non-zero temperature, a brief review . . . 1733
2 Homogeneous system . . . 1734
3 Spontaneous magnetism of repulsive fermions . . . 1737
4 Bosons: equation of state, attractive instability . . . 1745

****

XVI FIELD OPERATOR **1751**
A Definition of the field operator . . . 1752
B Symmetric operators . . . 1755
C Time evolution of the field operator (Heisenberg picture) . . . 1763
D Relation to field quantization . . . 1765

READER'S GUIDE FOR COMPLEMENTS **1767**

A_{XVI} Spatial correlations in an ideal gas of bosons or fermions **1769**
1 System in a Fock state . . . 1769
2 Fermions in the ground state . . . 1771
3 Bosons in a Fock state . . . 1775

B_{XVI} Spatio-temporal correlation functions, Green's functions **1781**
1 Green's functions in ordinary space . . . 1781
2 Fourier transforms . . . 1790
3 Spectral function, sum rule . . . 1795

C_{XVI} Wick's theorem **1799**
1 Demonstration of the theorem . . . 1799
2 Applications: correlation functions for an ideal gas . . . 1804

****

XVII PAIRED STATES OF IDENTICAL PARTICLES **1811**
A Creation and annihilation operators of a pair of particles . . . 1813
B Building paired states . . . 1818
C Properties of the kets characterizing the paired states . . . 1822
D Correlations between particles, pair wave function . . . 1830
E Paired states as a quasi-particle vacuum; Bogolubov-Valatin transformations 1836

READER'S GUIDE FOR COMPLEMENTS 1843

A_{XVII} Pair field operator for identical particles 1845
1 Pair creation and annihilation operators . . . 1846
2 Average values in a paired state . . . 1851
3 Commutation relations of field operators . . . 1861

B_{XVII} Average energy in a paired state 1869
1 Using states that are not eigenstates of the total particle number . . . 1869
2 Hamiltonian . . . 1871
3 Spin 1/2 fermions in a singlet state . . . 1874
4 Spinless bosons . . . 1881

C_{XVII} Fermion pairing, BCS theory 1889
1 Optimization of the energy . . . 1890
2 Distribution functions, correlations . . . 1899
3 Physical discussion . . . 1914
4 Excited states . . . 1919

D_{XVII} Cooper pairs 1927
1 Cooper model . . . 1927
2 State vector and Hamiltonian . . . 1927
3 Solution of the eigenvalue equation . . . 1929
4 Calculation of the binding energy for a simple case . . . 1929

E_{XVII} Condensed repulsive bosons 1933
1 Variational state, energy . . . 1935
2 Optimization . . . 1937
3 Properties of the ground state . . . 1940
4 Bogolubov operator method . . . 1950

****

XVIII REVIEW OF CLASSICAL ELECTRODYNAMICS 1957
A Classical electrodynamics . . . 1959
B Describing the transverse field as an ensemble of harmonic oscillators . . . 1968

READER'S GUIDE FOR COMPLEMENTS 1977

A_{XVIII} Lagrangian formulation of electrodynamics 1979
1 Lagrangian with several types of variables . . . 1980
2 Application to the free radiation field . . . 1986
3 Lagrangian of the global system field + interacting particles . . . 1992

****

XIX QUANTIZATION OF ELECTROMAGNETIC RADIATION 1997
A Quantization of the radiation in the Coulomb gauge 1999
B Photons, elementary excitations of the free quantum field 2004
C Description of the interactions 2009

READER'S GUIDE FOR COMPLEMENTS 2017

A_{XIX} Momentum exchange between atoms and photons 2019
1 Recoil of a free atom absorbing or emitting a photon 2020
2 Applications of the radiation pressure force: slowing and cooling atoms . . . 2025
3 Blocking recoil through spatial confinement 2036
4 Recoil suppression in certain multi-photon processes 2040

B_{XIX} Angular momentum of radiation 2043
1 Quantum average value of angular momentum for a spin 1 particle 2044
2 Angular momentum of free classical radiation as a function of normal variables 2047
3 Discussion 2050

C_{XIX} Angular momentum exchange between atoms and photons 2055
1 Transferring spin angular momentum to internal atomic variables 2056
2 Optical methods 2058
3 Transferring orbital angular momentum to external atomic variables 2065

****

XX ABSORPTION, EMISSION AND SCATTERING OF PHOTONS BY ATOMS 2067
A A basic tool: the evolution operator 2068
B Photon absorption between two discrete atomic levels 2073
C Stimulated and spontaneous emissions 2080
D Role of correlation functions in one-photon processes 2084
E Photon scattering by an atom 2085

READER'S GUIDE FOR COMPLEMENTS 2095

A_{XX} A multiphoton process: two-photon absorption 2097
1 Monochromatic radiation 2097
2 Non-monochromatic radiation 2101
3 Discussion 2105

B_{XX} Photoionization 2109
1 Brief review of the photoelectric effect 2110
2 Computation of photoionization rates 2112
3 Is a quantum treatment of radiation necessary to describe photoionization? . 2118
4 Two-photon photoionization 2123
5 Tunnel ionization by intense laser fields 2126

C_{XX} Two-level atom in a monochromatic field. Dressed-atom method 2129
1 Brief description of the dressed-atom method . . . 2130
2 Weak coupling domain . . . 2137
3 Strong coupling domain . . . 2141
4 Modifications of the field. Dispersion and absorption . . . 2147

D_{XX} Light shifts: a tool for manipulating atoms and fields 2151
1 Dipole forces and laser trapping . . . 2151
2 Mirrors for atoms . . . 2153
3 Optical lattices . . . 2153
4 Sub-Doppler cooling. Sisyphus effect . . . 2155
5 Non-destructive detection of a photon . . . 2159

E_{XX} Detection of one- or two-photon wave packets, interference 2163
1 One-photon wave packet, photodetection probability . . . 2165
2 One- or two-photon interference signals . . . 2167
3 Absorption amplitude of a photon by an atom . . . 2174
4 Scattering of a wave packet . . . 2176
5 Example of wave packets with two entangled photons . . . 2181

****

XXI QUANTUM ENTANGLEMENT, MEASUREMENTS, BELL'S INEQUALITIES 2187
A Introducing entanglement, goals of this chapter . . . 2188
B Entangled states of two spin-1/2 systems . . . 2190
C Entanglement between more general systems . . . 2193
D Ideal measurement and entangled states . . . 2196
E "Which path" experiment: can one determine the path followed by the photon in Young's double slit experiment? . . . 2202
F Entanglement, non-locality, Bell's theorem . . . 2204

READER'S GUIDE FOR COMPLEMENTS 2215

A_{XXI} Density operator and correlations; separability 2217
1 Von Neumann statistical entropy . . . 2217
2 Differences between classical and quantum correlations . . . 2221
3 Separability . . . 2223

B_{XXI} GHZ states, entanglement swapping 2227
1 Sign contradiction in a GHZ state . . . 2227
2 Entanglement swapping . . . 2232

C_{XXI} Measurement induced relative phase between two condensates 2237
1 Probabilities of single, double, etc. position measurements . . . 2239
2 Measurement induced enhancement of entanglement . . . 2242
3 Detection of a large number Q of particles . . . 2245

D_{XXI} Emergence of a relative phase with spin condensates; macroscopic non-locality and the EPR argument **2253**
1 Two condensates with spins 2254
2 Probabilities of the different measurement results 2255
3 Discussion 2259

****

APPENDICES **2267**

IV Feynman path integral **2267**
1 Quantum propagator of a particle 2267
2 Interpretation in terms of classical histories 2272
3 Discussion; a new quantization rule 2274
4 Operators 2276

V Lagrange multipliers **2281**
1 Function of two variables 2281
2 Function of N variables 2283

VI Brief review of Quantum Statistical Mechanics **2285**
1 Statistical ensembles 2285
2 Intensive or extensive physical quantities 2292

VII Wigner transform **2297**
1 Delta function of an operator 2299
2 Wigner distribution of the density operator (spinless particle) 2299
3 Wigner transform of an operator 2310
4 Generalizations 2318
5 Discussion: Wigner distribution and quantum effects 2319

BIBLIOGRAPHY OF VOLUME III **2325**

INDEX **2333**

Chapter XV

Creation and annihilation operators for identical particles

A General formalism 1592
A-1 Fock states and Fock space 1593
A-2 Creation operators $a^\dagger$ 1596
A-3 Annihilation operators a 1597
A-4 Occupation number operators (bosons and fermions) 1598
A-5 Commutation and anticommutation relations 1599
A-6 Change of basis 1601
B One-particle symmetric operators 1603
B-1 Definition 1603
B-2 Expression in terms of the operators a and $a^\dagger$ 1604
B-3 Examples 1606
B-4 Single particle density operator 1607
C Two-particle operators 1608
C-1 Definition 1608
C-2 A simple case: factorization 1609
C-3 General case 1610
C-4 Two-particle reduced density operator 1610
C-5 Physical discussion; consequences of the exchange 1611

Introduction

For a system composed of identical particles, the particle numbering used in Chapter XIV, the last chapter of Volume II [2], does not really have much physical significance. Furthermore, when the particle number gets larger than a few units, applying the symmetrization postulate to numbered particles often leads to complex calculations. For

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.
© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

example, computing the average value of a symmetric operator requires the symmetrization of the bra, the ket, and finally the operator, which introduces a large number of terms[1]. They seem different, a priori, but at the end of the computation many are found to be equal, or sometimes cancel each other. Fortunately, these lengthy calculations may be avoided using an equivalent method based on creation and annihilation operators in a "Fock space". The simple commutation (or anticommutation) rules satisfied by these operators are the expression of the symmetrization (or antisymmetrization) postulate. The non-physical particle numbering is replaced by assigning "occupation numbers" to individual states, which is more natural for treating identical particles.

The method described in this chapter and the following is sometimes called "second quantization"[2]. It deals with operators that no longer conserve the particle number, hence acting in a state space larger than those we have previously considered; this new space is called the "Fock space" (§ A). These operators which change the particle number appear mainly in the course of calculations, and often regroup at the end, keeping the total particle number constant. Examples will be given (§ B) for one-particle symmetric operators, such as the total linear momentum or angular momentum of a system of identical particles. We shall then study two-particle symmetric operators (§ C), such as the energy of a system of interacting identical particles, their spatial correlation function, etc. In quantum statistical mechanics, the Fock space is well adapted to computations performed in the "grand canonical" ensemble, where the total number of particles may fluctuate since the system is in contact with an external reservoir. Furthermore, as we shall see in the following chapters, the Fock space is very useful for describing physical processes where the particle number changes, as in photon absorption or emission.

A. General formalism

We denote $\mathcal{E}_N$ the state space of a system of N distinguishable particles, which is the tensor product of N individual state spaces $\mathcal{E}_1$:

$$\mathcal{E}_N = \mathcal{E}_1(1) \otimes \mathcal{E}_1(2) \otimes .. \otimes \mathcal{E}_1(N) \tag{A-1}$$

Two sub-spaces of $\mathcal{E}_N$ are particularly important for identical particles, as they contain all their accessible physical states: the space $\mathcal{E}_S(N)$ of the completely symmetric states for bosons, and the space $\mathcal{E}_A(N)$ of the completely antisymmetric states for fermions. The projectors onto these two sub-spaces are given by relations (B-49) and (B-50) of Chapter XIV:

$$S_N = \frac{1}{N!} \sum_\alpha P_\alpha \tag{A-2}$$

and:

$$A_N = \frac{1}{N!} \sum_\alpha \varepsilon_\alpha P_\alpha \tag{A-3}$$

[1] For a one-particle symmetric operator, which includes the sum of N terms, both the ket and bra contain $N!$ terms. The matrix element will therefore involve $N\,(N!)^2$ terms, a very large number once N exceeds a few units.

[2] A commonly accepted but a somewhat illogical expression, since no new quantification comes in addition to that of the usual postulates of Quantum Mechanics; its essential ingredient is the symmetrization of identical particles.

where the P_α are the $N!$ permutation operators for the N particles, and ε_α the parity of P_α (in this chapter we have added for clarity the index N to the projectors S and A defined in Chapter XIV).

A-1. Fock states and Fock space

Starting from an arbitrary orthonormal basis $\{|u_k\rangle\}$ of the state space for one particle, we constructed in § C-3-d of Chapter XIV a basis of the state space for N identical particles. Its vectors are characterized by the occupation numbers n_i, with:

$$n_1 + n_2 + .. + n_k + .. = N \tag{A-4}$$

where n_1 is the occupation number of the first basis vector $|u_1\rangle$ (i.e. the number of particles in $|u_1\rangle$), n_2 that of $|u_2\rangle$, ..,n_k that of $|u_k\rangle$. In this series of numbers, some (even many) may be zero: a given state has no particular reason to always be occupied. It is therefore often easier to specify only the non-zero occupation numbers, which will be noted $n_i, n_j, .., n_l, ..$. This series indicates that the first basis state that has at least one particle is $|u_i\rangle$ and it contains n_i particles; the second occupied state is $|u_j\rangle$ with a population n_j, etc. As in (A-4), these occupation numbers add up to N.

Comment:

In this chapter we constantly use subscripts of different types, which should not be confused. The subscripts i, j, k, l, ..denote different basis vectors $\{|u_i\rangle\}$ of the state space $\mathcal{E}_1$ of a single particle; they span values given by the dimension of this state space, which often goes to infinity. They should not be confused with the subscripts used to number the particles, which can take N different values, and are labeled q, q', etc. Finally the subscript α distinguishes the different permutations of the N particles, and can therefore take $N!$ different values.

A-1-a. Fock states for identical bosons

For bosons, the basis vectors can be written as in (C-15) of Chapter XIV:

$$\begin{aligned} &|n_i, n_j, .., n_l, ..\rangle \\ &\quad = c\ S_N\, |1 : u_i;\ 2 : u_i;\ ..;\ n_i : u_i;\ \ n_i + 1 : u_j;\ ..\ \ n_i + n_j : u_j;\ ..\rangle \end{aligned} \tag{A-5}$$

where c is a normalization constant; on the right-hand side, n_i particles occupy the state $|u_i\rangle$, n_j the state $|u_j\rangle$, etc... (because of symmetrization, their order does not matter).

Let us calculate the norm of the right-hand term. It is composed of $N!$ terms, coming from each of the $N!$ permutations included in S_N, but only some of them are orthogonal to each other: all the permutations P_α leading to redistributions of the n_i first particles among themselves, of the next n_j particles among themselves, etc. yield the same initial ket. On the other hand, if a permutation changes the individual state of one (or more than one) particle, it yields a different ket, actually orthogonal to the initial ket. This means that the different permutations contained in S_N can be grouped into families of $n_i!n_j!..n_l!..$ equivalent permutations, all yielding the same ket; taking into account the factor $N!$ appearing in the definition of S_N, the coefficient in front of this ket becomes $n_i!n_j!..n_l!../N!$ and its contribution to the norm of the ket is equal to the square of this number. On the other hand, the number of orthogonal kets is

$N!/n_i!n_j!..n_l!..$ Consequently if c was equal to 1 in formula (A-5), the ket thus defined would have a norm equal to:

$$\frac{N!}{n_i!n_j!..n_l!..}\left[\frac{1}{N!}n_i!n_j!..n_l!..\right]^2 = \frac{n_i!n_j!..n_l!..}{N!} \tag{A-6}$$

We shall therefore choose for c the inverse of the square root of that number, leading to the normalized ket:

$$|n_i,\ n_j,..,n_l,..\rangle = \sqrt{\frac{N!}{n_i!n_j!..n_l!..}}\ S_N\,|1:u_i;\ 2:u_i;\ ..;\ n_i:u_i;\ \ n_i+1:u_j;..;\ n_i+n_j:u_j;\ ..\rangle \tag{A-7}$$

These states are called the "*Fock states*", for which the occupation numbers are well defined.

For the Fock states, it is sometimes handy to use a slightly different but equivalent notation. In (A-7), these states are defined by specifying the occupation numbers of all the states that are actually occupied ($n_i \geq 1$). Another option would be to indicate all the occupation numbers including those which are zero [3] – this is what we have explicitly done in (A-4). We then write the same kets as:

$$|n_1, n_2, .., n_i, .., n_l, ..\rangle \tag{A-8}$$

Another possibility is to specify a list of N occupied states, where u_i is repeated n_i times, u_j repeated n_j times, etc. :

$$|\underbrace{u_i, u_i,\ \ldots}_{n_i\text{-times}}, \underbrace{u_j, u_j,\ \ldots}_{n_j\text{-times}},\ \ldots\ , u_l,\ \ldots\rangle \tag{A-9}$$

As we shall see later, this latter notation is sometimes useful in computations involving both bosons and fermions.

A-1-b. Fock states for identical fermions

In the case of fermions, the operator A_N acting on a ket where two (or more) numbered particles are in the same individual state yields a zero result: there are no such states in the physical space $\mathcal{E}_A(N)$. Hence we concentrate on the case where all the occupation numbers are either 1 or 0. We denote $|u_i\rangle$, $|u_j\rangle$,..,$|u_l\rangle$, .. all the states having an occupation number equal to 1. The equivalent for fermions of formula (A-7) is written:

$$|u_i, u_j, .., u_l, ..\rangle = \begin{cases} \sqrt{N!}\ A_N\,|1:u_i;\ 2:u_j;\ ..;\ q:u_l;\ ..\rangle & \text{if all the } u_i \text{ are different} \\ 0 & \text{if two } u_i \text{ are identical} \end{cases} \tag{A-10}$$

[3] Remember that, by convention, $0! = 1$.

Taking into account the $1/N!$ factor appearing in definition (A-3) of A_N, the right-hand side of this equation is a linear superposition, with coefficients $1/\sqrt{N!}$, of $N!$ kets which are all orthogonal to each other (as we have chosen an orthonormal basis for the individual states $\{|u_k\rangle\}$); hence its norm is equal to 1. Consequently, Fock states for fermions are defined by (A-10). Contrary to bosons, the main concern is no longer how many particles occupy a state, but whether a state is occupied or not. Another difference with the boson case is that, for fermions, the order of the states matters. If for instance the first two states u_i and u_j are exchanged, we get the opposite ket:

$$|u_j, u_i, .., u_l, ..\rangle = -\,|u_i, u_j, .., u_l, ..\rangle \tag{A-11}$$

but it obviously does not change the physical meaning of the ket.

A-1-c. Fock space

The Fock states are the building blocks used to construct this whole chapter. We have until now considered separately the spaces $\mathcal{E}_{S,A}(N)$ associated with different values of the particle number N. We shall now regroup them into a single space, called the "*Fock space* ", using the direct sum[4] formalism. For bosons:

$$\mathcal{E}^S_{\rm Fock} = \mathcal{E}_S(0) \oplus \mathcal{E}_S(1) \oplus \mathcal{E}_S(2) \oplus .. \oplus \mathcal{E}_S(N) \oplus .. \tag{A-12}$$

and, for fermions:

$$\mathcal{E}^A_{\rm Fock} = \mathcal{E}_A(0) \oplus \mathcal{E}_A(1) \oplus \mathcal{E}_A(2) \oplus .. \oplus \mathcal{E}_A(N) \oplus .. \tag{A-13}$$

(the sums go to infinity). In both cases, we have included on the right-hand side a first term associated with a total number of particles equal to zero. The corresponding space, $\mathcal{E}_{S,A}(0)$, is defined as a one-dimensional space, containing a single state called "vacuum" and denoted $|0\rangle$ or $|\text{vac}\rangle$. For bosons as well as fermions, an orthonormal basis for the Fock space can be built with the Fock states $|n_1,\ n_2, ..., n_i, ., n_l..\rangle$, relaxing the constraint (A-4): the occupation numbers may then take on any (integer) values, including zeros for all, which corresponds to the vacuum ket $|0\rangle$. Linear combinations of all these basis vectors yield all the vectors of the Fock space, including linear superpositions of kets containing different particle numbers. It is not essential to attribute a physical interpretation to such superpositions since they can be considered as intermediate states of the calculation. Obviously, the Fock space contains many kets with well defined particle numbers: all those belonging to a single sub-space $\mathcal{E}_S(N)$ for bosons, or $\mathcal{E}_A(N)$ for fermions. Two kets having different particle numbers N are necessarily orthogonal; for example, all the kets having a non-zero total population are orthogonal to the vacuum state.

Comments:

(i) Contrary to the distinguishable particle case, the Fock space is not the tensor product of the spaces of states associated with particles numbered 1, 2,..., q, etc. First of all, for a

[4]The direct sum of two spaces $\mathcal{E}_P$ (with dimension P) and $\mathcal{E}_Q$ (with dimension Q) is a space $\mathcal{E}_{P+Q}$ with dimension $P+Q$, spanned by all the linear combinations of a vector from the first space with a vector from the second. A basis for $\mathcal{E}_{P+Q}$ may be simply obtained by grouping together a basis for $\mathcal{E}_P$ and one for $\mathcal{E}_Q$. For example, vectors of a two-dimensional plane belong to a space that is the direct sum of the one-dimensional spaces for the vectors of two axes of that plane.

fixed N, it only includes the totally symmetric (or antisymmetric) subspace of this tensor product; furthermore, the Fock space is the direct sum of such subspaces associated with each value of the particle number N.

The Fock space is, however, the tensor product of Fock spaces $\mathcal{E}_{\rm Fock}^{u_i}$ associated with the individual orthogonal states $|u_i\rangle$, each $\mathcal{E}_{\rm Fock}^{u_i}$ being spanned by the kets $|n_i\rangle$ where n_i takes on all integer values (from zero to infinity for bosons, from zero to one for fermions):

$$\mathcal{E}_{\rm Fock}^{S,A} = \mathcal{E}_{\rm Fock}^{u_1} \otimes \mathcal{E}_{\rm Fock}^{u_2} \otimes .. \otimes \mathcal{E}_{\rm Fock}^{ui} \otimes .. \tag{A-14}$$

This is because the Fock states, which are a basis for $\mathcal{E}_{\rm Fock}^{S,A}$, may be written as the tensor product:

$$|n_1,\ n_2, .., n_i, ..\rangle = |n_1\rangle \otimes |n_2\rangle \otimes .. \otimes |n_i\rangle \otimes .. \tag{A-15}$$

It is often said that each individual state defines a "mode" of the system of identical particles. Decomposing the Fock state into a tensor product allows considering the modes as describing different and distinguishable variables. This will be useful on numerous occasions (see for example Complements $\rm B_{XV}$, $\rm D_{XV}$ and $\rm E_{XV}$).

(ii) One should not confuse a Fock state with an arbitrary state of the Fock space. The occupation numbers of individual states are all well defined in a Fock state (also called "number state"), whereas an arbitrary state of the Fock space is a linear superposition of these eigenstates, with several non-zero coefficients.

A-2. Creation operators $a^\dagger$

Choosing a basis of individual states $\{|u_i\rangle\}$, we now define the action in the Fock space of the creation operator[5] $a_{u_i}^\dagger$ on a particle in the state $|u_i\rangle$.

A-2-a. Bosons

For bosons, we introduce the linear operator $a_{u_i}^\dagger$ defined by:

$$a_{u_i}^\dagger\ |n_1, n_2, .., n_i, ..\rangle = \sqrt{n_i + 1}\ |n_1, n_2, .., n_i + 1, ..\rangle \tag{A-16}$$

As all the states of the Fock space may be obtained by a linear superposition of $|n_1,\ n_2, .., n_i, ..\rangle$, the action of $a_{u_i}^\dagger$ is defined in the entire space. It adds a particle to the system, which goes from a state of $\mathcal{E}_S(N)$ to a state of $\mathcal{E}_S(N+1)$, and in particular from the vacuum to a state having one single occupied state.

Creation operators acting on the vacuum allow building occupied states. Recurrent application of (A-16), leads to:

$$|n_1, n_2, .., n_i, ..\rangle = \frac{1}{\sqrt{n_1! n_2! .. n_i! ..}} \left[a_{u_1}^\dagger\right]^{n_1} \left[a_{u_2}^\dagger\right]^{n_2} .. \left[a_{u_i}^\dagger\right]^{n_i} .. |0\rangle \tag{A-17}$$

Comment:

Why was the factor $\sqrt{n_i + 1}$ introduced in (A-16)? We shall see later (§ B) that, together with the factors of (A-7), it simplifies the computations.

[5] A similar notation was used for the harmonic oscillator.

A-2-b. Fermions

For fermions, we define the operator $a^\dagger_{u_i}$ by:

$$a^\dagger_{u_i} \ |u_j, .., u_k, .., u_l, ..\rangle = |u_i, u_j, .., u_k, .., u_l, ..\rangle \tag{A-18}$$

where the newly created state $|u_i\rangle$ appears first in the list of states in the ket on the right-hand side. If we start from a ket where the individual state $|u_i\rangle$ is already occupied ($n_i = 1$), the action of $a^\dagger_{u_i}$ leads to zero, as in this case (A-10) gives:

$$a^\dagger_{u_i} \ |u_j, .., u_i, .., u_l, ..\rangle = |u_i, u_j, .., u_i, .., u_l, ..\rangle = 0 \tag{A-19}$$

Formulas (A-16) and (A-17) are also valid for fermions, with all the occupation numbers equal to 0 or 1(or else both members are zero).

Comment:

Definition (A-18) must not depend on the specific order of the individual states $u_j, .., u_k, .., u_l, ..$ in the ket on which the operator $a^\dagger_{u_i}$ acts. It can be easily verified that any permutation of the states simply multiply by its parity both members of the equality. It therefore remains valid independently of the order chosen for the individual states in the initial ket.

A-3. Annihilation operators a

We now study the Hermitian conjugate operator of $a^\dagger_{u_i}$, that we shall simply call a_{u_i} since taking twice in a row the Hermitian conjugate of an operator brings you back to the initial operator.

A-3-a. Bosons

For bosons, we deduce from (A-16) that the only non-zero matrix elements of $a^\dagger_{u_i}$ in the Fock states orthonormal basis are:

$$\langle n_1, n_2, .., n_i + 1, ..| \, a^\dagger_{u_i} \, |n_1, n_2, .., n_i, ..\rangle = \sqrt{n_i + 1} \tag{A-20}$$

They link two vectors having equal occupation numbers except for n_i, which increases by one going from the ket to the bra.

The matrix elements of the Hermitian conjugate of $a^\dagger_{u_i}$ are obtained from relation (A-20), using the general definition (B-49) of Chapter II. The only non-zero matrix elements of a_{u_i} are thus:

$$\langle n_1, n_2, .., n_i, ..| \, a_{u_i} \, |n_1, n_2, .., n_i + 1, ..\rangle = \sqrt{n_i + 1} \tag{A-21}$$

Since the basis we use is complete, we can deduce the action of the operators a_i on kets having given occupation numbers:

$$a_{u_i} \ |n_1, n_2, .., n_i, ..\rangle = \sqrt{n_i} \, |n_1, n_2, .., n_i - 1, ..\rangle \tag{A-22}$$

(note that we have replaced n_i by $n_i - 1$). As opposed to $a^\dagger_{u_i}$, which adds a particle in the state $|u_i\rangle$, the operator a_{u_i} takes one away; it yields zero when applied on a ket where the state $|u_i\rangle$ is empty to begin with, such as the vacuum state:

$$a_{u_i} \, |0\rangle = 0 \tag{A-23}$$

We call a_{u_i} "*the annihilation operator*" for the state $|u_i\rangle$.

A-3-b. Fermions

For fermions, relation (A-18) allows writing the matrix elements:

$$\langle u_i, u_j, .., u_k, ..u_l, ..|\, a^\dagger_{u_i}\, |u_j, .., u_k, .., u_l, ..\rangle = 1 \tag{A-24}$$

The only non-zero elements are those where all the individual occupied states are left unchanged in the bra and the ket, except for the state u_i only present in the bra, but not in the ket. As for the occupation numbers, none change, except for n_i which goes from 0 (in the ket) to 1 (in the bra).

The Hermitian conjugation operation then yields the action of the corresponding annihilation operator:

$$a_{u_i}\ \ |u_i, u_j, u_k, .., u_l, ..\rangle = |u_j, u_k, .., u_l, ..\rangle \tag{A-25}$$

or, if initially the state $|u_i\rangle$ is not occupied:

$$a_{u_i}\ \ |u_j, u_k, .., u_l, ..\rangle = 0 \tag{A-26}$$

Relations (A-22) and (A-23) are also valid for fermions, with the usual condition that all occupation numbers should be equal to 0 or 1; otherwise, the relations amount to $0 = 0$.

Comment:

To use relation (A-25) when the state $|u_i\rangle$ is already occupied but not listed in the first position, we first have to bring it there; if it requires an odd permutation, a change of sign will occur. For example:

$$a_{u_2}\, |u_1, u_2\rangle = -\, |u_1\rangle \tag{A-27}$$

For fermions, the operators a and $a^\dagger$ therefore act on the individual state that is listed in the first position in the N-particle ket; a destroys the first state in the list, and $a^\dagger$ creates a new state placed at the beginning of the list. Forgetting this could lead to errors in sign.

A-4. Occupation number operators (bosons and fermions)

Consider the operator $\widehat{n}_{u_i}$ defined by:

$$\widehat{n}_{u_i} = a^\dagger_{u_i} a_{u_i} \tag{A-28}$$

and its action on a Fock state. For bosons, if we apply successively formulas (A-22) and (A-16), we see that this operator yields the same Fock state, but multiplied by its occupation number n_i. For fermions, if $|u_i\rangle$ is empty in the Fock state, relation (A-26) shows that the action of the operator $\widehat{n}_{u_i}$ yields zero. If the state $|u_i\rangle$ is already occupied, we must first permute the states to bring $|u_i\rangle$ to the first position, which may eventually change the sign in front of the Fock space ket. The successive application on this ket of (A-25) and (A-19) shows that the action of the operator $\widehat{n}_{u_i}$ leaves this ket unchanged; we then move the state $|u_i\rangle$ back to its initial position, which may introduce a second change in sign, canceling the first one. We finally obtain for fermions the same result as for bosons, except that the n_i can only take the values 1 and 0. In both cases the Fock states are the eigenvectors of the operator $\widehat{n}_{u_i}$ with the occupation numbers as

eigenvalues; consequently, this operator is named the "*occupation number operator of the state* $|u_i\rangle$". The operator $\widehat{N}$ associated with the total number of particles is simply the sum:

$$\widehat{N} = \sum_i \widehat{n}_{u_i} = \sum_i a^\dagger_{u_i} a_{u_i} \tag{A-29}$$

A-5. Commutation and anticommutation relations

Creation and annihilation operators have very simple commutation (for the bosons) and anticommutation (for the fermions) properties, which make them easy tools for taking into account the symmetrization or antisymmetrization of the state vectors.

To simplify the notation, each time the equations refer to a single basis of individual states $|u_i\rangle$, we shall write a_i instead of a_{u_i}. If, however, it can lead to ambiguity, we will return to the full notation.

A-5-a. Bosons: commutation relations

Consider, for bosons, the two operators $a^\dagger_i$ and $a^\dagger_j$. If both subscripts i and j are different, they correspond to orthogonal states $|u_i\rangle$ and $|u_j\rangle$. Using twice (A-16) then yields:

$$a^\dagger_i a^\dagger_j \,|n_1, n_2, .., n_i, .., n_j, ..\rangle = \sqrt{n_i + 1}\sqrt{n_j + 1}\;\; |n_1, n_2, .., n_i + 1, .., n_j + 1, ..\rangle \tag{A-30}$$

Changing the order of the operators yields the same result. As the Fock states form a basis, we can deduce that the commutator of $a^\dagger_i$ and $a^\dagger_j$ is zero if $i \neq j$. In the same way, it is easy to show that both operator products $a_i a_j$ and $a_j a_i$ acting on the same ket yield the same result (a ket having two occupation numbers lowered by 1); a_i and a_j thus commute if $i \neq j$. Finally the same procedure allows showing that a_i and $a^\dagger_j$ commute if $i \neq j$. Now, if $i = j$, we must evaluate the commutator of a_i and $a^\dagger_i$. Let us apply (A-16) and (A-22) successively, first in that order, and then in the reverse order:

$$\begin{aligned} a_i a^\dagger_i \,|n_1, n_2, .., n_i, ..\rangle &= (n_i + 1)\,|n_1, n_2, .., n_i, ..\rangle \\ a^\dagger_i a_i \,|n_1, n_2, .., n_i, ..\rangle &= (n_i)\,|n_1, n_2, .., n_i, ..\rangle \end{aligned} \tag{A-31}$$

The commutator of a_i and $a^\dagger_i$ is therefore equal to 1 for all the values of the subscript i. All the previous results are summarized in three equalities valid for bosons:

$$[a_i, a_j] = 0 \qquad \left[a^\dagger_i, a^\dagger_j\right] = 0 \qquad \left[a_i, a^\dagger_j\right] = \delta_{ij} \tag{A-32}$$

A-5-b. Fermions: anticommutation relations

For fermions, let us first assume that the subscripts i and j are different. The successive action of $a^\dagger_i$ and $a^\dagger_j$ on an occupation number ket only yields a non-zero ket if $n_i = n_j = 0$; using twice (A-18) leads to:

$$a^\dagger_i a^\dagger_j \;\; |u_k, u_l .., u_m, ..\rangle = |u_i, u_j, u_k, u_l, .., u_m, ..\rangle \tag{A-33}$$

but, if we change the order:

$$a^\dagger_j a^\dagger_i \;\; |u_k, u_l, .., u_m, ..\rangle = |u_j, u_i, u_k, u_l, .., u_m, ..\rangle = -\,|u_i, u_j, u_k, u_l, .., u_m, ..\rangle \tag{A-34}$$

Consequently the sign change that goes with the permutation of the two individual states leads to:

$$a_i^\dagger a_j^\dagger = -a_j^\dagger a_i^\dagger \qquad \text{if } i \neq j \tag{A-35}$$

If we define the anticommutator $[A, B]_+$ of two operators A and B by:

$$[A, B]_+ = AB + BA \tag{A-36}$$

(A-35) may be written as:

$$\left[a_i^\dagger, a_j^\dagger\right]_+ = 0 \qquad \text{if } i \neq j \tag{A-37}$$

Taking the Hermitian conjugate of (A-35), we get:

$$a_j a_i = -a_i a_j \qquad \text{if } i \neq j \tag{A-38}$$

which can be written as:

$$[a_i, a_j]_+ = 0 \qquad \text{if } i \neq j \tag{A-39}$$

Finally, we show by the same method that the anticommutator of a_i and $a_j^\dagger$ is zero except when it acts on a ket where $n_i = 1$ and $n_j = 0$; those two occupation numbers are then interchanged. The computation goes as follows:

$$a_j^\dagger a_i \;\; |u_i, u_l, ..u_m, ..\rangle = a_j^\dagger \, |u_l, ..u_m, ..\rangle = |u_j, u_l, ..u_m, ..\rangle \tag{A-40}$$

and:

$$\begin{aligned} a_i a_j^\dagger \, |u_i, u_l, .., u_m, ..\rangle &= a_i \, |u_j, u_i, u_l, .., u_m, ..\rangle \\ &= -a_i \, |u_i, u_j, u_l, ...u_m, ..\rangle = - \, |u_j, u_l, .., u_m, ..\rangle \end{aligned} \tag{A-41}$$

Adding those two equations yields zero, hence proving that the anticommutator is zero:

$$\left[a_i, a_j^\dagger\right]_+ = 0 \qquad \text{si } i \neq j \tag{A-42}$$

In the case where $i = j$, the limitation on the occupation numbers (0 or 1) leads to:

$$\left[a_i\right]^2 = 0 \qquad \text{and} \qquad \left[a_i^\dagger\right]^2 = 0 \tag{A-43}$$

Equalities (A-37) and (A-39) are still valid if i and j are equal. We are now left with the computation of the anticommutator of a_i and $a_i^\dagger$. Let us first examine the product $a_i a_i^\dagger$; it yields zero if applied to a ket having an occupation number $n_i = 1$, but leaves unchanged any ket with $n_i = 0$, since the particle created by $a_i^\dagger$ is then annihilated by a_i. We get the inverse result for the product $a_i^\dagger a_i$ where the order has been inverted: it yields zero if $n_i = 0$, and leaves the ket unchanged if $n_i = 1$. Finally, whatever the occupation number ket is, one of the terms of the anticommutator yields zero, the other 1, and the net result is always 1. Therefore:

$$\left[a_i, a_i^\dagger\right]_+ = 1 \tag{A-44}$$

All the previous results valid for fermions are summarized in the following three relations, which are for fermions the equivalent of relations (A-32) for bosons:

$$[a_i, a_j]_+ = 0 \qquad \left[a_i^\dagger, a_j^\dagger\right]_+ = 0 \qquad \left[a_i, a_j^\dagger\right]_+ = \delta_{ij} \tag{A-45}$$

A-5-c. Common relations for bosons and fermions

To regroup the results valid for bosons and fermions in common relations, we introduce the notation:

$$[A,B]_{-\eta} = AB - \eta\ BA \tag{A-46}$$

with:

$$\begin{array}{ll} \eta = & 1 \quad \text{for bosons} \\ \eta = & -1 \quad \text{for fermions} \end{array} \tag{A-47}$$

so that (A-46) is the commutator of A and B for bosons, and their anticommutator for fermions. We then have:

$$\begin{array}{l} \left[a_i, a_j\right]_{-\eta} = 0 \quad \text{for all } i \text{ and } j \text{ (equal or different)} \\ \left[a_i^\dagger, a_j^\dagger\right]_{-\eta} = 0 \quad \text{for all } i \text{ and } j \text{ (equal or different)} \end{array} \tag{A-48}$$

and the only non-zero combinations are:

$$\left[a_i, a_j^\dagger\right]_{-\eta} = \delta_{ij} \qquad\qquad \left[a_i^\dagger, a_j\right]_{-\eta} = -\eta\delta_{ij} \tag{A-49}$$

A-6. Change of basis

What are the effects on the creation and annihilation operators of a change of basis for the individual states? The operators $a_{u_i}^\dagger$ and a_{u_i} have been introduced by their action on the Fock states, defined by relations (A-7) and (A-10) for which a given basis of individual states $\{|u_i\rangle\}$ was chosen. One could also choose any another orthonormal basis $\{|v_s\rangle\}$ and define in the same way bases for the Fock state and creation $a_{v_s}^\dagger$ and annihilation a_{v_s} operators. What is the relation between these new operators and the ones we defined earlier with the initial basis?

For creation operators acting on the vacuum state $|0\rangle$, the answer is quite straightforward: the action of $a_{v_s}^\dagger$ on $|0\rangle$ yields a one-particle ket, which can be written as:

$$a_{v_s}^\dagger |0\rangle = |1 : v_s\rangle = \sum_i \langle u_i|\ v_s\rangle\ |1 : u_i\rangle = \sum_i \langle u_i|\ v_s\rangle\ a_{u_i}^\dagger |0\rangle \tag{A-50}$$

This result leads us to expect a simple linear relation of the type:

$$a_{v_s}^\dagger = \sum_i \langle u_i|\ v_s\rangle\ \ a_{u_i}^\dagger \tag{A-51}$$

with its Hermitian conjugate:

$$a_{v_s} = \sum_i \langle v_s|\ u_i\rangle\ \ a_{u_i} \tag{A-52}$$

Equation (A-51) implies that creation operators are transformed by the same unitary relation as the individual states. Commutation or anticommutation relations are then conserved, since:

$$\left[a_{v_s}, a_{v_t}^\dagger\right]_{-\eta} = \sum_{i,j} \langle v_s|\ u_i\rangle\ \langle u_j|\ v_t\rangle\ \left[a_{u_i}, a_{u_j}^\dagger\right]_{-\eta} = \sum_{i,j} \langle v_s|\ u_i\rangle\ \langle u_j|\ v_t\rangle\ \delta_{ij} \tag{A-53}$$

which amounts to (as expected):

$$\left[a_{v_s}, a^\dagger_{v_t}\right]_{-\eta} = \sum_i \langle v_s|\ u_i\rangle \langle u_i|\ v_t\rangle = \delta_{st} \tag{A-54}$$

Furthermore, it is straightforward to show that the creation operators commute (or anticommute), as do the annihilation operators.

Equivalence of the two bases

We have not yet shown the complete equivalence of the two bases, which can be done following two different approaches. In the first one, we use (A-51) and (A-52) to define the creation and annihilation operators in the new basis. The associated Fock states are defined by replacing the $a^\dagger_{u_i}$ by the $a^\dagger_{v_s}$ in relations (A-17) for the bosons, and (A-18) for the fermions. We then have to show that these new Fock states are still related to the states with numbered particles as in (A-18) for bosons, and (A-10) for fermions. This will establish the complete equivalence of the two bases.

We shall follow a second approach where the two bases are treated completely symmetrically. Replacing in relations (A-7) and (A-10) the u_i by the v_s, we construct the new Fock basis. We next define the operators $a^\dagger_{v_s}$ by transposing relations (A-17) and (A-18) to the new basis. We then must verify that these operators obey relation (A-51), without limiting ourselves, as in (A-50), to their action on the vacuum state.

(i) Bosons

Relations (A-7) and (A-17) lead to:

$$\begin{aligned}&\left[a^\dagger_{u_i}\right]^{n_i}\left[a^\dagger_{u_j}\right]^{n_j} .. |0\rangle \\ &= \sqrt{N!}\ S_N\ \ |1:u_i;\ 2:u_i;\ ..;\ n_i:u_i;\ \ n_i+1:u_j;..;\ n_i+n_j:u_j;\ ..\rangle\end{aligned} \tag{A-55}$$

where, on the right-hand side, the n_i first particles occupy the same individual state u_i, the following n_j particles, numbered from n_i+1 to n_i+n_j, the individual state u_j, etc. The equivalent relation in the second basis can be written:

$$\begin{aligned}&\left[a^\dagger_{v_s}\right]^{p_s}\left[a^\dagger_{v_t}\right]^{p_t} .. |0\rangle \\ &= \sqrt{N!}\ S_N\ \ |1:v_s;\ 2:v_s;\ ..;\ p_s:v_s;\ \ p_s+1:v_t;..;\ p_s+p_t:v_t;\ ..\rangle\end{aligned} \tag{A-56}$$

with:

$$n_i + n_j + .. = p_s + p_t + .. = N \tag{A-57}$$

Replacing on the right-hand side of (A-56), the first ket $|v_s\rangle$ by:

$$|v_s\rangle = \sum_i \langle u_i|\ v_s\rangle\, |u_i\rangle \tag{A-58}$$

we obtain:

$$\sum_i \langle u_i|\ v_s\rangle\, \sqrt{N!}\ S_N\, |1:u_i;\ 2:v_s;\ ..;\ p_s:v_s;\ \ p_s+1:v_t;..;\ p_s+p_t:v_t;\ ..\rangle \tag{A-59}$$

Following the same procedure for all the basis vectors of the right-hand side, we can replace it by:

$$\sum_i \langle u_i|\ v_s\rangle \sum_{i'} \langle u_{i'}|\ v_s\rangle \,..\sum_j \langle u_j|\ v_t\rangle \sum_{j'} \langle u_{j'}|\ v_t\rangle \,..$$
$$\times\sqrt{N!}\ S_N\,|1:u_i;\ 2:u_{i'};\ ..;\ p_s+1:u_j;\ p_s+2:u_{j'};\ ..\rangle \tag{A-60}$$

or else[6], taking into account (A-55):

$$\left[\sum_i \langle u_i|\ v_s\rangle\, a^\dagger_{u_i}\right]\left[\sum_{i'} \langle u_{i'}|\ v_s\rangle\, a^\dagger_{u_{i'}}\right] ..\left[\sum_j \langle u_j|\ v_t\rangle\, a^\dagger_{u_j}\right]\left[\sum_{j'} \langle u_{j'}|\ v_t\rangle\, a^\dagger_{u_{j'}}\right] ..|0\rangle \tag{A-61}$$

We have thus shown that the operators $\left[a^\dagger_{v_s}\right]^{p_s}\left[a^\dagger_{v_t}\right]^{p_t}$ .. act on the vacuum state in the same way as the operators defined by (A-51), raised to the powers p_s, p_t , ..

When the occupation numbers p_s, p_t, .. can take on any values, the kets (A-56) span the entire Fock space. Writing the previous equality for $p_s$and p_s+1, we see that the action on all the basis kets of $a^\dagger_{v_s}$ and of $\sum_i \langle u_i|\ v_s\rangle\, a^\dagger_{u_i}$ yields the same result, establishing the equality between these two operators. Relation (A-52) can be readily obtained by Hermitian conjugation.

(ii) Fermions

The demonstration is identical, with the constraint that the occupation numbers are 0 or 1. As this requires no changes in the operator or state order, it involves no sign changes.

B. One-particle symmetric operators

Using creation and annihilation operators makes it much easier to deal, in the Fock space, with physical operators that are thus symmetric (§ C-4-a-β of Chapter XIV). We first study the simplest of such operators, those which act on a single particle and are called "one-particle operators".

B-1. Definition

Consider an operator $\widehat{f}$ defined in the space of individual states; $\widehat{f}(q)$ acts in the state space of particle q. It could be for example the momentum of the q-th particle, or its angular momentum with respect to the origin. We now build the operator associated with the total momentum of the N-particle system, or its total angular momentum, which is the sum over q of all the $\widehat{f}(q)$ associated with the individual particles.

A one-particle symmetric operator acting in the space $\mathcal{E}_S(N)$ for bosons - or $\mathcal{E}_A(N)$ for fermions - is therefore defined by:

$$\widehat{F}^{(N)} = \sum_{q=1}^{N} \widehat{f}(q) \tag{B-1}$$

[6]In this relation, the first p_s sums are identical, as are the next p_t sums, etc.

(contrary to states, which are symmetric for bosons and antisymmetric for fermions, the physical operators are always symmetric). The operator $\widehat{F}$ acting in the Fock space is defined as the operator $\widehat{F}^{(N)}$ acting either in $\mathcal{E}_S(N)$ or in $\mathcal{E}_A(N)$, depending on the specific case. Since the basis for the entire Fock space is the union of the bases of these spaces for all values of N, the operator $\widehat{F}$ is thus well defined in the direct sum of all these subspaces. To summarize:

$$\widehat{F}^{(N)} \quad ; \quad N = 1, 2, 3, ... \quad \Longrightarrow \widehat{F} \tag{B-2}$$

Using (B-1) directly to compute the matrix elements of $\widehat{F}$ often leads to tedious manipulations. Starting with an operator involving numbered particles, we place it between states with numbered particles; we then symmetrize the bra, the ket, and take into account the symmetry of the operator (*cf.* footnote 1). This introduces several summations (on the particles and on the permutations) that have to be properly regrouped to be simplified. We will now show that expressing $\widehat{F}$ in terms of creation and annihilation operators avoids all these intermediate calculations, taking nevertheless into account all the symmetry properties.

B-2. Expression in terms of the operators a and $a^\dagger$

We choose a basis $\{|u_i\rangle\}$ for the individual states. The matrix elements f_{kl} of the one-particle operator $\widehat{f}$ are given by:

$$f_{kl} = \langle u_k | \, \widehat{f} \, | u_l \rangle \tag{B-3}$$

They can be used to expand the operator itself as follows:

$$\widehat{f}(q) = \sum_{k,l} |q : u_k\rangle \langle q : u_k| \, \widehat{f}(q) \, |q : u_l\rangle \langle q : u_l| = \sum_{k,l} f_{kl} \; |q : u_k\rangle \langle q : u_l| \tag{B-4}$$

B-2-a. Action of $F^{(N)}$ on a ket with N particles

Using in (B-1) the expression (B-4) for $\widehat{f}(q)$ leads to:

$$\widehat{F}^{(N)} = \sum_{k,l} f_{kl} \sum_{q=1}^{N} \; |q : u_k\rangle \langle q : u_l| \tag{B-5}$$

The action of $\widehat{F}^{(N)}$ on a symmetrized ket written as (A-9) therefore includes a sum over k and l of terms:

$$\left[\sum_{q=1}^{N} \; |q : u_k\rangle \langle q : u_l| \right] \quad |u_i, u_i, .., u_j, u_j, ..\rangle \tag{B-6}$$

with coefficients f_{kl}. Let us use (A-7) or (A-10) to compute this ket for given values of k and l. As the operator contained in the bracket is symmetric with respect to the exchange of particles, it commutes with the two operators S_N and A_N (§ C-4-a-β of Chapter XIV)), and the ket can be written as:

$$\sqrt{\frac{N!}{n_i! n_j! .. n_m! ..}} \, S_N \, A_N \left[\sum_{q=1}^{N} \; |q : u_k\rangle \langle q : u_l| \right] \\ |1 : u_i; \; 2 : u_i; ..; \; n_i : u_i; \; n_i + 1 : u_j; \; ..; \; q : u_m; \; ..\rangle \tag{B-7}$$

In the summation over q, the only non-zero terms are those for which the individual state $|u_l\rangle$ coincides with the individual state $|u_m\rangle$ occupied in the ket on the right by the particle labeled q; there are n_l different values of q that obey this condition (i.e. none or one for fermions). For these n_l terms, the operator $|q:u_k\rangle\langle q:u_l|$ transforms the state $|u_m\rangle$ into $|u_k\rangle$, then S_N (or A_N) reconstructs a symmetrized (but not normalized) ket:

$$\sqrt{\frac{N!}{n_i!n_j!..n_m!..}}\begin{matrix}S_N\\A_N\end{matrix}\ |1:u_i;\ ..;\ n_i+1:u_j;\ ..;\ q:u_k;\ ..\rangle \tag{B-8}$$

This ket is always the same for all the numbers q among the n_l selected ones (for fermions, this term might be zero, if the state $|u_k\rangle$ was already occupied in the initial ket). We shall then distinguish two cases:

(i) For $k \neq l$, and for bosons, the ket written in (B-8) equals:

$$\sqrt{\frac{n_k+1}{n_l}}\,|u_i,u_i,..,u_j,u_j,..,u_k,..\rangle \tag{B-9}$$

where the square root factor comes from the variation in the occupation numbers n_k and n_l, which thus change the numerical coefficients in the definition (A-7) of the Fock states. As this ket is obtained n_l times, this factor becomes $\sqrt{(n_k+1)n_l}$. This is exactly the factor obtained by the action on the same symmetrized ket of the operator $a_k^\dagger a_l$, which also removes a particle from the state $|u_l\rangle$ and creates a new one in the state $|u_k\rangle$. Consequently, the operator $a_k^\dagger a_l$ reproduces exactly the same effect as the sum over q.

For fermions, the result is zero except when, in the initial ket, the state $|u_l\rangle$ was occupied by a particle, and the state $|u_k\rangle$ empty, in which case no numerical factor appears; as before, this is exactly what the action of the operator $a_k^\dagger a_l$ would do.

(ii) if $k = l$, for bosons the only numerical factor involved is n_l, coming from the number of terms in the sum over q that yields the same symmetrized ket. For fermions, the only condition that yields a non-zero result is for the state $|u_l\rangle$ to be occupied, which also leads to the factor n_l. In both cases, the sum over q amounts to the action of the operator $a_k^\dagger a_l$.

We have shown that:

$$\sum_{q=1}^{N}\ |q:u_k\rangle\langle q:u_l| = a_k^\dagger a_l \tag{B-10}$$

The summation over k and l in (B-5) then yields:

$$\widehat{F}^{(N)} = \sum_{k,l}\ f_{kl}\ \ a_k^\dagger a_l = \sum_{k,l}\langle u_k|\,\widehat{f}\,|u_l\rangle\ \ a_k^\dagger a_l \tag{B-11}$$

B-2-b. Expression valid in the entire Fock space

The right-hand side of (B-11) contains an expression completely independent of the space $\mathcal{E}_S(N)$ or $\mathcal{E}_A(N)$ in which we defined the action of the operator $\widehat{F}^{(N)}$. Since we defined operator $\widehat{F}$ as acting as $\widehat{F}^{(N)}$ in each of these subspaces having fixed N, we can simply write:

$$\boxed{\widehat{F} = \sum_{k,l}\langle u_k|\,\widehat{f}\,|u_l\rangle\ \ a_k^\dagger a_l} \tag{B-12}$$

This is the expression of one-particle symmetric operators we were looking for. Its form is valid for any value of N and the particles are no longer numbered; it contains equal numbers of creation and annihilation operators, which only act on occupation numbers.

Comment:

Choosing the proper basis $\{|u_i\rangle\}$, it is always possible to diagonalize the Hermitian operator $\widehat{f}$ and write:

$$f_{kl} = f_k \times \delta_{kl} \tag{B-13}$$

Equality (B-11) is then simply written as:

$$\widehat{F} = \sum_k f_k \ a_k^\dagger a_k = \sum_k f_k \ \widehat{n}_k \tag{B-14}$$

where $\widehat{n}_k = \widehat{n}_{u_k}$ is the occupation number operator in the state $|u_k\rangle$ defined in (A-28).

B-3. Examples

A first very simple example is the operator $\widehat{N}$, already described in (A-29), and corresponding to the total number of particles:

$$\widehat{N} = \sum_i \widehat{n}_i = \sum_i a_i^\dagger a_i \tag{B-15}$$

As expected, this operator does not depend on the basis $\{|u_i\rangle\}$ chosen to count the particles, as we now show. Using the unitary transformations of operators (A-51) and (A-52), and with the full notation for the creation and annihilation operators to avoid any ambiguity, we get:

$$\sum_i a_{u_i}^\dagger a_{u_i} = \sum_i \sum_{s,t} \langle v_s|\ u_i\rangle \ \langle u_i|\ v_t\rangle \ a_{v_s}^\dagger a_{v_t} = \sum_{s,t} \delta_{st} \ a_{v_s}^\dagger a_{v_t} \tag{B-16}$$

which shows that:

$$\widehat{N} = \sum_i a_{u_i}^\dagger a_{u_i} = \sum_s a_{v_s}^\dagger a_{v_s} \tag{B-17}$$

For a spinless particle one can also define the operator corresponding to the probability density at point $\mathbf{r}_0$:

$$\widehat{f} = |\mathbf{r}_0\rangle \ \langle \mathbf{r}_0| \tag{B-18}$$

Relation (B-12) then leads to the "particle local density" (or "single density") operator:

$$\widehat{D}(\mathbf{r}_0) = \sum_{k,l} u_k^*(\mathbf{r}_0) u_l(\mathbf{r}_0) \ a_k^\dagger a_l \tag{B-19}$$

The same procedure as above shows that this operator is independent of the basis $\{|u_i\rangle\}$ chosen in the individual states space.

Let us assume now that the chosen basis is formed by the eigenvectors $|\mathbf{k}_i\rangle$ of a particle's momentum $\hbar\mathbf{k}_i$, and that the corresponding annihilation operators are noted $a_{\mathbf{k}_i}$. The operator associated with the total momentum of the system can be written as:

$$\widehat{\mathbf{P}} = \sum_{\mathbf{k}_i} \hbar\mathbf{k}_i \; a^\dagger_{\mathbf{k}_i} a_{\mathbf{k}_i} = \sum_{\mathbf{k}_i} \hbar\mathbf{k}_i \; \widehat{n}_{\mathbf{k}_i} \tag{B-20}$$

As for the kinetic energy of the particles, its associated operator is expressed as:

$$\widehat{H}_0 = \sum_{\mathbf{k}_i} \frac{\hbar^2\mathbf{k}_i^2}{2m} \; a^\dagger_{\mathbf{k}_i} a_{\mathbf{k}_i} = \sum_{\mathbf{k}_i} \frac{\hbar^2\mathbf{k}_i^2}{2m} \; \widehat{n}_{\mathbf{k}_i} \tag{B-21}$$

B-4. Single particle density operator

Consider the average value $\left\langle \widehat{F} \right\rangle$ of a one-particle operator $\widehat{F}$ in an arbitrary N-particle quantum state. It can be expressed, using relation (B-12), as a function of the average values of operator products $a^\dagger_k a_l$:

$$\left\langle \widehat{F} \right\rangle = \sum_{k,l} \langle u_k | \, \widehat{f} \, | u_l \rangle \;\; \left\langle a^\dagger_k a_l \right\rangle \tag{B-22}$$

This expression is close to that of the average value of an operator for a physical system composed of a single particle. Remember (Complement $\mathrm{E_{III}}$, § 4-b) that if a system is described by a single particle density operator $\widehat{\rho}_1\,(1)$, the average value of any operator $\widehat{f}\,(1)$ is written as:

$$\left\langle \widehat{f}\,(1) \right\rangle = \mathrm{Tr}\left\{ \widehat{f}\,(1)\,\widehat{\rho}_1\,(1) \right\} = \sum_{k,l} \langle u_k | \, \widehat{f} \, | u_l \rangle \;\; \langle u_l | \, \widehat{\rho}_1 \, | u_k \rangle \tag{B-23}$$

The above two expressions can be made to coincide if, for the system of identical particles, we introduce a "density operator reduced to a single particle" $\widehat{\rho}_1$ whose matrix elements are defined by:

$$\langle u_l | \, \widehat{\rho}_1 \, | u_k \rangle = \left\langle a^\dagger_k a_l \right\rangle \tag{B-24}$$

This reduced operator allows computing average values of all the single particle operators as if the system consisted only of a single particle:

$$\left\langle \widehat{F} \right\rangle = \mathrm{Tr}\left\{ \widehat{f} \; \widehat{\rho}_1 \right\} \tag{B-25}$$

where the trace is taken in the state space of a single particle.

The trace of the reduced density operator thus defined is not equal to unity, but to the average particle number as can be shown using (B-24) and (B-15):

$$\mathrm{Tr}\,\{\widehat{\rho}_1\} = \sum_k \left\langle a^\dagger_k a_k \right\rangle = \left\langle \widehat{N} \right\rangle \tag{B-26}$$

This normalization convention can be useful. For example, the diagonal matrix element of $\widehat{\rho}_1$ in the position representation is simply the average of the particle local density defined in (B-19):

$$\langle \mathbf{r}_0 | \, \widehat{\rho}_1 \, | \mathbf{r}_0 \rangle = \left\langle \widehat{D}(\mathbf{r}_0) \right\rangle \tag{B-27}$$

It is however easy to choose a different normalization for the reduced density operator: its trace can be made equal to 1 by dividing the right-hand side of definition (B-24) by the factor $\left\langle \widehat{N} \right\rangle$.

C. Two-particle operators

We now extend the previous results to the case of two-particle operators.

C-1. Definition

Consider a physical quantity involving two particles, labeled q and q'. It is associated with an operator $\widehat{g}(q, q')$ acting in the state space of these two particles (the tensor product of the two individual state's spaces). Starting from this binary operator, the easiest way to obtain a symmetric N-particle operator is to sum all the $\widehat{g}(q, q')$ over all the particles q and q', where the two subscripts q and q' range from 1 to N. Note, however, that in this sum all the terms where $q = q'$ add up to form a one-particle operator of exactly the same type as those studied in § B-1. Consequently, to obtain a real two-particle operator we shall exclude the terms where $q = q'$ and define:

$$\widehat{G}^{(N)} = \frac{1}{2} \sum_{q,q'=1;\ q \neq q'}^{N} \widehat{g}(q, q') \tag{C-1}$$

The factor $1/2$ present in this expression is arbitrary but often handy. If for example the operator describes an interaction energy that is the sum of the contributions of all the distinct pairs of particles, $\widehat{g}(q, q')$ and $\widehat{g}(q', q)$ corresponding to the same pair are equal and appear twice in the sum over q and q': the factor $1/2$ avoids counting them twice. Whenever $\widehat{g}(q, q') = \widehat{g}(q', q)$, it is equivalent to write $\widehat{G}^{(N)}$ in the form:

$$\widehat{G}^{(N)} = \sum_{q<q'}^{N} \widehat{g}(q, q') \tag{C-2}$$

As with the one-particle operators, expression (C-1) defines symmetric operators separately in each physical state's space having a given particle number N. This definition may be extended to the entire Fock space, which is their direct sum over all N. This results in a more general operator $\widehat{G}$, following the same scheme as for (B-2):

$$\widehat{G}^{(N)} \quad ; \quad N = 1, 2, 3, ... \quad \Longrightarrow \widehat{G} \tag{C-3}$$

C-2. A simple case: factorization

Let us first assume the operator $\widehat{g}(q,q')$ can be factored as:

$$\widehat{g}(q,q') = \widehat{f}(q)\widehat{h}(q') \tag{C-4}$$

The operator written in (C-1) then becomes:

$$\widehat{G}^{(N)} = \frac{1}{2}\sum_{q,q'=1;\ q\neq q'}^{N} \widehat{f}(q)\widehat{h}(q') = \frac{1}{2}\left[\sum_{q=1}^{N}\widehat{f}(q)\sum_{q'=1}^{N}\widehat{h}(q') - \sum_{q=1}^{N}\widehat{f}(q)\widehat{h}(q)\right] \tag{C-5}$$

The right-hand side of this expression starts with a product of one-particle operators, each of which can be replaced, following (B-11), by its expression as a function of the creation and annihilation operators:

$$\sum_{q=1}^{N}\widehat{f}(q) = \sum_{i,k}\ \langle u_i|\,\widehat{f}\,|u_k\rangle\ a_i^\dagger a_k \qquad \text{and} \qquad \sum_{q'=1}^{N}\widehat{h}(q') = \sum_{j,l}\ \langle u_j|\,\widehat{h}\,|u_l\rangle\ a_j^\dagger a_l \tag{C-6}$$

As for the last term on the right-hand side of (C-5), it is already a single particle operator:

$$\sum_{q=1}^{N}\widehat{f}(q)\widehat{h}(q) = \sum_{i,l}\ \langle u_i|\,\widehat{f}\,\widehat{h}\,|u_l\rangle\ a_i^\dagger a_l \tag{C-7}$$

This leads to:

$$\widehat{G}^{(N)} = \frac{1}{2}\left[\sum_{i,j,k,l}\langle u_i|\,\widehat{f}\,|u_k\rangle\,\langle u_j|\,\widehat{h}\,|u_l\rangle\left(a_i^\dagger a_k\right)\left(a_j^\dagger a_l\right) - \sum_{i,l}\langle u_i|\,\widehat{f}\,\widehat{h}\,|u_l\rangle\ a_i^\dagger a_l\right] \tag{C-8}$$

We can then use general relations (A-49) to transform the operator product:

$$\left(a_i^\dagger a_k\right)\left(a_j^\dagger a_l\right) = \eta\ a_i^\dagger a_j^\dagger a_k a_l + \delta_{jk}\ a_i^\dagger a_l = a_i^\dagger a_j^\dagger a_l a_k + \delta_{jk}\ a_i^\dagger a_l \tag{C-9}$$

Including this form in the first term on the right-hand side of (C-8) yields, for the δ_{jk} contribution:

$$\sum_{i,l}\sum_{k}\langle u_i|\,\widehat{f}\,|u_k\rangle\,\langle u_k|\,\widehat{h}\,|u_l\rangle\ a_i^\dagger a_l = \sum_{i,l}\ \langle u_i|\,\widehat{f}\,\widehat{h}\,|u_l\rangle\ a_i^\dagger a_l \tag{C-10}$$

which exactly cancels the second term of (C-8). Consequently, we are left with:

$$\widehat{G}^{(N)} = \frac{1}{2}\sum_{i,j,k,l}\langle u_i|\,\widehat{f}\,|u_k\rangle\,\langle u_j|\,\widehat{h}\,|u_l\rangle\ a_i^\dagger a_j^\dagger a_l a_k \tag{C-11}$$

As the right-hand side of this expression has the same form in all spaces having a fixed N, it is also valid for the operator $\widehat{G}$ acting in the entire Fock space.

C-3. General case

Any two-particle operator $\widehat{g}(q,q')$ may be decomposed as a sum of products of single particle operators:

$$\widehat{g}(q,q') = \sum_{\alpha,\beta} c_{\alpha,\beta}\ \widehat{f}_\alpha(q)\widehat{h}_\beta(q') \tag{C-12}$$

where the coefficients $c_{\alpha,\beta}$ are numbers[7]. Hence expression (C-1) can be written as:

$$\widehat{G}^{(N)} = \frac{1}{2}\sum_{q,q'=1;\ q\neq q'}^{N} \widehat{g}(q,q') = \frac{1}{2}\sum_{\alpha,\beta} c_{\alpha,\beta}\sum_{q=1}^{N}\sum_{q'=1;\ q'\neq q}^{N} \widehat{f}_\alpha(q)\widehat{h}_\beta(q') \tag{C-13}$$

In this linear combination with coefficients $c_{\alpha,\beta}$, each term (corresponding to a given α and β) is of the form (C-5) and can therefore be replaced by expression (C-11). This leads to:

$$\widehat{G}^{(N)} = \frac{1}{2}\sum_{\alpha,\beta} c_{\alpha,\beta}\sum_{i,j,k,l} \langle u_i|\,\widehat{f_\alpha}\,|u_k\rangle\,\langle u_j|\,\widehat{h_\beta}\,|u_l\rangle\ a_i^\dagger a_j^\dagger a_l a_k \tag{C-14}$$

The right-hand side of this equation has the same form in all the spaces of fixed N; hence it is valid in the entire Fock space. Furthermore, we recognize in the summation over α and β the matrix element of $\widehat{g}$ as defined by (C-12):

$$g_{i,j,k,l} = \langle 1:u_i;\ 2:u_j|\,\widehat{g}(1,2)\,|1:u_k;\ 2:u_l\rangle = \sum_{\alpha,\beta} c_{\alpha,\beta}\ \langle u_i|\,\widehat{f_\alpha}\,|u_k\rangle\,\langle u_j|\,\widehat{h_\beta}\,|u_l\rangle \tag{C-15}$$

The final result is then:

$$\boxed{\widehat{G} = \frac{1}{2}\sum_{i,j,k,l} \langle 1:u_i;\ 2:u_j|\,\widehat{g}(1,2)\,|1:u_k;\ 2:u_l\rangle\ a_i^\dagger a_j^\dagger a_l a_k} \tag{C-16}$$

which is the general expression for a two-particle symmetric operator.

As for the one-particle operators, each term of expression (C-16) for the two-particle operators contains equal numbers of creation and annihilation operators. Consequently, these symmetric operators do not change the total number of particles, as was obvious from their initial definition.

C-4. Two-particle reduced density operator

Relation (C-16) implies that the average value of any two-particle operator may be written as:

$$\left\langle \widehat{G} \right\rangle = \frac{1}{2}\sum_{i,j,k,l} \langle 1:u_i;\ 2:u_j|\,\widehat{g}(1,2)\,|1:u_k;\ 2:u_l\rangle\ \left\langle a_i^\dagger a_j^\dagger a_l a_k \right\rangle \tag{C-17}$$

[7] The two-particle state space is the tensor product of the two spaces of individual states (see § F-4-b of Chapter II). In the same way, the space of operators acting on two particles is the tensor product of the spaces of operators acting separately on these particles. For example, the operator for the interaction potential between two particles can be decomposed as a sum of products of two operators: the first one is a function of the position of the first particle, and the second one of the position of the second particle.

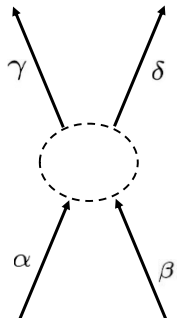


Figure 1: Physical interaction between two identical particles: initially in the states $|u_{k_\alpha}\rangle$ and $|u_{k_\beta}\rangle$ (schematized by the letters α and β), the particles are transferred to the states $|u_{k_\gamma}\rangle$ and $|u_{k_\delta}\rangle$ (schematized by the letters γ and δ)

This expression is similar to the average value of an operator $\widehat{g}\,(1,2)$ for a two-particle system having a density operator $\widehat{\rho}_2\,(1,2)$:

$$\langle \widehat{g}\,(1,2)\rangle = \sum_{i,j,k,l} \langle 1:u_i;\ 2:u_j|\,\widehat{g}(1,2)\,|1:u_k;\ 2:u_l\rangle \times \langle 1:u_k;\ 2:u_l|\,\widehat{\rho}_2\,(1,2)\,|1:u_i;\ 2:u_j\rangle \tag{C-18}$$

which leads us to define a two-particle reduced density operator $\widehat{\rho}_2$:

$$\langle 1:u_k;\ 2:u_l|\,\widehat{\rho}_2\,|1:u_i;\ 2:u_j\rangle = \left\langle a_i^\dagger a_j^\dagger a_l a_k \right\rangle \tag{C-19}$$

In this definition we have left out the factor 1/2 of (C-17) since this will lead to a normalization of $\widehat{\rho}_2$ often more handy: the matrix element of $\widehat{\rho}_2$ in the position representation yields directly the double density (as well as the field correlation function that we shall study in § B-3-b of Chapter XVI). The trace of $\widehat{\rho}_2$ is then written:

$$\begin{aligned}\mathrm{Tr}\,\{\widehat{\rho}_2\} &= \sum_{i,j} \left\langle a_i^\dagger a_j^\dagger a_j a_i \right\rangle = \sum_{i,j} \left\langle a_i^\dagger a_i a_j^\dagger a_j - \delta_{ij} a_i^\dagger a_i \right\rangle \\ &= \left\langle \widehat{N}\left(\widehat{N}-1\right)\right\rangle \end{aligned} \tag{C-20}$$

It is obviously possible to divide the right-hand side of the definition of $\widehat{\rho}_2$ either by the factor 2, or else by the factor $\left\langle \widehat{N}\left(\widehat{N}-1\right)\right\rangle$ if we wish its trace to be equal to 1.

C-5. Physical discussion; consequences of the exchange

As mentioned in the introduction of this chapter, the equations no longer contain labeled particles, permutations, symmetrizers and antisymmetrizers; the total number of particles N has also disappeared. We may now continue the discussion begun in § D-2 of Chapter XIV concerning the exchange terms, but in a more general way since we no longer specify the total particle number N.

C-5-a. Two terms in the matrix elements

Consider a physical process (schematized in Figure 1) where, in a system of N identical particles, an interaction produces a transfer from the two states $\left|u_{k_\alpha}\right\rangle$ and $\left|u_{k_\beta}\right\rangle$ towards the two states $\left|u_{k_\gamma}\right\rangle$ and $\left|u_{k_\delta}\right\rangle$; we assume that the four states we are dealing with are different. In the summation over i, j, k, l of (C-16), the only terms involved in this process are those where the bra contains either $i = k_\gamma$ and $j = k_\delta$, or the opposite $i = k_\delta$ and $j = k_\gamma$; as for the ket, it must contain either $k = k_\alpha$ and $l = k_\beta$, or the opposite $k = k_\beta$ and $l = k_\alpha$. We are then left with four terms:

$$\begin{aligned}
&\tfrac{1}{2}\left\langle 1 : u_{k_\gamma};\ 2 : u_{k_\delta}\right|\widehat{g}\left|1 : u_{k_\alpha};\ 2 : u_{k_\beta}\right\rangle\ a^\dagger_{k_\gamma} a^\dagger_{k_\delta} a_{k_\beta} a_{k_\alpha}\\
&\tfrac{1}{2}\left\langle 1 : u_{k_\delta};\ 2 : u_{k_\gamma}\right|\widehat{g}\left|1 : u_{k_\beta};\ 2 : u_{k_\alpha}\right\rangle\ a^\dagger_{k_\delta} a^\dagger_{k_\gamma} a_{k_\alpha} a_{k_\beta}\\
&\tfrac{1}{2}\left\langle 1 : u_{k_\gamma};\ 2 : u_{k_\delta}\right|\widehat{g}\left|1 : u_{k_\beta};\ 2 : u_{k_\alpha}\right\rangle\ a^\dagger_{k_\gamma} a^\dagger_{k_\delta} a_{k_\alpha} a_{k_\beta}\\
&\tfrac{1}{2}\left\langle 1 : u_{k_\delta};\ 2 : u_{k_\gamma}\right|\widehat{g}\left|1 : u_{k_\alpha};\ 2 : u_{k_\beta}\right\rangle\ a^\dagger_{k_\delta} a^\dagger_{k_\gamma} a_{k_\beta} a_{k_\alpha}
\end{aligned} \tag{C-21}$$

However, since the numbers used to label the particles are dummy variables, the first two matrix elements shown in (C-21) are equal and so are the last two. In addition, the product of creation and annihilation operators obey the following relations, for bosons ($\eta = 1$) as well as for fermions ($\eta = -1$):

$$\begin{aligned}
a^\dagger_{k_\delta} a^\dagger_{k_\gamma} a_{k_\alpha} a_{k_\beta} &= a^\dagger_{k_\gamma} a^\dagger_{k_\delta} a_{k_\beta} a_{k_\alpha}\\
a^\dagger_{k_\gamma} a^\dagger_{k_\delta} a_{k_\alpha} a_{k_\beta} &= a^\dagger_{k_\delta} a^\dagger_{k_\gamma} a_{k_\beta} a_{k_\alpha} = \eta\ a^\dagger_{k_\gamma} a^\dagger_{k_\delta} a_{k_\beta} a_{k_\alpha}
\end{aligned} \tag{C-22}$$

These relations are obvious for bosons since we only commute either creation operators or annihilation operators. For fermions, as we assumed all the states were different, the anticommutation of operators a or of operators $a^\dagger$, leads to sign changes; these may cancel out depending on whether the number of anticommutations is even or odd. If we now double the sum of the first and last term of (C-21), we obtain the final contribution to (C-16):

$$\begin{aligned}
&a^\dagger_{k_\gamma} a^\dagger_{k_\delta} a_{k_\beta} a_{k_\alpha}\\
&\qquad\times\ \Big[\left\langle 1 : u_{k_\gamma};\ 2 : u_{k_\delta}\right|\widehat{g}\left|1 : u_{k_\alpha};\ 2 : u_{k_\beta}\right\rangle\\
&\qquad\qquad+\ \eta\left\langle 1 : u_{k_\delta};\ 2 : u_{k_\gamma}\right|\widehat{g}\left|1 : u_{k_\alpha};\ 2 : u_{k_\beta}\right\rangle\Big]
\end{aligned} \tag{C-23}$$

Hence we are left with two terms whose relative sign depends on the nature (bosons or fermions) of the identical particles. They correspond to a different "switching point" for the incoming and outgoing individual states (Fig. 2).

For bosons, the product of the 4 operators in (C-23) acting on an occupation number ket introduces the square root:

$$\sqrt{n_{k_\alpha} n_{k_\beta}\left(n_{k_\gamma}+1\right)\left(n_{k_\delta}+1\right)} \tag{C-24}$$

For large occupation numbers, this square root may considerably increase the value of the matrix element. For fermions, however, this amplification effect does not occur. Furthermore, if the direct and exchange matrix elements of $\widehat{g}$ are equal, they will cancel each other in (C-23) and the corresponding transition amplitude of this process will be zero.

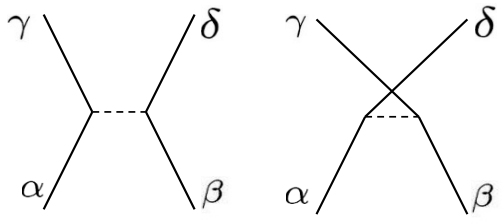


Figure 2: Two diagrams representing schematically the two terms appearing in equation (C-23); they differ by an exchange of the individual states of the exit particles. They correspond, in a manner of speaking, to a different "switching point" for the incoming and outgoing states. The solid lines represent the particles' free propagation, and the dashed lines their binary interaction.

C-5-b. Particle interaction energy, the direct and exchange terms

Many physics problems involve computing the average particle interaction energy. For the sake of simplicity, we shall only study here spinless particles (or, equivalently, particles being in the same internal spin state so that the corresponding quantum number does not come into play) and assume their interactions to be binary. These interactions are then described by an operator $\widehat{W}_{\rm int}$, diagonal in the $\{|\mathbf{r}_1, \mathbf{r}_2, ...\mathbf{r}_N\rangle\}$ basis (eigenstates of all the particles' positions), which multiplies each of these states by the function:

$$W_{\rm int}(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N) = \sum_{q<q'} W_2(\mathbf{r}_q, \mathbf{r}_{q'}) \tag{C-25}$$

In this expression, the function $W_2(\mathbf{r}_q, \mathbf{r}_{q'})$ yields the diagonal matrix elements of the operator $\widehat{W}_2(\mathbf{R}_q, \mathbf{R}_{q'})$ associated with the two-particle (q, q') interaction, where $\mathbf{R}_q$ is the quantum operator associated with the classical position $\mathbf{r}_q$. The matrix elements of this operator in the $|u_k; u_l\rangle$ basis is simply obtained by inserting a closure relation for each of the two positions. This leads to:

$$\begin{aligned}&\langle 1:u_i;\ 2:u_j|\,\widehat{W}_2(\mathbf{R}_1,\mathbf{R}_2)\,|1:u_k;\ 2:u_l\rangle \\ &\qquad = \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2\ W_2(\mathbf{r}_1,\mathbf{r}_2)\ u_i^*(\mathbf{r}_1) u_j^*(\mathbf{r}_2) u_k(\mathbf{r}_1) u_l(\mathbf{r}_2)\end{aligned} \tag{C-26}$$

α. General expression:

Replacing in (C-16) operator $\widehat{g}(1,2)$ by $\widehat{W}_{\rm int}(\mathbf{R}_1, \mathbf{R}_2)$ and taking (C-26) into account, we get:

$$\widehat{W}_{\rm int} = \frac{1}{2} \sum_{i,j,k,l} \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2\ W_2(\mathbf{r}_1,\mathbf{r}_2)\ u_i^*(\mathbf{r}_1) u_j^*(\mathbf{r}_2) u_k(\mathbf{r}_1) u_l(\mathbf{r}_2)\ a_i^\dagger a_j^\dagger a_l a_k \tag{C-27}$$

We can thus write the average value of the interaction energy in any normalized state $|\Phi\rangle$ as:

$$\left\langle \widehat{W}_{\rm int} \right\rangle = \langle\Phi|\, \widehat{W}_2\, |\Phi\rangle = \frac{1}{2} \int {\rm d}^3 r_1 \int {\rm d}^3 r_2 \; W_2(\mathbf{r}_1, \mathbf{r}_2)\; G_2(\mathbf{r}_1, \mathbf{r}_2) \tag{C-28}$$

where $G_2(\mathbf{r}_1, \mathbf{r}_2)$ is the spatial correlation function defined by:

$$G_2(\mathbf{r}_1, \mathbf{r}_2) = \sum_{i,j,k,l} u_i^*(\mathbf{r}_1) u_j^*(\mathbf{r}_2) u_k(\mathbf{r}_1) u_l(\mathbf{r}_2) \;\; \langle\Phi|\, a_i^\dagger a_j^\dagger a_l a_k\, |\Phi\rangle \tag{C-29}$$

Consequently, knowing the correlation function $G_2(\mathbf{r}_1, \mathbf{r}_2)$ associated with the quantum state $|\Phi\rangle$ allows computing directly, by a double spatial integration, the average interaction energy in that state.

Actually, as we shall see in more detail in § B-3 of Chapter XVI, $G_2(\mathbf{r}_1, \mathbf{r}_2)$ is simply the *double density*, equal to the probability density of finding any particle in $\mathbf{r}_1$ and another one in $\mathbf{r}_2$. The physical interpretation of (C-28) is simple: the average interaction energy is equal to the sum over all the particles' pairs of the interaction energy $W_{\rm int}(\mathbf{r}_1, \mathbf{r}_2)$ of a pair, multiplied by the probability of finding such a pair at points $\mathbf{r}_1$ and $\mathbf{r}_2$ (the factor 1/2 avoids the double counting of each pair).

β. Specific case: the Fock states

Let us assume the state $|\Phi\rangle$ is a Fock state, with specified occupation numbers n_i:

$$|\Phi\rangle = |n_1 : u_1;\; n_2 : u_2;\; ..;\; n_i : u_i;\; ..\rangle \tag{C-30}$$

We can compute explicitly, as a function of the n_i, the average values:

$$\langle\Phi|\, a_i^\dagger a_j^\dagger a_l a_k\, |\Phi\rangle \tag{C-31}$$

contained in (C-29). We first notice that to get a non-zero result, the two operators $a^\dagger$ must create particles in the same states from which they were removed by the two annihilation operators a. Otherwise the action of the four operators on the ket $|\Phi\rangle$ will yield a new Fock state orthogonal to the initial one, and hence a zero result. We must therefore impose either $i = k$ and $j = l$, or the opposite $i = l$ and $j = k$, or eventually the special case where all the subscripts are equal. The first case leads to what we call the "direct term", and the second, the "exchange term". We now compute their values.

(i) Direct term, $i = k$ and $j = l$, shown on the left diagram of Figure 3. If $i = j = k = l$, the four operators acting on $|\Phi\rangle$ reconstruct the same ket, multiplied by the factor $n_i\,(n_i - 1)$; this yields a zero result for fermions. If $i \neq j$, we can move the operator $a_k = a_i$ just to the right of the first operator $a_i^\dagger$ to form the particle number operator $\hat{n}_i$. This permutation in the operators' order does not change anything: for bosons, we are moving commuting operators, and for fermions, two anticommutations introduce two minus signs, which cancel each other. The same goes for the operators with subscript j, leading to the particle number $\hat{n}_j$. Finally, the direct term is equal to:

$$G_2^{\rm dir}(\mathbf{r}_1, \mathbf{r}_2) = \sum_{i \neq j} |u_i(\mathbf{r}_1)|^2\, |u_j(\mathbf{r}_2)|^2\, n_i n_j + \sum_i n_i\,(n_i - 1)\, |u_i(\mathbf{r}_1)|^2\, |u_i(\mathbf{r}_2)|^2 \tag{C-32}$$

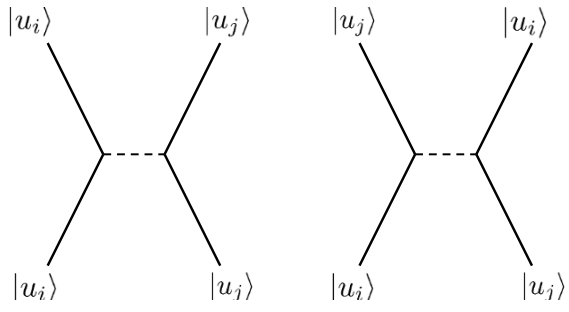


Figure 3: Schematic representation of a direct term (left diagram where each particle remains in the same individual state) and an exchange term (right diagram where the particles exchange their individual states). As in Figure 2, the solid lines represent the particles free propagation, and the dashed lines their binary interaction.

where the second sum is zero for fermions (n_i is equal to 0 or 1).

(ii) Exchange term, $i = l$ and $j = k$, shown on the right diagram of Figure 3. The case where all four subscripts are equal is already included in the direct term. To get the operators' product $\hat{n}_i\hat{n}_j$ starting from the product $a_i^\dagger a_j^\dagger a_i a_j$, we just have to permute the two central operators $a_j^\dagger a_i$; when $i \neq j$ this operation is of no consequence for bosons, but introduces a change of sign for fermions (anticommutation). The exchange term can therefore be written as $\eta G_2^{\rm ex}(\mathbf{r}_1, \mathbf{r}_2)$, with:

$$G_2^{\rm ex}(\mathbf{r}_1, \mathbf{r}_2) = \sum_{i \neq j} n_i n_j \left[u_i^*(\mathbf{r}_1) u_j^*(\mathbf{r}_2) u_i(\mathbf{r}_2) u_j(\mathbf{r}_1) \right] \tag{C-33}$$

Finally, the spatial correlation function (or double density) $G_2(\mathbf{r}_1, \mathbf{r}_2)$ is the sum of the direct and exchange terms:

$$G_2(\mathbf{r}_1, \mathbf{r}_2) = G_2^{\rm dir}(\mathbf{r}_1, \mathbf{r}_2) + \eta G_2^{\rm ex}(\mathbf{r}_1, \mathbf{r}_2) \tag{C-34}$$

where the factor η in front of the exchange term is 1 for bosons and -1 for fermions. The direct term only contains the product $|u_i(\mathbf{r}_1)|^2 |u_j(\mathbf{r}_2)|^2$ of the probability densities associated with the individual wave functions $u_i(\mathbf{r}_1)$ and $u_j(\mathbf{r}_2)$; it corresponds to non-correlated particles. We must add to it the exchange term, which has a more complex mathematical form and reveals correlations between the particles, even when they do not interact with each other. These correlations come from explicitly taking into account the fact that the particles are identical (symmetrization or antisymmetrization of the state vector). They are sometimes called "statistical correlations " and their spatial dependence will be studied in more detail in Complement $\mathrm{A_{XVI}}$.

Conclusion

The creation and annihilation operators introduced in this chapter lead to compact and general expressions for operators acting on any particle number N. These expressions involve the occupation numbers of the individual states but the particles are no longer

numbered. This considerably simplifies the computations performed on "N-body systems", like N interacting bosons or fermions. The introduction of approximations such as the mean field approximation used in the Hartree-Fock method (Complement $\mathrm{D_{XV}}$) will also be facilitated.

We have shown the complete equivalence between this approach and the one where we explicitly take into account the effect of permutations between numbered particles. It is important to establish this link for the study of certain physical problems. In spite of the overwhelming efficiency of the creation and annihilation operator formalism, the labeling of particles is sometimes useful or cannot be avoided. This is often the case for numerical computations, dealing with numbers or simple functions that require numbered particles and which, if needed, will be symmetrized (or antisymmetrized) afterwards.

In this chapter, we have only considered creation and annihilation operators with discrete subscripts. This comes from the fact that we have only used discrete bases $\{|u_i\rangle\}$ or $\{|v_j\rangle\}$ for the individual states. Other bases could be used, such as the position eigenstates $\{|\mathbf{r}\rangle\}$ of a spinless particle. The creation and annihilation operators will then be labeled by a continuous subscript $\mathbf{r}$. Fields of operators are thus introduced at each space point: they are called "field operators" and will be studied in the next chapter.

COMPLEMENTS OF CHAPTER XV, READER'S GUIDE

A_{XV} : PARTICLES AND HOLES	In an ideal gas of fermions, one can define creation and annihilation operators of holes (absence of a particle). Acting on the ground state, these operators allow building excited states. This is an important concept in condensed matter physics. *Easy to grasp, this complement can be considered to be a preliminary to Complement E_{XV}.*
B_{XV} : IDEAL GAS IN THERMAL EQUILIBRIUM; QUANTUM DISTRIBUTION FUNCTIONS	Studying the thermal equilibrium of an ideal gas of fermions or bosons, we introduce the distribution functions characterizing the physical properties of a particle or of a pair of particles. These distribution functions will be used in several other complements, in particular G_{XV} and H_{XV}. Bose-Einstein condensation is introduced in the case of bosons. The equation of state is discussed for both types of particles.

The list of complements continues on the next page

Series of four complements, discussing the behavior of particles interacting through a mean field created by all the others. Important, since the mean field concept is largely used throughout many domains of physics and chemistry.

C_{XV} : CONDENSED BOSON SYSTEM, GROSS-PITAEVSKII EQUATION

C_{XV} : This complement shows how to use a variational method for studying the ground state of a system of interacting bosons. The system is described by a one-particle wave function in which all the particles of the system accumulate. This wave function obeys the Gross-Pitaevskii equation.

D_{XV} : TIME-DEPENDENT GROSS-PITAEVSKII EQUATION

D_{XV} : This complement generalizes the previous one to the case where the Gross-Pitaevskii wave function is time-dependent. This allows us to obtain the excitation spectrum (Bogolubov spectrum), and to discuss metastable flows (superfluidity).

E_{XV} : FERMION SYSTEM, HARTREE-FOCK APPROXIMATION

E_{XV} : An ensemble of interacting fermions can be treated by a variational method, the Hartree-Fock approximation, which plays an essential role in atomic, molecular and solid state physics. In this approximation, the interaction of each particle with all the others is replaced by a mean field created by the other particles. The correlations introduced by the interactions are thus ignored, but the fermions' indistinguishability is accurately treated. This allows computing the energy levels of the system to an approximation that is satisfactory in many situations.

F_{XV} : FERMIONS, TIME-DEPENDENT HARTREE-FOCK APPROXIMATION

F_{XV} : We often have to study an ensemble of fermions in a time-dependent situation, as for example electrons in a molecule or a solid subjected to an oscillating electric field. The Hartree-Fock mean field method also applies to time-dependent problems. It leads to a set of coupled equations of motion involving a Hartree-Fock mean field potential, very similar to the one encountered for time-independent problems.

The list of complements continues on the next page

The mean field approximation can also be used to study the properties, at thermal equilibrium, of systems of interacting fermions or bosons. The variational method amounts to optimizing the one-particle reduced density operator. It permits generalizing to interacting particles a number of results obtained for an ideal gas (Complement B_{XV}).

G_{XV} : FERMIONS OR BOSONS: MEAN FIELD THERMAL EQUILIBRIUM	G_{XV}: The trial density operator at non-zero temperature can be optimized using a variational method. This leads to self-consistent Hartree-Fock equations, of the same type as those derived in Complement E_{XV}. We thus obtain an approximate value for the thermodynamic potential.
H_{XV} : APPLICATIONS OF THE MEAN FIELD METHOD FOR NON-ZERO TEMPERATURES (FERMIONS AND BOSONS)	H_{XV}: This complement discusses various applications of the method described in the previous complement: spontaneous magnetism of an ensemble of repulsive fermions, equation of state for bosons and instability in the presence of attractive interactions.

Complement A_{XV}

Particles and holes

1 Ground state of a non-interacting fermion gas 1621
2 New definition for the creation and annihilation operators 1622
3 Vacuum excitations 1623

Creation and annihilation operators are frequently used in solid state physics where the notion of particle and hole plays an important role. A good example is the study of metals or semiconductors, where we talk about an electron-hole pair created by photon absorption. A hole means an absence of a particle, but it has properties similar to a particle, like a mass, a momentum, an energy; the holes obey the same fermion statistics as the electrons they replace. Using creation or annihilation operators allows a better understanding of the hole concept. We will remain in the simple framework of a free particle gas, but the concepts can be generalized to the case of particles placed in an external potential or a Hartree-Fock mean potential (Complement E_{XV}).

1. Ground state of a non-interacting fermion gas

Consider a system of non-interacting fermions in their ground state. We assume for simplicity that they are all in the same spin state, and thus introduce no spin index (generalization to several spin states is fairly simple). As we showed in Complement C_{XIV}, this system in its ground state is described by a state where all the occupation numbers of the individual states having an energy lower than the Fermi energy E_F are equal to 1, and all the other individual states are empty. In momentum space, the only occupied states are all the individual states whose wave vector $\mathbf{k}$ is included in a sphere (called the "Fermi sphere") of radius k_F (the "Fermi radius") given by[1]:

$$\frac{\hbar^2 (k_F)^2}{2m} = E_F = \frac{\hbar^2}{2m} \left[\frac{6\pi^2 N}{L^3}\right]^{2/3} \tag{1}$$

where we have used the notation of formula (7) in Complement C_{XIV}: E_F is the Fermi energy (proportional to the particle density to the power 2/3), and L the edge length of the cube containing the N particles. When the system is in its ground state, all the individual states inside the Fermi sphere are occupied, whereas all the other individual states are empty. Choosing for the individual states basis $\{|u_i\rangle\}$ the plane wave basis, noted $\{|u_{\mathbf{k}_i}\rangle\}$ to explicit the wave vector $\mathbf{k}_i$, the occupation numbers are:

$$\begin{array}{ll} n_{\mathbf{k}_i} = 1 & \text{if } |\mathbf{k}_i| \leq k_F \\ n_{\mathbf{k}_i} = 0 & \text{if } |\mathbf{k}_i| > k_F \end{array} \tag{2}$$

[1] In Complement C_{XIV} we had assumed that both spin states of the electron gas were occupied, whereas this is not the case here. This explains why the bracket in formula (1) contains the coefficient $6\pi^2 N$ instead of $3\pi^2 N$.

In a macroscopic system, the number of occupied states is very large, of the order of the Avogadro number ($\simeq 10^{23}$). The ground state energy is given by:

$$E_0 = \sum_{|\mathbf{k}_i| \leq k_F} e_{\mathbf{k}_i} \tag{3}$$

with:

$$e_{\mathbf{k}_i} = \frac{\hbar^2 (k_i)^2}{2m} \tag{4}$$

The sum over $\mathbf{k}_i$ in (3) must be interpreted as a sum over all the $\mathbf{k}_i$ values that obey the boundary conditions in the box of volume L^3, as well as the restriction on the length of the vector $\mathbf{k}_i$ which must be smaller or equal to k_F.

2. New definition for the creation and annihilation operators

We now consider this ground state as a new "vacuum" $\widetilde{|0\rangle}$ and introduce creation operators that, acting on this vacuum, create excited states for this system. We define:

$$\begin{array}{l} \text{if } |\mathbf{k}_i| \leq k_F \begin{cases} b^\dagger_{\mathbf{k}_i} = a_{\mathbf{k}_i} \\ b_{\mathbf{k}_i} = a^\dagger_{\mathbf{k}_i} \end{cases} \\ \text{if } |\mathbf{k}_i| > k_F \begin{cases} c_{\mathbf{k}_i} = a_{\mathbf{k}_i} \\ c^\dagger_{\mathbf{k}_i} = a^\dagger_{\mathbf{k}_i} \end{cases} \end{array} \tag{5}$$

Outside the Fermi sphere, the new operators $c^\dagger_{\mathbf{k}_i}$ and $c_{\mathbf{k}_i}$ are therefore simple operators of creation (or annihilation) of a particle in a momentum state that is not occupied in the ground state. Inside the Fermi sphere, the results are just the opposite: operator $b^\dagger_{\mathbf{k}_i}$ creates a missing particle, that we shall call a "hole"; the adjoint operator $b_{\mathbf{k}_i}$ repopulates that level, hence destroying the hole. It is easy to show that the anticommutation relations for the new operators are:

$$\begin{array}{l} \left[b_{\mathbf{k}_i}, b_{\mathbf{k}_j}\right]_+ = \left[b^\dagger_{\mathbf{k}_i}, b^\dagger_{\mathbf{k}_j}\right]_+ = 0 \\ \left[b_{\mathbf{k}_i}, b^\dagger_{\mathbf{k}_j}\right]_+ = \delta_{\mathbf{k}_i \mathbf{k}_j} \end{array} \tag{6}$$

as well as:

$$\begin{array}{l} \left[c_{\mathbf{k}_i}, c_{\mathbf{k}_j}\right]_+ = \left[c^\dagger_{\mathbf{k}_i}, c^\dagger_{\mathbf{k}_j}\right]_+ = 0 \\ \left[c_{\mathbf{k}_i}, c^\dagger_{\mathbf{k}_j}\right]_+ = \delta_{\mathbf{k}_i \mathbf{k}_j} \end{array} \tag{7}$$

which are the same as for ordinary fermions. Finally, the cross anticommutation relations are:

$$\left[c_{\mathbf{k}_i}, b^\dagger_{\mathbf{k}_j}\right]_+ = \left[c^\dagger_{\mathbf{k}_i}, b_{\mathbf{k}_j}\right]_+ = \left[c_{\mathbf{k}_i}, b_{\mathbf{k}_j}\right]_+ = \left[c^\dagger_{\mathbf{k}_i}, b^\dagger_{\mathbf{k}_j}\right]_+ = 0 \tag{8}$$

3. Vacuum excitations

Imagine, for example, that with this new point of view we apply an annihilation operator $b_{\mathbf{k}_i}$, with $|\mathbf{k}_i| \leq k_F$, to the "new vacuum" $\widetilde{|0\rangle}$. The result must be zero since it is impossible to annihilate a non-existent hole. From the old point of view and according to (5), this amounts to applying the creation operator $a^\dagger_{\mathbf{k}_i}$ to a system where the individual state $|\mathbf{k}_i\rangle$ is already occupied, and the result is indeed zero, as expected. On the other hand, if we apply the creation operator $b^\dagger_{\mathbf{k}_i}$, with $|\mathbf{k}_i| \leq k_F$, to the new vacuum, the result is not zero: from the old point of view, it removes a particle from an occupied state, and in the new point of view it creates a hole that did not exist before. The two points of view are consistent.

Instead of talking about particles and holes, one can also use a general term, excitations (or "quasi-particles"). The creation operator of an excitation of $|\mathbf{k}_i| \leq k_F$ is the creation operator $b^\dagger_{\mathbf{k}_i} = a_{\mathbf{k}_i}$ of a hole ; the creation operator of an excitation of $|\mathbf{k}_i| > k_F$ is the creation operator $c^\dagger_{\mathbf{k}_i} = a^\dagger_{\mathbf{k}_i}$ of a particle. The vacuum state defined initially is a common eigenvector of all the particle annihilation operators, with eigenvalues zero; in a similar way, the new vacuum state $\widetilde{|0\rangle}$ is a common eigenvector of all the excitation annihilation operators. We therefore call it the "quasi-particle vacuum".

As we have neglected all particle interactions, the system Hamiltonian is written as:

$$\widehat{H} = \sum_{\mathbf{k}_i} e_{\mathbf{k}_i} n_{\mathbf{k}_i} = \sum_{\mathbf{k}_i} e_{\mathbf{k}_i} a^\dagger_{\mathbf{k}_i} a_{\mathbf{k}_i} = \sum_{|\mathbf{k}_i| \leq k_F} e_{\mathbf{k}_i} b_{\mathbf{k}_i} b^\dagger_{\mathbf{k}_i} + \sum_{|\mathbf{k}_i| > k_F} e_{\mathbf{k}_i} c^\dagger_{\mathbf{k}_i} c_{\mathbf{k}_i} \tag{9}$$

Taking into account the anticommutation relations between the operators $b_{\mathbf{k}_i}$ and $b^\dagger_{\mathbf{k}_i}$ we can rewrite this expression as:

$$\widehat{H} - E_0 = - \sum_{|\mathbf{k}_i| \leq k_F} e_{\mathbf{k}_i} b^\dagger_{\mathbf{k}_i} b_{\mathbf{k}_i} + \sum_{|\mathbf{k}_i| > k_F} e_{\mathbf{k}_i} c^\dagger_{\mathbf{k}_i} c_{\mathbf{k}_i} \tag{10}$$

where E_0 has been defined in (3) and simply shifts the origin of all the system energies. Relation (10) shows that holes (excitations with $|\mathbf{k}_i| \leq k_F$) have a negative energy, as expected since they correspond to missing particles. Starting from its ground state, to increase the system energy keeping the particle number constant, we must apply the operator $b^\dagger_{\mathbf{k}_i} c^\dagger_{\mathbf{k}_j}$ that creates both a particle and a hole: the system energy is then increased by the quantity $e_j - e_i$; inversely, to decrease the system energy, the adjoint operator $c_{\mathbf{k}_j} b_{\mathbf{k}_i}$ must be applied.

Comments:

(i) We have discussed the notion of hole in the context of free particles, but nothing in the previous discussion requires the one-particle energy spectrum to be simply quadratic as in (4). In semi-conductor physics for example, particles often move in a periodic potential, and occupy states in the "valence band" when their energy is lower than the Fermi level E_F, whereas the others occupy the "conduction band", separated from the previous band by an "energy gap". Sending a photon with an energy larger than this gap allows the creation of an electron-hole pair, easily studied in the formalism we just introduced.

A somewhat similar case occurs when studying the relativistic Dirac wave equation, where two energy continuums appear: one with energies greater than the electron rest energy mc^2 (where m is the electron mass, and c the speed of light), and one for negative energies less than $-mc^2$ associated with the positron (the antiparticle of the electron, having the opposite charge). The energy spectrum is relativistic, and thus different from formula (4), even inside each of those two continuums. However, the general formalism remains valid, the operators $b^\dagger_{\mathbf{k}_i}$ and $b_{\mathbf{k}_i}$ describing now, respectively, the creation and annihilation of a positron. The Dirac equation however leads to difficulties by introducing for example an infinity of negative energy states, assumed to be all occupied to avoid problems. A proper treatment of this type of relativistic problems must be done in the framework of quantum field theory.

(ii) An arbitrary N-particle Fock state $|\Phi\rangle$ does not have to be the ground state to be formally considered as a "quasi-particle vacuum". We just have to consider any annihilation operator on an already occupied individual state as a creation operator of a hole (i.e. of an excitation); we then define the corresponding hole (or excitation) annihilation operators, which all have in common the eigenvector $|\Phi\rangle$ with eigenvalue zero. This comment will be useful when studying the Wick theorem (Complement C_{XVI}). In § E of Chapter XVII, we shall see another example of a quasi-particle vacuum, but where, this time, the new annihilation operators are no longer acting on individual states but on states of pairs of particles.

Complement B_{XV}

Ideal gas in thermal equilibrium; quantum distribution functions

1 Grand canonical description of a system without interactions 1626
1-a Density operator . . . 1626
1-b Grand canonical partition function, grand potential . . . 1627
2 Average values of symmetric one-particle operators . . . 1628
2-a Fermion distribution function . . . 1629
2-b Boson distribution function . . . 1629
2-c Common expression . . . 1630
2-d Characteristics of Fermi-Dirac and Bose-Einstein distributions 1630
3 Two-particle operators . . . 1631
3-a Fermions . . . 1632
3-b Bosons . . . 1632
3-c Common expression . . . 1634
4 Total number of particles . . . 1635
4-a Fermions . . . 1635
4-b Bosons . . . 1637
5 Equation of state, pressure . . . 1640
5-a Fermions . . . 1640
5-b Bosons . . . 1641

This complement studies the average values of one- or two-particle operators for an ideal gas, in thermal equilibrium. It includes a discussion of several useful properties of the Fermi-Dirac and Bose-Einstein distribution functions, already introduced in Chapter XIV.

To describe thermal equilibrium, statistical mechanics often uses the grand canonical ensemble, where the particle number may fluctuate, with an average value fixed by the chemical potential μ (cf. Appendix VI, where you will find a number of useful concepts for reading this complement). This potential plays, with respect to the particle number, a role similar to the role the inverse of the temperature term $\beta = 1/k_B T$ plays with respect to the energy (k_B is the Boltzmann constant). In quantum statistical mechanics, Fock space is a good choice for the grand canonical ensemble as it easily allows changing the total number of particles. As a direct application of the results of §§ B and C of Chapter XV, we shall compute the average values of symmetric one- or two-particle operators for a system of identical particles in thermal equilibrium.

We begin in § 1 with the density operator for non-interacting particles, and then show in §§ 2 and 3 that the average values of the symmetric operators may be expressed in terms of the Fermi-Dirac and Bose-Einstein distribution functions, increasing their application range and hence their importance. In § 5, we shall study the equation of state for an ideal gas of fermions or bosons at temperature T and contained in a volume $\mathcal{V}$.

1. Grand canonical description of a system without interactions

We first recall how a system of non-interacting particles is described, in quantum statistical mechanics, by the grand canonical ensemble; more details on this subject can be found in Appendix VI, § 1-c.

1-a. Density operator

Using relations (42) and (43) of Appendix VI, we can write the grand canonical density operator ρ_{eq} (whose trace has been normalized to 1) as:

$$\rho_{eq} = \frac{1}{Z} e^{-\beta\left(\widehat{H}-\mu\widehat{N}\right)} \tag{1}$$

where Z is the grand canonical partition function:

$$Z = \mathrm{Tr}\left\{ e^{-\beta\left(\widehat{H}-\mu\widehat{N}\right)} \right\} \tag{2}$$

In these relations, $\beta = 1/(k_B T)$ is the inverse of the absolute temperature T multiplied by the Boltzmann constant k_B, and μ, the chemical potential (which may be fixed by a large reservoir of particles). Operators $\widehat{H}$ and $\widehat{N}$ are, respectively, the system Hamiltonian and the particle number operator defined by (B-15) in Chapter XV.

Assuming the particles do not interact, equation (B-1) of Chapter XV allows writing the system Hamiltonian $\widehat{H}$ as a sum of one-particle operators, in each subspace having a total number of particles equal to N:

$$\widehat{H} = \sum_{q=1}^{N} \widehat{h}(q) \tag{3}$$

Let us call $\{|u_k\rangle\}$ the basis of the individual states that are the eigenstates of the operator $\widehat{h}$. Noting $a_k^\dagger$ and a_k the creation and annihilation operators of a particle in these states, $\widehat{H}$ may be written as in (B-14):

$$\widehat{H} = \sum_k e_k \; a_k^\dagger a_k = \sum_k e_k \; \widehat{n}_k \tag{4}$$

where the e_k are the eigenvalues of $\widehat{h}$. Operator (1) can also be written as:

$$\rho_{eq} = \frac{1}{Z} e^{-\beta \sum_k (e_k-\mu) \; a_k^\dagger a_k} = Z^{-1} \prod\nolimits_k e^{-\beta(e_k-\mu) \; a_k^\dagger a_k} \tag{5}$$

We shall now compute the average values of all the one- or two-particle operators for a system described by the density operator (1).

1-b. Grand canonical partition function, grand potential

In statistical mechanics, the "grand potential" Φ associated with the grand canonical equilibrium is defined as the (natural) logarithm of the partition function, multiplied by $-k_B T$ (*cf.* Appendix VI, § 1-c-β):

$$\Phi = -k_B T \ln Z \tag{6}$$

where Z is given by (2). The trace appearing in this equation is easily computed in the basis of the Fock states built from the individual states $\{|u_k\rangle\}$, as we now show. The trace of a tensor product of operators (Chapter II, § F-2-b) is simply the product of the traces of each operator. The Fock space has the structure of a tensor product of the spaces associated with each of the $|u_k\rangle$ (each being spanned by kets having a population n_k ranging from zero to infinity – see comment (i) of § A-1-c in Chapter XV); we must thus compute a product of traces in each of these spaces. For a fixed k, we sum all the diagonal elements over all the values of n_k, then take the product over all k's, which leads to:

$$Z = \prod_k \sum_{n_k} \exp\left[-\beta n_k \left(e_k - \mu\right)\right] \tag{7}$$

α. *Fermions*

For fermions, as n_k can only take the values 0 or 1 (two identical fermions never occupy the same individual state), we get:

$$Z_{\text{fermions}} = \prod_k \left[1 + e^{-\beta(e_k - \mu)}\right] \tag{8}$$

and:

$$\Phi_{\text{fermions}} = -k_B T \sum_k \ln \left[1 + e^{-\beta(e_k - \mu)}\right] \tag{9}$$

The index k must be summed over all the individual states. In case these states are also labeled by orbital and spin subscripts, these must also be included in the summation. Let us consider for example particles having a spin S and contained in a box of volume $\mathcal{V}$ with periodic boundary conditions. The individual stationary states may be written as $|\mathbf{k}, \nu\rangle$, where $\mathbf{k}$ obeys the periodic boundary conditions (Complement C_{XIV}, § 1-c) and the subscript ν takes $(2S+1)$ values. Assuming the particles to be free in the box (no spin Hamiltonian), each ν value yields the same contribution to Φ_{fermions}; in the large volume limit, expression (9) then becomes:

$$\Phi_{\text{fermions}} = -\left(2S+1\right) k_B T \frac{\mathcal{V}}{\left(2\pi\right)^3} \int \mathrm{d}^3 k \, \ln \left[1 + e^{-\beta(e_k - \mu)}\right] \tag{10}$$

β. *Bosons*

For bosons, the summation over n_k in (7) goes from $n_k = 0$ to infinity, which introduces a geometric series whose sum is readily computed. We therefore get:

$$Z_{\text{bosons}} = \prod_k \frac{1}{1 - e^{-\beta(e_k - \mu)}} \tag{11}$$

which leads to:

$$\Phi_{\text{bosons}} = k_B T \sum_k \ln \left[1 - e^{-\beta(e_k - \mu)}\right] \tag{12}$$

For a system of free particles with spin S, confined in a box with periodic boundary conditions, we obtain, in the large volume limit:

$$\Phi_{\text{bosons}} = (2S+1)\, k_B T \frac{\mathcal{V}}{(2\pi)^3} \int \mathrm{d}^3 k \, \ln \left[1 - e^{-\beta(e_k - \mu)}\right] \tag{13}$$

In a general way, for fermions as well as bosons, the grand potential directly yields the pressure P, as shown in relation (61) of AppendixVI:

$$\Phi = -P\mathcal{V} \tag{14}$$

Using the proper derivatives with respect to the equilibrium parameters (temperature, chemical potential, volume), it also yields the other thermodynamic quantities such as the energy, the specific heats, etc.

2. Average values of symmetric one-particle operators

Symmetric quantum operators for one, and then for two particles, were introduced in a general way in Chapter XV (§§ B and C). The general expression for a one-particle operator $\widehat{F}$ is given by equation (B-12) of that chapter. We can thus write:

$$\left\langle \widehat{F} \right\rangle = \sum_{i,j} \langle u_i | \, \widehat{f} \, | u_j \rangle \left\langle a_i^\dagger a_j \right\rangle \tag{15}$$

with, when the state of the system is given by the density operator (1):

$$\left\langle a_i^\dagger a_j \right\rangle = \mathrm{Tr} \left\{ \rho_{eq} a_i^\dagger a_j \right\} = Z^{-1} \, \mathrm{Tr} \left\{ e^{-\beta\left(\widehat{H} - \mu \widehat{N}\right)} a_i^\dagger a_j \right\} \tag{16}$$

This trace can be computed in the Fock state basis $|n_1, .., n_i .., n_j, ..\rangle$ associated with the eigenstates basis $\{|u_k\rangle\}$ of $\widehat{h}$. If $i \neq j$, operator $a_i^\dagger a_j$ destroys a particle in the individual state $|u_j\rangle$ and creates another one in the different state $|u_i\rangle$; it therefore transforms the Fock state $|n_1, .., n_i, .., n_j, ..\rangle$ into a different, hence orthogonal, Fock state $|n_1, .., n_i - 1.., n_j + 1, ..\rangle$. Operator ρ_{eq} then acts on this ket, multiplying it by a constant. Consequently, if $i \neq j$, all the diagonal elements of the operator whose trace is taken in (16) are zero; the trace is therefore zero. If $i = j$, this average value may be computed as for the partition function, since the Fock space has the structure of a tensor product of individual state's spaces. The trace is the product of the i value contribution by all the other k values contributions. We can thus write, in a general way:

$$\left\langle a_i^\dagger a_j \right\rangle = \delta_{ij} \; Z^{-1} \left\{ \sum_{n_i} n_i \exp\left[-\beta n_i \left(e_i - \mu\right)\right] \right\} \times \prod_{k \neq i} \left\{ \sum_{n_k} \exp\left[-\beta n_k \left(e_k - \mu\right)\right] \right\} \tag{17}$$

For $i = j$, this expression yields the average particle number in the individual state $|u_i\rangle$.

2-a. Fermion distribution function

As the occupation number only takes the values 0 and 1, the first bracket in expression (17) is equal to $\left[e^{-\beta(e_i-\mu)}\right]$; as for the other modes ($k \neq i$) contribution, in the second bracket, it has already been computed when we determined the partition function. We therefore obtain:

$$\left\langle a_i^\dagger a_j \right\rangle = \delta_{ij}\ Z^{-1} \left[e^{-\beta(e_i-\mu)}\right] \times \prod_{k\neq i} \left[1 + e^{-\beta(e_k-\mu)}\right] \tag{18}$$

Multiplying both the numerator and denominator by $1 + e^{-\beta(e_i-\mu)}$ allows reconstructing the function Z in the numerator, and, after simplification by Z, we get:

$$\left\langle a_i^\dagger a_j \right\rangle = \delta_{ij} \frac{e^{-\beta(e_i-\mu)}}{1 + e^{-\beta(e_i-\mu)}} = \delta_{ij}\ f_\beta^{FD}(e_i - \mu) \tag{19}$$

We find again the Fermi-Dirac distribution function f_β^{FD} (§ 1-b of Complement C_{XIV}):

$$f_\beta^{FD}(e-\mu) = \frac{e^{-\beta(e-\mu)}}{1 + e^{-\beta(e-\mu)}} = \frac{1}{e^{\beta(e-\mu)} + 1} \tag{20}$$

This distribution function gives the average population of each individual state $|u_i\rangle$ with energy e; its value is always less than 1, as expected for fermions.

The average value at thermal equilibrium of any one-particle operator is now readily computed by using (19) in relation (15).

2-b. Boson distribution function

The mode $j = i$ contribution can be expressed as:

$$\begin{aligned}\sum_{n_i=0}^{\infty} n_i\ \exp\left[-\beta n_i\,(e_i-\mu)\right] &= \frac{1}{\beta}\frac{\partial}{\partial\mu}\sum_{n_i=0}^{\infty}\exp\left[-\beta n_i\,(e_i-\mu)\right] \\ &= \frac{1}{\beta}\frac{\partial}{\partial\mu}\frac{1}{1-\exp\left[-\beta\,(e_i-\mu)\right]}\end{aligned} \tag{21}$$

We then get:

$$\left\langle a_i^\dagger a_j \right\rangle = \frac{\delta_{i,j}}{Z}\frac{e^{-\beta(e_i-\mu)}}{\left[1-e^{-\beta(e_i-\mu)}\right]^2} \times \prod_{k\neq i}\frac{1}{1-e^{-\beta(e_k-\mu)}} \tag{22}$$

which, using (11), amounts to:

$$\left\langle a_i^\dagger a_j \right\rangle = \delta_{ij}\ f_\beta^{BE}(e_i-\mu) \tag{23}$$

where the Bose-Einstein distribution function f_β^{BE} is defined as:

$$f_\beta^{BE}(e-\mu) = \frac{e^{-\beta(e-\mu)}}{1-e^{-\beta(e-\mu)}} = \frac{1}{e^{\beta(e-\mu)}-1} \tag{24}$$

This distribution function gives the average population of the individual state $|u_i\rangle$ with energy e. The only constraint of this population, for bosons, is to be positive. The chemical potential is always less than the lowest individual energy e_k. In case this energy is zero, μ must always be negative. This avoids any divergence of the function f_β^{BE}.

Hence for bosons, the average value of any one-particle operator is obtained by inserting (23) into relation (15).

2-c. Common expression

We define the function f_β as equal to either the function f_β^{FD} for fermions, or the function f_β^{BE} for bosons. We can write for both cases:

$$f_\beta(e-\mu) = \frac{1}{e^{\beta(e-\mu)} - \eta} \tag{25}$$

where the number η is defined as:

$$\begin{aligned} &\eta = -1 \text{ for fermions} \\ &\eta = +1 \text{ for bosons} \end{aligned} \tag{26}$$

2-d. Characteristics of Fermi-Dirac and Bose-Einstein distributions

We already gave in Complement C_{XIV} (Figure 3) the form of the Fermi-Dirac distribution. Figure 1 shows both the variations of this distribution and the Bose-Einstein distribution. For the sake of comparison, it also includes the variations of the classical Boltzmann distribution:

$$f_\beta^{\text{Boltzmann}}(e-\mu) = e^{-\beta(e-\mu)} \tag{27}$$

which takes on intermediate values between the two quantum distributions. For a non-interacting gas contained in a box with periodic boundary conditions, the lowest possible energy e is zero and all the others are positive. Exponential $e^{\beta(e_i-\mu)}$ is therefore always greater than $e^{-\beta\mu}$. We are now going to distinguish several cases, starting with the most negative values for the chemical potential.

(i) For a negative value of $\beta\mu$, with a modulus large compared to 1 (i.e. for $\mu \ll -k_BT$, which corresponds to the right-hand side of the figure), the exponential in the denominator of (25) is always much larger than 1 (whatever the energy e), and the distribution reduces to the classical Boltzmann distribution (27). Bosons and fermions have practically the same distribution; the gas is said to be "non-degenerate".

(ii) For a fermion system, the chemical potential has no upper boundary, but the population of an individual state can never exceed 1. If μ is positive, with $\mu \gg k_BT$:

– for low values of the energy, the factor 1 is much larger than the exponential term; the population of each individual state is almost equal to 1, its maximum value.

– if the energy e_i increases to values of the order μ, the population decreases and when $e_i \gg \mu$, it becomes practically equal to the value predicted by the Boltzmann exponential (27).

Most of the particles occupy, however, the individual states having an energy less or comparable to μ, whose population is close to 1. The fermion system is said to be "degenerate".

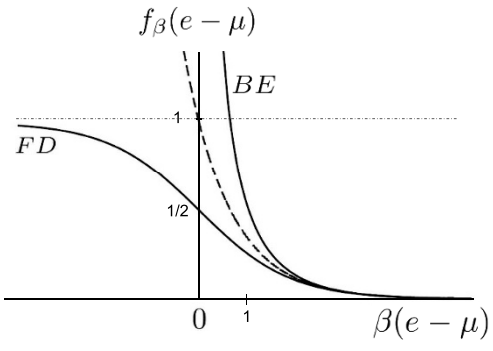


Figure 1: Quantum distribution functions of Fermi-Dirac f_β^{FD} (for fermions, lower curve) and of Bose-Einstein f_β^{BE} (for bosons, upper curve) as a function of the dimensionless variable $\beta(e-\mu)$; the dashed line intermediate curve represents the classical Boltzmann distribution $e^{-\beta(e-\mu)}$. In the right-hand side of the figure, corresponding to large negative values of μ, the particle number is small (the low density region) and the two distributions practically join the Boltzmann distribution. The system is said to be non-degenerate, or classical. As μ increases, we reach the central and left hand side of the figure, and the distributions become more and more different, reflecting the increasing gas degeneracy. For bosons, μ cannot be larger than the one-particle ground state energy, assumed to be zero in this case. The divergence observed for $\mu = 0$ corresponds to Bose-Einstein condensation. For fermions, the chemical potential μ can increase without limit, and for all the energy values, the distribution function tends towards 1 (but never exceeding 1 due to the Pauli exclusion principle).

(iii) For a boson system, the chemical potential cannot be larger than the lowest e_0 individual energy value, which we assumed to be zero. As μ tends towards zero through negative values and $-k_BT \ll \mu < 0$, the distribution function denominator becomes very small leading to very large populations of the corresponding states. The boson gas is then said to be "degenerate". On the other hand, for energies of the order or larger than μ, and as was the case for fermions, the boson distribution becomes practically equal to the Boltzmann distribution.

(iv) Finally, for situations intermediate between the extreme cases described above, the gas is said to be "partially degenerate".

3. Two-particle operators

For a two-particle symmetric operator $\widehat{G}$ we must use formula (C-16) of Chapter XV, which yields:

$$\left\langle \widehat{G} \right\rangle = \frac{1}{2} \sum_{i,j,k,l} \langle 1:u_i;\ 2:u_j |\, \widehat{g}(1,2) \,| 1:u_k;\ 2:u_l \rangle \ \left\langle a_i^\dagger a_j^\dagger a_l a_k \right\rangle \tag{28}$$

with:

$$\left\langle a_i^\dagger a_j^\dagger a_l a_k \right\rangle = \frac{1}{Z} \mathrm{Tr} \left\{ e^{-\beta\left(\widehat{H}-\mu\widehat{N}\right)} a_i^\dagger a_j^\dagger a_l a_k \right\} \tag{29}$$

As the exponential operator in the trace is diagonal in the Fock basis states $|n_1, .., n_i, .., n_j, ..\rangle$, this trace will be non-zero on the double condition that the states i and j associated with the creation operator be exactly the same as the states k and l associated with the annihilation operators, whatever the order. In other words, to get a non-zero trace, we must have either $i = l$ and $j = k$, or $i = k$ and $j = l$, or both.

3-a. Fermions

As two fermions cannot occupy the same quantum state, the product $a_i^\dagger a_j^\dagger$ is zero if $i = j$; we therefore assume $i \neq j$ which allows, using for ρ_{eq} expression (5) (which is a product), to perform independent calculations for the different modes. The case $i = l$ and $j = k$ yields, using the anticommutation relations:

$$a_i^\dagger a_j^\dagger a_i a_j = -\left(a_i^\dagger a_i\right)\left(a_j^\dagger a_j\right) \tag{30}$$

and the case $i = k$ and $j = l$ yields:

$$a_i^\dagger a_j^\dagger a_j a_i = +\left(a_i^\dagger a_i\right)\left(a_j^\dagger a_j\right) \tag{31}$$

We begin with term (30). As i and j are different, operators $a_i^\dagger a_i$ and $a_j^\dagger a_j$ act on different modes, which belong to different factors in the density operator (5). The average value of the product is thus simply the product of the average values:

$$\left\langle \left(a_i^\dagger a_i\right)\left(a_j^\dagger a_j\right) \right\rangle = \left\langle a_i^\dagger a_i \right\rangle \left\langle a_j^\dagger a_j \right\rangle \tag{32}$$

$$= f_\beta^{FD}(e_i - \mu) f_\beta^{FD}(e_j - \mu) \tag{33}$$

As for the second term (31), it is just the opposite of the first one. Consequently, we finally get:

$$\left\langle a_i^\dagger a_j^\dagger a_l a_k \right\rangle = \left[\delta_{ik}\delta_{jl} - \delta_{il}\delta_{jk}\right] \; f_\beta^{FD}(e_i - \mu) f_\beta^{FD}(e_j - \mu) \tag{34}$$

The first term on the right-hand side is called the direct term. The second one is the exchange term, and has a minus sign, as expected for fermions.

3-b. Bosons

For bosons, the operators a commute with each other.

α. Average value calculation

If $i \neq j$, a calculation, similar to the one we just did, yields:

$$\left\langle a_i^\dagger a_j^\dagger a_l a_k \right\rangle = \left[\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\right] \; f_\beta^{BE}(e_i - \mu) f_\beta^{BE}(e_j - \mu) \tag{35}$$

which differs in two ways from (34): the result now involves the Bose-Einstein distribution, and the exchange term is positive.

If $i = j$, only one individual state comes into a new calculation, which we now perform. Using for ρ_{eq} expression (5) we get, after summing as in (11) a geometric series:

$$\left\langle a_i^\dagger a_i^\dagger a_i a_i \right\rangle = \frac{1}{Z} \sum_{n_i=0}^{\infty} n_i \left(n_i - 1\right) \ \exp\left[-\beta n_i \left(e_i - \mu\right)\right] \prod_{k \neq i} \frac{1}{1 - e^{-\beta(e_k - \mu)}} \tag{36}$$

The sum appearing in this equation can be written as:

$$\begin{aligned} \sum_{n_i=0}^{\infty} n_i \left(n_i - 1\right) \ \exp\left[-\beta n_i \left(e_i - \mu\right)\right] &= \left[\frac{1}{\beta^2} \frac{\partial^2}{\partial \mu^2} - \frac{1}{\beta} \frac{\partial}{\partial \mu}\right] \sum_{n_i=0}^{\infty} \exp\left[-\beta n_i \left(e_i - \mu\right)\right] \\ &= \left[\frac{1}{\beta^2} \frac{\partial^2}{\partial \mu^2} - \frac{1}{\beta} \frac{\partial}{\partial \mu}\right] \frac{1}{1 - e^{-\beta(e_i - \mu)}} \end{aligned} \tag{37}$$

The first order derivative term yields:

$$-\frac{1}{\beta} \frac{\partial}{\partial \mu} \frac{1}{1 - e^{-\beta(e_i - \mu)}} = \frac{e^{-\beta(e_i - \mu)}}{\left[1 - e^{\beta(e_i - \mu)}\right]^2} \tag{38}$$

and the second order derivative term is:

$$\frac{1}{\beta^2} \frac{\partial^2}{\partial \mu^2} \frac{1}{1 - e^{-\beta(e_i - \mu)}} = -\frac{e^{-\beta(e_i - \mu)}}{\left[1 - e^{-\beta(e_i - \mu)}\right]^2} + 2 \frac{\left[e^{-\beta(e_i - \mu)}\right]^2}{\left[1 - e^{-\beta(e_i - \mu)}\right]^3} \tag{39}$$

Summing these two terms yields:

$$\frac{2 \left[e^{-\beta(e_i - \mu)}\right]^2}{\left[1 - e^{-\beta(e_i - \mu)}\right]^3} = \frac{2}{1 - e^{-\beta(e_i - \mu)}} \left[f_\beta^{BE}(e_i - \mu)\right]^2 \tag{40}$$

Multiplying by $1/\left[1 - e^{-\beta(e_i - \mu)}\right]$ the product at the end of the right-hand side of (36) yields the partition function Z, which cancels out the first factor $1/Z$. We are then left with:

$$\left\langle a_i^\dagger a_i^\dagger a_l a_k \right\rangle = 2 \ \delta_{il} \delta_{ik} \left[f_\beta^{BE}(e_i - \mu)\right]^2 \tag{41}$$

This result proves that (35) remains valid even in the case $i = j$.

β. Physical discussion: occupation number fluctuations

For two different physical states i and j, the average value $\left\langle a_i^\dagger a_j^\dagger a_j a_i \right\rangle$ for an ideal gas is simply equal to the product of the average values $\left\langle a_i^\dagger a_i \right\rangle = f_\beta^{BE}(e_i - \mu)$ and $\left\langle a_j^\dagger a_j \right\rangle = f_\beta^{BE}(e_j - \mu)$; this is a consequence of the total absence of interaction between the particles. The same is true for the average value $\left\langle a_i^\dagger a_j^\dagger a_i a_j \right\rangle$.

Now if $i = j$, we note the factor 2 in relation (41). As we now show, this factor leads to the presence of strong fluctuations associated with the operator $\widehat{n}_i$, the particle number in the state $|u_i\rangle$. The calculations shows that:

$$\left\langle (\widehat{n}_i)^2 \right\rangle = \langle a_i^\dagger a_i a_i^\dagger a_i \rangle = \langle a_i^\dagger a_i^\dagger a_i a_i \rangle + \langle a_i^\dagger a_i \rangle$$
$$= 2\left[f_\beta^{BE}(e_i - \mu) \right]^2 + f_\beta^{BE}(e_i - \mu) \tag{42a}$$

The square of the root mean square deviation Δn_i is therefore given by:

$$(\Delta n_i)^2 = \left\langle (\widehat{n}_i)^2 \right\rangle - \langle \widehat{n}_i \rangle^2 = \left[f_\beta^{BE}(e_i - \mu) \right]^2 + f_\beta^{BE}(e_i - \mu) = \langle \widehat{n}_i \rangle^2 + \langle \widehat{n}_i \rangle \tag{42b}$$

The fluctuations of this operator are therefore larger than its average value, which implies that the population of each state $|u_i\rangle$ is necessarily poorly defined[1] at thermal equilibrium. This is particularly true for large $\langle \widehat{n}_i \rangle$: in an ideal boson gas, a largely populated individual state is associated with a very large population fluctuation. This is due to the shape of the Bose-Einstein distribution (24), a decreasing exponential which is maximum at the origin: the most probable occupation number is always $n_i = 0$. Hence it is impossible to get a very large average $\langle \widehat{n}_i \rangle$ without introducing a distribution spreading over many n_i values. Complement H_{XV} (§ 4-a) discusses certain consequences of these fluctuations for an ideal gas. It also shows that as soon as a weak repulsive particle interaction is introduced, the fluctuations greatly diminish and almost completely disappear, since their presence would lead to a very large increase in the potential energy.

3-c. Common expression

To summarize, we can write in all cases:

$$\left\langle a_i^\dagger a_j^\dagger a_l a_k \right\rangle = \left[\delta_{ik}\delta_{jl} + \eta\ \delta_{il}\delta_{jk} \right]\ f_\beta(e_i - \mu) f_\beta(e_j - \mu) \tag{43}$$

with:

$$\begin{cases} \text{for fermions} & \eta = -1, \quad f_\beta = f_\beta^{FD} \\ \text{for bosons} & \eta = +1, \quad f_\beta = f_\beta^{BE} \end{cases} \tag{44}$$

As shown in relation (C-19) of Chapter XV, this average value is simply the matrix element $\langle 1 : u_k;\ 2 : u_l | \widehat{\rho}_2 | 1 : u_i;\ 2 : u_j \rangle$ of the two-particle reduced density operator. To get the general expression for the average of any symmetric two-particle operator, we simply use (43) in (28). Consequently, for independent particles, the average values of all these operators are simply expressed in terms of the quantum Fermi-Dirac and Bose-Einstein distribution functions.

Complement C_{XVI} will show how the Wick theorem allows generalizing these results to operators dealing with any number of particles.

[1] A physical observable is said to have a well defined value in a given quantum state if, in this state, its root mean square deviation is small compared to its average value.

4. Total number of particles

The operator $\widehat{N}$ corresponding to the total number of particles is given by the sum over all the individual states:

$$\widehat{N} = \sum_{i=1}^{\infty} a_i^{\dagger} a_i \tag{45}$$

and its average value is given by:

$$\left\langle \widehat{N} \right\rangle = \mathrm{Tr}\left\{ \widehat{N} \rho_{eq} \right\} = \sum_{i=1}^{\infty} f_\beta(e_i - \mu) \tag{46}$$

As f_β increases as a function of μ, the total number of particles is controlled (for fixed β) by the chemical potential.

4-a. Fermions

For the sake of simplicity, we study the ideal gas properties without taking into account the spin, which assumes that all particles are in the same spin state (the spin can easily be accounted for by adding the contributions of the different individual spin states). For a large physical system, the energy levels are very close and the discrete sum in (46) can be replaced by an integral. This leads to:

$$\left\langle \widehat{N} \right\rangle = N_{ig}^{FD}(\beta, \mu) \tag{47}$$

where the function $N_{ig}^{FD}(\beta, \mu)$ is defined as (the subscript ig stands for ideal gas):

$$N_{ig}^{FD}(\beta, \mu) = \frac{\mathcal{V}}{(2\pi)^3} \int \mathrm{d}^3 k \frac{1}{e^{\beta(e_k - \mu)} + 1} \tag{48}$$

Figure 2 shows the variations of the function $N_{ig}^{FD}(\beta, \mu)$ as a function of μ, for fixed values of β and the volume $\mathcal{V}$.

To deal with dimensionless quantities, one often introduces the "thermal wavelength" λ_T as:

$$\lambda_T = \hbar \sqrt{\frac{2\pi}{m k_B T}} = \hbar \sqrt{\frac{2\pi\beta}{m}} \tag{49}$$

We can then use in the integral of (48) the dimensionless variable:

$$\kappa = \frac{k \lambda_T}{2\sqrt{\pi}} \tag{50}$$

and write:

$$N_{ig}^{FD}(\beta, \mu) = \frac{\mathcal{V}}{(\lambda_T)^3} \; I_{3/2}(\beta\mu) \tag{51}$$

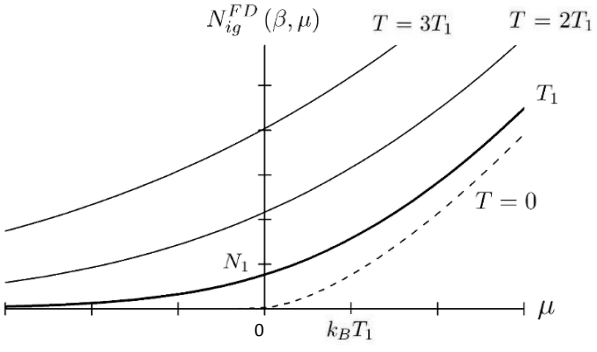


Figure 2: Variations of the particle number $N_{ig}^{FD}(\beta,\mu)$ for an ideal fermion gas, as a function of the chemical potential μ, and for different fixed temperatures T ($\beta = 1/(k_BT)$). For $T = 0$ (lower dashed line curve), the particle number is zero for negative values of μ, and proportional to $\mu^{3/2}$ for positive values of μ. For a non-zero temperature $T = T_1$ (thick line curve), the curve is above the previous one, and never goes to zero. Also shown are the curves obtained for temperatures twice ($T = 2T_1$) and three times ($T = 3T_1$) as large. The units chosen for the axes are the thermal energy k_BT_1 associated with the thick line curve, and the particle number $N_1 = \mathcal{V}/(\lambda_{T_1})^3$, where λ_{T_1} is the thermal wavelength at temperature T_1.
Largely negative values of μ correspond to the classical region where the fermion gas is not degenerate; the classical ideal gas equations are then valid to a good approximation. In the region where $\mu >> k_BT$, the gas is largely degenerate and a Fermi sphere shows up clearly in the momentum space; the total number of particles has only a slight temperature dependence and varies approximately as $\mu^{3/2}$.
This figure was kindly contributed by Geneviève Tastevin.

with[2]:

$$I_{3/2}(\beta\mu) = \pi^{-3/2}\int \mathrm{d}^3\kappa \frac{1}{e^{\kappa^2-\beta\mu}+1} = \frac{2}{\sqrt{\pi}}\int_0^\infty \mathrm{d}x \frac{\sqrt{x}}{e^{x-\beta\mu}+1} \tag{52}$$

where, in the second equality, we made the change of variable:

$$x = \kappa^2 \tag{53}$$

Note that the value of $I_{3/2}$ only depends on a dimensionless variable, the product $\beta\mu$.

If the particles have a spin 1/2, both contributions $\left\langle \widehat{N}_+ \right\rangle$ and $\left\langle \widehat{N}_- \right\rangle$ from the two spin states must be added to (46); in the absence of an external magnetic field, the

[2]The subscript 3/2 refers to the subscript used for more general functions $f_m(z)$, often called the Fermi functions in physics. They are defined by $f_m(z) = \sum_{l=1}^{\infty}(-1)^{l+1}z^l/l^m$, where z is the "fugacity" $z = e^{\beta\mu}$. Expanding in terms of z the function $1/\left[1+z^{-1}e^x\right] = ze^{-x}/\left[1+ze^{-x}\right]$ and using the properties of the Euler Gamma function, it can be shown that $I_{3/2}(\beta\mu) = f_{3/2}(z)$.

individual particle energies do not depend on their spin direction, and the total particle number is simply doubled:

$$\left\langle \widehat{N} \right\rangle = \left\langle \widehat{N}_+ \right\rangle + \left\langle \widehat{N}_- \right\rangle = 2N_{ig}^{FD}(\beta, \mu) \tag{54}$$

4-b. Bosons

For the sake of simplicity, we shall also start with spinless particles, but including several spin states is fairly straightforward. For bosons, we must use the Bose-Einstein distribution (24) and their average number is therefore:

$$\left\langle \widehat{N} \right\rangle = \sum_k f_\beta^{BE}(e_k - \mu) = \sum_k \frac{1}{e^{\beta(e_k-\mu)} - 1} \tag{55}$$

We impose periodic boundary conditions in a cubic box of edge length L. The lowest individual energy[3] is $e_k = 0$. Consequently, for expression (55) to be meaningful, μ must be negative or zero:

$$\mu \leq 0 \tag{56}$$

Two cases are possible, depending on whether the boson system is condensed or not.

α. *Non-condensed bosons*

When the parameter μ takes on a sufficiently negative value (much lower than the opposite of the individual energy e_1 of the first excited level), the function in the summation (55) is sufficiently regular for the discrete summation to be replaced by an integral (in the limit of large volumes). The average particle number is then written as:

$$\left\langle \widehat{N} \right\rangle = N_{ig}^{BE}(\beta, \mu) \tag{57}$$

with:

$$N_{ig}^{BE}(\beta, \mu) = \frac{\mathcal{V}}{(2\pi)^3} \int \mathrm{d}^3k \; \frac{1}{e^{\beta(e_k-\mu)} - 1} \tag{58}$$

Performing the same change of variables as above, this expression becomes:

$$N_{ig}^{BE}(\beta, \mu) = \frac{\mathcal{V}}{(\lambda_T)^3} K_{3/2}(\beta\mu) \tag{59}$$

with[4]:

$$K_{3/2}(\beta\mu) = \pi^{-3/2} \int \mathrm{d}^3\kappa \frac{1}{e^{\kappa^2-\beta\mu} - 1} = \frac{2}{\sqrt{\pi}} \int_0^\infty \mathrm{d}x \frac{\sqrt{x}}{e^{x-\beta\mu} - 1} \tag{60}$$

[3] Defining other boundary conditions on the box walls will lead in general to a non-zero ground state energy; choosing that value as the common origin for the energies and the chemical potential will leave the following computations unchanged.

[4] The subscript 3/2 refers to the subscript used for the functions $g_m(z)$, often called, in physics, the Bose functions (or the polylogarithmic functions). They are defined by the series $g_m(z) = \sum_{l=1}^{\infty} z^l/l^m$. The exact value of the number ζ defined in (61) is thus given by the series $\sum_{l=1}^{\infty} l^{-3/2}$.

The variations of $N_{ig}^{BE}(\beta,\mu)$ as a function of μ are shown in Figure 3. Note that the total particle number tends towards a limit $\zeta \times \left(\mathcal{V}/\lambda_T^3\right)$ as μ tends towards zero through negative values, where ζ is the number:

$$\zeta = K_{3/2}(0) = 2,612... \tag{61}$$

As the function increases with μ, we can write:

$$N_{ig}^{BE}(\beta,\mu) \leq \zeta \frac{\mathcal{V}}{(\lambda_T)^3} \tag{62}$$

There exists an insurmountable upper limit for the total particle number of a non-condensed ideal Bose gas.

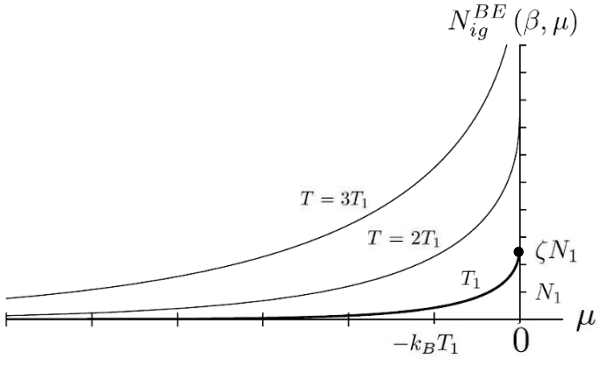


Figure 3: Variations of the total particle number $N_{ig}^{BE}(\beta,\mu)$ in a non-condensed ideal Bose gas, as a function of μ and for fixed $\beta = 1/(k_BT)$. The chemical potential is always negative, and the figure shows curves corresponding to several temperatures $T = T_1$ (thick line), $T = 2T_1$ and $T = 3T_1$. Units on the axes are the same as in Figure 2: the thermal energy k_BT_1 associated with curve $T = T_1$, and the particle number $N_1 = \mathcal{V}/(\lambda_{T_1})^3$, where λ_{T_1} is the thermal wavelength for this same temperature T_1. As the chemical potential tends towards zero, the particle numbers tend towards a finite value. For $T = T_1$, this value is equal to ζN_1 (shown as a dot on the vertical axis), where ζ is given by (61). This figure was kindly contributed by Geneviève Tastevin

β. Condensed bosons

As μ gets closer to zero, the population N_0 of the ground state becomes:

$$N_0 \equiv f_\beta^{BE}(e_0 - \mu) = \frac{1}{e^{-\beta\mu} - 1} \underset{\mu\to 0}{\sim} \frac{1}{\beta\,|\mu|} \tag{63}$$

This population diverges in the limit $\mu = 0$ and, when $|\mu|$ gets small enough, it can become arbitrarily large. It can, for example, become proportional[5] to the volume $\mathcal{V}$, in which case it adds a finite contribution $N_0/\mathcal{V}$ to the particle numerical density (particle number per unit volume) as $\mathcal{V} \to \infty$.

This particularity is limited to the ground state, which, in this case, plays a very different role than the other levels. Let us show, for example, that the first excited state population does not yield a similar effect. Assuming the system to be contained in a cubic box[6] of edge length L, the population of the first excited energy level $e_1 \sim \pi^2\hbar^2/(2mL^2)$ can be written as:

$$f_\beta^{BE}(e_1 - \mu) = \frac{1}{e^{\beta(e_1-\mu)} - 1} \underset{\mu\to 0}{\to} \frac{1}{\beta e_1} \sim \left(\frac{L}{\lambda_T}\right)^2 \tag{64}$$

(we assume the box to be large enough so that $L \gg \lambda_T$, which means $\beta e_1 \ll 1$); this population can therefore be proportional only to the square of L, i.e. to the volume to the power 2/3. It shows that this first excited level cannot make a contribution to the particle density in the limit $L \to \infty$; the same is true for all the other excited levels whose contributions are even smaller. The only arbitrary contribution to the density comes from the ground state.

This arbitrarily large value as $\mu \to 0$ obviously does not appear in relation (59), which predicts that the density $N_{ig}^{BE}(\beta, \mu)/\mathcal{V}$ is always less than a finite value as shown by (62). This is not surprising: as the population varies radically from the first energy level to the next, we can no longer compute the average particle number by replacing in (55) the discrete summation by an integral and a more precise calculation is necessary. Actually, only the ground state population must be treated separately, and the summation over all the excited states (of which none contributes to the density divergence) can still be replaced by an integral as before. Consequently, to get the total population of the physical system we simply add the integral on the right-hand side of (57) to the contribution N_0 of the ground level:

$$\left\langle \widehat{N} \right\rangle = N_{ig}^{BE}(\beta, 0) + N_0 \tag{65}$$

where N_0 is defined in (63).

As $\mu \to 0$, the total population of all the excited levels (others than the ground level) remains practically constant and equal to its upper limit (62); only the ground state has a continuously increasing population N_0, which becomes comparable to the total population of all the excited states when the right-hand sides of (63) and (62) are of the same order of magnitude:

$$\mu \gtrsim -k_B T \frac{\lambda_T^3}{\mathcal{V}} \Rightarrow N_0 \gtrsim \frac{\mathcal{V}}{\lambda_T^3} \tag{66}$$

(μ being of course always negative). When this condition is satisfied, a significant fraction of the particles accumulates in the individual ground level, which is said to have a

[5] The limit where $\mathcal{V} \to \infty$ while the density remains constant is often called the "thermodynamic limit".

[6] As above, we assume periodic conditions on the box walls. Another choice would be to impose zero values for the wave functions on the walls: the numerical coefficients of the individual energies would be changed, but not the line of reasoning.

"macroscopic population" (proportional to the volume). We can even encounter situations where the majority of the particles all occupy the same quantum state. This phenomenon is called "Bose-Einstein condensation" (it was predicted by Einstein in 1935, following Bose's studies of quantum statistics applicable to photons). It occurs when the total density n_{tot} reaches the maximum predicted by formula (62), that is:

$$n_{tot} = \frac{\zeta}{\lambda_T^3} \simeq \frac{2.612}{\lambda_T^3} \tag{67}$$

This condition means that the average distance between particles is of the order of the thermal wavelength λ_T.

Initially, Bose-Einstein condensation was considered to be a mathematical curiosity rather than an important physical phenomenon. Later on, people realized that it played an important role in superfluid liquid Helium 4, although this was a system with constantly interacting particles, hence far from an ideal gas. For a dilute gas, Bose-Einstein condensation was observed for the first time in 1995, and in a great number of later experiments.

5. Equation of state, pressure

The "equation of state" of a fluid at thermal equilibrium is the relation that links, for a given particle number N, its pressure P, volume V, and temperature $T = 1/k_B\beta$. We have just studied the variations of the total particle number. We shall now examine the pressure of a fermion or boson ideal gas.

5-a. Fermions

The grand canonical potential of a fermion ideal gas is given by (9). Equation (14) indicates that, for a system at thermal equilibrium, this grand potential is equal to the opposite of the product of the volume $\mathcal{V}$ and the pressure P. We thus have:

$$\begin{aligned} P\mathcal{V} &= k_B T \sum_k \ln \left[1 + e^{-\beta(e_k - \mu)}\right] \\ &= \frac{\mathcal{V}}{(2\pi)^3} k_B T \int \mathrm{d}^3 k \, \ln \left[1 + e^{-\beta(e_k - \mu)}\right] \end{aligned} \tag{68}$$

(where the second equality is valid in the limit of large volumes). Simplifying by $\mathcal{V} = L^3$, we get the pressure of a fermion system contained in a box of macroscopic dimension:

$$\begin{aligned} P &= (2\pi)^{-3} k_B T \int \mathrm{d}^3 k \, \ln \left[1 + e^{-\beta(e_k - \mu)}\right] \\ &= \frac{1}{\beta \lambda_T^3} I_{5/2}(\beta\mu) \end{aligned} \tag{69}$$

with:

$$\begin{aligned} I_{5/2}(\beta\mu) &= \pi^{-3/2} \int \mathrm{d}^3 \kappa \, \ln \left[1 + e^{\beta\mu - \kappa^2}\right] \\ &= \frac{2}{\sqrt{\pi}} \int_0^\infty \mathrm{d}x \, \sqrt{x} \, \ln \left[1 + e^{\beta\mu - x}\right] \end{aligned} \tag{70}$$

where x has been defined in (53).

To obtain the equation of state, we must find a relation between the pressure P, the volume $\mathcal{V}$, and the temperature T of the physical system, assuming the particle number to be fixed. We have, however, used the grand canonical ensemble (cf. Appendix VI), where the temperature is determined by the parameter β and the volume is fixed, but where the particle number can vary: its average value is a function of a parameter, the chemical potential μ (for fixed values of β and $\mathcal{V}$). Mathematically, the pressure P appears as a function of $\mathcal{V}$, T and μ and not as the function of $\mathcal{V}$, T and the particle number we were looking for. We can nevertheless vary μ, and obtain values of the pressure and particle number of the system and consequently explore, point by point, the equation of state in this parametric form. To obtain an explicit form of the equation of state would require the elimination of the chemical potential using both (47) and (69); there is generally no algebraic solution, and people just use the parametric form of the equation of state, which allows computing all the possible state variables. There also exists a "virial expansion" in powers of the fugacity $e^{\beta\mu}$, which allows the explicit elimination of μ at all the successive orders; its description is beyond the scope of this book.

5-b. Bosons

The pressure of an ideal boson gas is derived from the grand potential (12), taking into account its relation (14) to the pressure P and volume $\mathcal{V}$:

$$\begin{aligned} P\mathcal{V} &= -k_B T \sum_k \ln \left[1 - e^{-\beta(e_k-\mu)}\right] \\ &= -\frac{\mathcal{V}}{(2\pi)^3} k_B T \int d^3k \, \ln \left[1 - e^{-\beta(e_k-\mu)}\right] \end{aligned} \tag{71}$$

(the second relation being valid in the limit of large volumes). This leads to:

$$\begin{aligned} P &= -\frac{k_B T}{(2\pi)^3} \int d^3k \, \ln \left[1 - e^{-\beta(e_k-\mu)}\right] \\ &= \frac{1}{\beta\lambda_T^3} K_{5/2}(\beta\mu) \end{aligned} \tag{72}$$

with:

$$\begin{aligned} K_{5/2}(\beta\mu) &= -\pi^{-3/2} \int \mathrm{d}^3\kappa \, \ln \left[1 - e^{\beta\mu-\kappa^2}\right] \\ &= \frac{-2}{\sqrt{\pi}} \int_0^\infty \mathrm{d}x \, \sqrt{x} \, \ln \left[1 - e^{\beta\mu-x}\right] \end{aligned} \tag{73}$$

As $\mu \to 0$, the contribution P_0 of the ground level to the pressure written in (71) is:

$$P_0 = -\frac{k_B T}{\mathcal{V}} \ln \left[1 - e^{\beta\mu}\right] \simeq -\frac{k_B T}{\mathcal{V}} \ln \left[-\beta\mu\right] \tag{74}$$

When the chemical potential tends towards zero as in (66), it leads to:

$$P_0 \simeq \frac{k_B T}{\mathcal{V}} \ln \frac{\mathcal{V}}{\lambda_T^3} \tag{75}$$

which therefore goes to zero in the limit of large volumes. For a large system, the ground level contribution to the pressure remains negligible compared to that of all the other individual energy levels, whose number gets bigger as the system gets larger. Contrary to what we encountered for the average total particle number, the condensed particles' contribution to the pressure goes to zero in the limit of large volumes.

As we have seen for fermions, the equation of state must be obtained by eliminating the chemical potential μ between equations (72) yielding the pressure and (65) yielding the total particle number. As opposed to an ideal fermion gas, whose particle number and pressure increase without limit as μ and the density increase, the pressure in a boson system is limited. As soon as the system condenses, only the particle number in the individual ground state continues to grow, but not the pressure. In other words, the physical system acquires an infinite compressibility, and becomes a "marginally pathological" system (a system whose pressure decreases with its volume is unstable). This pathology comes, however, from totally neglecting the bosons' interactions. As soon as repulsive interactions are introduced, no matter how small, the compressibility will take on a finite value and the pathology will disappear.

This complement is a nice illustration of the simplifications incurred by the systematic use, in the calculations, of the creation and annihilation operators. We shall see in the following complements that these simplifications still occur when taking into account the interactions, provided we stay in the framework of the mean field approximation. Complement B_{XVI} will even show that for an interacting system studied without using this approximation, the ideal gas distribution functions are still somewhat useful for expressing the average values of various physical quantities.

Complement C_{XV}

Condensed boson system, Gross-Pitaevskii equation

1 Notation, variational ket 1643

1-a Hamiltonian 1644

1-b Choice of the variational ket (or trial ket) 1644

2 First approach 1645

2-a Trial wave function for spinless bosons, average energy 1645

2-b Variational optimization 1646

3 Generalization, Dirac notation 1648

3-a Average energy 1648

3-b Energy minimization 1649

3-c Gross-Pitaevskii equation 1650

4 Physical discussion 1651

4-a Energy and chemical potential 1651

4-b Healing length 1652

4-c Another trial ket: fragmentation of the condensate 1654

The Bose-Einstein condensation phenomenon for an ideal gas (no interaction) of N identical bosons was introduced in § 4-b-β of Complement B_{XV}. We show in the present complement how to describe this phenomenon when the bosons interact. We shall look for the ground state of this physical system within the mean field approximation, using a variational method (see Complement E_{XI}). After introducing in § 1 the notation and the variational ket, we study in § 2 spinless bosons, for which the wave function formalism is simple and the introduction of the creation and annihilation operators does not lead to any major computation simplifications. This will lead us to a first version of the Gross-Pitaevskii equation. We will then come back in § 3 to Dirac notation and the creation operators, to deal with the more general case where each particle may have a spin. Defining the Gross-Pitaevskii potential operator, we shall obtain a more general version of that equation. Finally, some properties of the Gross-Pitaevskii equation will be discussed in § 4, as well as the role of the chemical potential, the existence of a relaxation (or "healing") length, and the energetic consequences of "condensate fragmentation" (these terms will be defined in § 4-c).

1. Notation, variational ket

We first define the notation and the variational family of state vectors that will lead to relatively simple calculations for a system of identical interacting bosons.

1-a. Hamiltonian

The Hamiltonian operator $\widehat{H}$ we consider is the sum of operators for the kinetic energy $\widehat{H}_0$, the one-body potential energy $\widehat{V}_{\rm ext}$, and the interaction energy $\widehat{W}_{\rm int}$:

$$\widehat{H} = \widehat{H}_0 + \widehat{V}_{\rm ext} + \widehat{W}_{\rm int} \tag{1}$$

The first term $\widehat{H}_0$ is simply the sum of the individual kinetic energy operators associated with each of the particles q:

$$\widehat{H}_0 = \sum_q K_0(q) \tag{2}$$

where :

$$K_0(q) = \frac{\mathbf{P}_q^2}{2m} \tag{3}$$

($\mathbf{P}_q$ is the momentum of particle q). Similarly, $\widehat{V}_{\rm ext}$ is the sum of the external potential operators $V_1(\mathbf{R}_q)$, each depending on the position operator $\mathbf{R}_q$ of particle q:

$$\widehat{V}_{\rm ext} = \sum_{q=1}^{N} V_1(\mathbf{R}_q) \tag{4}$$

Finally, $\widehat{W}_{\rm int}$ is the sum of the interaction energy associated with all the pairs of particles:

$$\widehat{W}_{\rm int} = \frac{1}{2} \sum_{q \neq q'=1}^{N} W_2(\mathbf{R}_q, \mathbf{R}_{q'}) \tag{5}$$

(this summation can also be written as a sum over $q < q'$, while removing the prefactor $1/2$).

1-b. Choice of the variational ket (or trial ket)

Let us choose an arbitrary normalized quantum state $|\theta\rangle$:

$$\langle\theta\,|\theta\rangle = 1 \tag{6}$$

and call $a_\theta^\dagger$ the associated creation operator. The N-particle variational kets we consider are defined by the family of all the kets that can be written as:

$$\widetilde{|\Psi\rangle} = \frac{1}{\sqrt{N!}} \left[a_\theta^\dagger\right]^N |0\rangle \tag{7}$$

where $|\theta\rangle$ can vary, only constrained by (6). Consider a basis $\{|\theta_k\rangle\}$ of the individual state space whose first vector is $|\theta_1\rangle = |\theta\rangle$. Relation (A-17) of Chapter XV shows that this ket is simply a Fock state whose only non-zero occupation number is the first one:

$$\widetilde{|\Psi\rangle} = |n_1 = N, n_2 = 0, n_3 = 0, ...\rangle \tag{8}$$

An assembly of bosons that occupy the same individual state is called a "Bose-Einstein condensate".

Relation (8) shows that the kets $\widetilde{|\Psi\rangle}$ are normalized to 1. We are going to vary $|\theta\rangle$, and therefore $\widetilde{|\Psi\rangle}$, so as to minimize the average energy:

$$\widetilde{E} = \widetilde{\langle\Psi|}\widehat{H}\widetilde{|\Psi\rangle} \tag{9}$$

2. First approach

We start with a simple case where the bosons have no spin. We can then use the wave function formalism and keep the computations fairly simple.

2-a. Trial wave function for spinless bosons, average energy

Assuming one single individual state to be populated, the wave function $\Psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$ is simply the product of N functions $\theta(\mathbf{r})$:

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N) = \theta(\mathbf{r}_1)\ \theta(\mathbf{r}_2) ... \ \theta(\mathbf{r}_N) \tag{10}$$

with:

$$\theta(\mathbf{r}) = \langle \mathbf{r} \,|\theta\rangle \tag{11}$$

This wave function is obviously symmetric with respect to the exchange of all particles and can be used for a system of identical bosons.

In the position representation, each operator $K_0(q)$ defined by (3) corresponds to $\left(-\hbar^2/2m\right)\Delta_{\mathbf{r}_q}$, where $\Delta_{\mathbf{r}_q}$ is the Laplacian with respect to the position $\mathbf{r}_q$; consequently, we have:

$$\left\langle \widehat{H}_0 \right\rangle = \frac{-\hbar^2}{2m} \sum_{q=1}^{N} \int \mathrm{d}^3 r_1 ... \int \mathrm{d}^3 r_q ... \int \mathrm{d}^3 r_N \\ \times \theta^*(\mathbf{r}_1) ... \theta^*(\mathbf{r}_q) ... \theta^*(\mathbf{r}_N) \times \theta(\mathbf{r}_1) ... \ \Delta\theta(\mathbf{r}_q) ... \ \theta(\mathbf{r}_N) \tag{12}$$

In this expression, all the integral variables others than $\mathbf{r}_q$ simply introduce the square of the norm of the function $\theta(\mathbf{r})$, which is equal to 1. We are just left with one integral over $\mathbf{r}_q$, in which $\mathbf{r}_q$ plays the role of a dummy variable, and thus yields a result independent of q. Consequently, all the q values give the same contribution, and we can write:

$$\left\langle \widehat{H}_0 \right\rangle = \frac{-\hbar^2}{2m} N \int \mathrm{d}^3 r \ \theta^*(\mathbf{r}) \Delta\theta(\mathbf{r}) \tag{13}$$

As for the one-body potential energy, a similar calculation yields:

$$\left\langle \widehat{V}_{\text{ext}} \right\rangle = N \int \mathrm{d}^3 r \ \theta^*(\mathbf{r}) \ V_1(\mathbf{r}) \ \theta(\mathbf{r}) \tag{14}$$

Finally, the interaction energy calculation follows the same steps, but we must keep two integral variables instead of one. The final result is proportional to the number $N(N-1)/2$ of pairs of integral variables:

$$\left\langle \widehat{W}_2 \right\rangle = \frac{N(N-1)}{2} \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \ \theta^*(\mathbf{r}) \theta^*(\mathbf{r}') \ W_2(\mathbf{r}, \mathbf{r}') \ \theta(\mathbf{r}) \theta(\mathbf{r}') \tag{15}$$

The variational average energy $\widetilde{E}$ is the sum of these three terms:

$$\widetilde{E} = \left\langle \widehat{H}_0 \right\rangle + \left\langle \widehat{V}_{\text{ext}} \right\rangle + \left\langle \widehat{W}_2 \right\rangle \tag{16}$$

2-b. Variational optimization

We now optimize the energy we just computed, so as to determine the wave functions $\theta(\mathbf{r})$ corresponding to its minimum value.

α. Variation of the wave function

Let us vary the function $\theta(\mathbf{r})$ by a quantity:

$$\theta(\mathbf{r}) \Rightarrow \theta(\mathbf{r}) + e^{i\chi}\ \delta f(\mathbf{r}) \tag{17}$$

where $\delta f(\mathbf{r})$ is an infinitesimal function and χ an arbitrary number. A priori, $\delta f(\mathbf{r})$ must be chosen to take into account the normalization constraint (6), which forces the integral of the $\theta(\mathbf{r})$ modulus squared to remain constant. We can, however, use the Lagrange multiplier method (Appendix V) to impose this constraint. We therefore introduce the multiplier μ (we shall see in § 4-a that this factor can be interpreted as the chemical potential) and minimize the function:

$$\widetilde{F} = \widetilde{E} - \mu \int \mathrm{d}^3 r\ \theta^*(\mathbf{r})\theta(\mathbf{r}) \tag{18}$$

This allows considering the infinitesimal variation $\delta f(\mathbf{r})$ to be free of any constraint. The variation $\delta\widetilde{F}$ of the function $\widetilde{F}$ is now the sum of 4 variations, coming from the three terms of (16) and from the integral in (18). For example, the variation of $\left\langle \widehat{H}_0 \right\rangle$ yields:

$$\delta \left\langle \widehat{H}_0 \right\rangle = -\frac{\hbar^2}{2m} N \int \mathrm{d}^3 r\ \left[e^{-i\chi}\ \delta f^*(\mathbf{r})\ \Delta\theta(\mathbf{r}) + e^{i\chi}\ \theta^*(\mathbf{r})\ \Delta\delta f(\mathbf{r}) \right] \tag{19}$$

which is the sum of a term proportional to $e^{-i\chi}$ and another proportional to $e^{i\chi}$. This is true for all 4 variations and the total variation $\delta\widetilde{F}$ can be expressed as the sum of two terms:

$$\delta\widetilde{F} = \delta c_1 e^{-i\chi} + \delta c_2 e^{i\chi} \tag{20}$$

the first being the $\delta f^*(\mathbf{r})$ contribution and the second, that of $\delta f(\mathbf{r})$. Now if $\widetilde{F}$ is stationary, $\delta\widetilde{F}$ must be zero whatever the choice of χ, which is real. Choosing for example $\chi = 0$ imposes $\delta c_1 + \delta c_2 = 0$, and the choice $\chi = \pi/2$ leads (after multiplication by i) to $\delta c_1 - \delta c_2 = 0$. Adding and subtracting the two relations shows that both coefficients δc_1 and δc_2 must be zero. In other words, we can impose $\delta\widetilde{F}$ to be zero as just $\theta^*(\mathbf{r})$ varies but not $\theta(\mathbf{r})$ – or the opposite[1].

β. Stationary condition: Gross-Pitaevskii equation

We choose to impose the variation $\delta\widetilde{F}$ to be zero as only $\theta^*(\mathbf{r})$ varies and for $\chi = 0$. We must first add contributions coming from (13) and (14), then from (15). For this last contribution, we must add two terms, one coming from the variations due to $\theta^*(\mathbf{r})$, and the other from the variation due to $\theta^*(\mathbf{r}')$. These two terms only differ by the notation

[1] This means that the stationary condition may be found by varying indifferently the real or imaginary part of $\theta(\mathbf{r})$.

in the integral variable and are thus equal: we just keep one and double it. We finally add the term due to the variation of the integral in (18), and we get:

$$\delta\widetilde{F} = N \int \mathrm{d}^3 r \; \delta f^*(\mathbf{r}) \; \times \\ \left\{ \left[\frac{-\hbar^2}{2m}\Delta + V_1(\mathbf{r}) - \mu \right] + (N-1) \int \mathrm{d}^3 r' \; W_2(\mathbf{r},\mathbf{r}') \; \theta^*(\mathbf{r}') \; \theta(\mathbf{r}') \right\} \theta(\mathbf{r}) \qquad (21)$$

This variation must be zero for any value of $\delta f^*(\mathbf{r})$; this requires the function that multiplies $\delta f^*(\mathbf{r})$ in the integral to be zero, and consequently that $\theta(\mathbf{r})$ be the solution of the following equation, written for $\varphi(\mathbf{r})$:

$$\boxed{\left\{ \left[-\frac{\hbar^2}{2m}\Delta + V_1(\mathbf{r}) \right] + (N-1) \int \mathrm{d}^3 r' \; W_2(\mathbf{r},\mathbf{r}') \;\; |\varphi(\mathbf{r}')|^2 \right\} \varphi(\mathbf{r}) = \mu \; \varphi(\mathbf{r})} \qquad (22)$$

This is the time-independent Gross-Pitaevskii equation. It is similar to an eigenvalue Schrödinger equation, but with a potential term:

$$V_1(\mathbf{r}) + (N-1) \int \mathrm{d}^3 r' \; W_2(\mathbf{r},\mathbf{r}') \;\; |\varphi(\mathbf{r}')|^2 \qquad (23)$$

which actually contains the wave function φ in the integral over $\mathrm{d}^3 r'$; it is therefore a nonlinear equation. The physical meaning of the potential term in W_2 is simply that, in the mean field approximation, each particle moves in the mean potential created by all the others, each of them being described by the same wave function $\varphi(\mathbf{r}')$; the factor $(N-1)$ corresponds to the fact that each particle interacts with $(N-1)$ other particles. The Gross-Pitaevskii equation is often used to describe the properties of a boson system in its ground state (Bose-Einstein condensate).

γ. *Zero-range potential*

The Gross-Pitaevskii equation is often written in conjunction with an approximation where the particle interaction potential has a microscopic range, very small compared to the distances over which the wave function $\varphi(\mathbf{r})$ varies. We can then substitute:

$$W_2(\mathbf{r},\mathbf{r}') \Longrightarrow g \; \delta(\mathbf{r}-\mathbf{r}') \qquad (24)$$

where the constant g is called the "coupling constant"; such a potential is sometimes known as a "contact potential" or, in other contexts, a "Fermi potential". We then get:

$$\left[-\frac{\hbar^2}{2m}\Delta + V_1(\mathbf{r}) + (N-1)\, g\, |\varphi(\mathbf{r})|^2 \right] \varphi(\mathbf{r}) = \mu\varphi(\mathbf{r}) \qquad (25)$$

Whether in this form[2] or in its more general form (22), the equation includes a cubic term in $\varphi(\mathbf{r})$. It may render the problem difficult to solve mathematically, but it also is the source of many interesting physical phenomena. This equation explains, for example, the existence of quantum vortices in superfluid liquid helium.

[2]Strictly speaking, in what is generally called the Gross-Pitaevskii equation, the coupling constant g is replaced by $4\pi\hbar^2 a_0/m$, where a_0 is the "scattering length"; this length is defined when studying the collision phase shift $\delta_l(k)$ (Chapter VIII, § *C*), as the limit of $\delta_0(k) \sim -ka_0$ when $k \to 0$. This scattering length is a function of the interaction potential $W_2(\mathbf{r},\mathbf{r}')$, but generally not merely proportional to it, as opposed to the matrix elements of $W_2(\mathbf{r},\mathbf{r}')$. It is then necessary to make a specific demonstration for this form of the Gross-Pitaevskii equation, using for example the "pseudo-potential" method.

δ. Other normalization

Rather than normalizing the wave function $\varphi(\mathbf{r})$ to 1 in the entire space, one sometimes chooses a normalization taking into account the particle number by setting:

$$\int \mathrm{d}^3 r \, |\varphi(\mathbf{r})|^2 = N \tag{26}$$

This amounts to multiplying by $\sqrt{N}$ the wave function we have used until now. At each point $\mathbf{r}$ of space, the particle (numerical) density $n(\mathbf{r})$ is then given by:

$$n(\mathbf{r}) = |\varphi(\mathbf{r})|^2 \tag{27}$$

With this normalization, the factor $(N-1)$ in (25) is replaced by $(N-1)/N$, which can generally be taken equal to 1 for large N. The Gross-Pitaevskii equation then becomes:

$$\boxed{\left[-\frac{\hbar^2}{2m}\Delta + V_1(\mathbf{r}) + g\,|\varphi(\mathbf{r})|^2\right]\varphi(\mathbf{r}) = \mu\varphi(\mathbf{r})} \tag{28}$$

As already mentioned, we shall see in § 4-a that μ is simply the chemical potential.

3. Generalization, Dirac notation

We now go back to the previous line of reasoning, but in a more general case where the bosons may have spins. The variational family is the set of the N-particle state vectors written in (7). The one-body potential may depend on the position $\mathbf{r}$, and, at the same time, act on the spin (particles in a magnetic field gradient, for example).

3-a. Average energy

To compute the average energy value $\widetilde{\langle\Psi|}\widehat{H}\widetilde{|\Psi\rangle}$, we use a basis $\{|\theta_k\rangle\}$ of the individual state space, whose first vector is $|\theta_1\rangle = |\theta\rangle$.

Using relation (B-12) of Chapter XV, we can write the average value $\left\langle \widehat{H}_0 \right\rangle$ as:

$$\left\langle \widehat{H}_0 \right\rangle = \sum_{k,l} \langle\theta_k|\, K_0\, |\theta_l\rangle \;\; \widetilde{\langle\Psi|} a_k^\dagger a_l \widetilde{|\Psi\rangle} \tag{29}$$

Since $\widetilde{|\Psi\rangle}$ is a Fock state whose only non-zero population is that of the state $|\theta_1\rangle$, the ket $a_k^\dagger a_l \widetilde{|\Psi\rangle}$ is non-zero only if $l = 1$; it is then orthogonal to $\widetilde{|\Psi\rangle}$ if $k \neq 1$. Consequently, the only term left in the summation corresponds to $k = l = 1$. As the operator $a_1^\dagger a_1$ multiplies the ket by its population N, we get:

$$\left\langle \widehat{H}_0 \right\rangle = N \;\; \langle\theta_1|\, K_0\, |\theta_1\rangle \tag{30}$$

With the same argument, we can write:

$$\left\langle \widehat{V}_{\mathrm{ext}} \right\rangle = N \;\; \langle\theta_1|\, V_1\, |\theta_1\rangle \tag{31}$$

Using relation (C-16) of Chapter XV, we can express the average value of the interaction energy as[3]:

$$\left\langle \widehat{W_2} \right\rangle = \frac{1}{2} \sum_{k,l,m,n} \langle 1:\theta_k; 2:\theta_l |\, W_2(1,2) \,| 1:\theta_m; 2:\theta_n \rangle \ \widetilde{\langle \Psi |} a_k^\dagger a_l^\dagger a_n a_m \widetilde{| \Psi \rangle} \tag{32}$$

In this case, for the second matrix element to be non-zero, both subscripts m and n must be equal to 1 and the same is true for both subscripts k and l (otherwise the operator will yield a Fock state orthogonal to $\widetilde{|\Psi\rangle}$). When all the subscripts are equal to 1, the operator multiplies the ket $\widetilde{|\Psi\rangle}$ by $N(N-1)$. This leads to:

$$\left\langle \widehat{W_2} \right\rangle = \frac{N(N-1)}{2} \langle 1:\theta_1; 2:\theta_1 |\, W_2(1,2) \,| 1:\theta_1; 2:\theta_1 \rangle \tag{33}$$

The average interaction energy is therefore simply the product of the number of pairs $N(N-1)/2$ that can be formed with N particles and the average interaction energy of a given pair.

We can replace $|\theta_1\rangle$ by $|\theta\rangle$, since they are equal. The variational energy, obtained as the sum of (30), (31) and (33), then reads:

$$\widetilde{E} = N \ \langle \theta | \left[K_0 + V_1 \right] | \theta \rangle + \frac{N(N-1)}{2} \langle 1:\theta; 2:\theta |\, W_2(1,2) \,| 1:\theta; 2:\theta \rangle \tag{34}$$

3-b. Energy minimization

Consider a variation of $|\theta\rangle$:

$$|\theta\rangle \Rightarrow |\theta\rangle + e^{i\chi} |\delta\alpha\rangle \tag{35}$$

where $|\delta\alpha\rangle$ is an arbitrary infinitesimal ket of the individual state space, and χ an arbitrary real number. To ensure that the normalization condition (6) is still satisfied, we impose $|\delta\alpha\rangle$ and $|\theta\rangle$ to be orthogonal:

$$\langle \delta\alpha \,|\theta\rangle = 0 \tag{36}$$

so that $\langle \theta \,|\theta\rangle$ remains equal to 1 (to the first order in $|\delta\alpha\rangle$). Inserting (35) into (34) to obtain the variation $\mathrm{d}\widetilde{E}$ of the variational energy, we get the sum of two terms: the first one comes from the variation of the ket $|\theta\rangle$, and is proportional to $e^{i\chi}$; the second one comes from the variation of the bra $\langle\theta|$, and is proportional to $e^{-i\chi}$. The result has the form:

$$\delta\widetilde{E} = \delta c_1 e^{i\chi} + \delta c_2 e^{-i\chi} \tag{37}$$

The stationarity condition for $\widetilde{E}$ must hold for any arbitrary real value of χ. As before (§ 2-b-α), it follows that both δc_1 and δc_2 are zero. Consequently, we can impose the variation $\widetilde{E}$ to be zero as just the bra $\langle\theta|$ varies (but not the ket $|\theta\rangle$), or the opposite.

Varying only the bra, we get the condition:

$$\begin{aligned} 0 = N \ \langle \delta\alpha | \left[K_0 + V_1 - \mu \right] | \theta \rangle + \frac{N(N-1)}{2} \big[& \langle 1:\delta\alpha; 2:\theta |\, W_2(1,2) \,| 1:\theta; 2:\theta \rangle \\ & + \langle 1:\theta; 2:\delta\alpha |\, W_2(1,2) \,| 1:\theta; 2:\theta \rangle \big] \end{aligned} \tag{38}$$

[3]We use the simpler notation $W_2(1,2)$ for $W_2(\mathbf{R}_1, \mathbf{R}_2)$.

As the interaction operator $W_2(1,2)$ is symmetric, the last two terms within the bracket in this equation are equal. We get (after simplification by N):

$$0 = \langle \delta\alpha | \left[K_0 + V_1 - \mu \right] | \theta \rangle + (N-1) \langle 1 : \delta\alpha; 2 : \theta | \, W_2(1,2) \, | 1 : \theta; 2 : \theta \rangle \tag{39}$$

3-c. Gross-Pitaevskii equation

To deal with equation (39), we introduce the Gross-Pitaevskii operator V_{GP}^{θ}, defined as a one-particle operator whose matrix elements in an arbitrary basis $\{|u_i\rangle\}$ are given by:

$$\langle u_i | \, V_{GP}^{\theta} \, | u_j \rangle = (N-1) \langle 1 : u_i; 2 : \theta | \, W_2(1,2) \, | 1 : u_j; 2 : \theta \rangle \tag{40}$$

which leads to:

$$\langle v | \, V_{GP}^{\theta} \, | v' \rangle = (N-1) \langle 1 : v; 2 : \theta | \, W_2(1,2) \, | 1 : v'; 2 : \theta \rangle \tag{41}$$

where $|v\rangle$ and $|v'\rangle$ are two arbitrary one-particle kets – this can be shown by expanding these two kets on the basis $\{|u_i\rangle\}$ and using relation (40). Note that this potential operator does not include an exchange term; this term does not exist when the two interacting particles are in the same individual quantum state. Equation (39) then becomes:

$$0 = \langle \delta\alpha | \left[K_0 + V_1 + V_{GP}^{\theta} \right] | \theta \rangle \tag{42}$$

This stationarity condition must be verified for any value of the bra $\langle \delta\alpha|$, with only the constraint that it must be orthogonal to $|\theta\rangle$ (according to relation (36)). This means that the ket resulting from the action of the operator $\left[K_0 + V_1 + V_{GP}^{\theta} \right]$ on $|\theta\rangle$ must have zero components on all the vectors orthogonal to $|\theta\rangle$; its only non-zero component must be on the ket $|\theta\rangle$ itself, which means it is necessarily proportional to $|\theta\rangle$. In other words, $|\theta\rangle$ must be an eigenvector of that operator, with eigenvalue μ (real since the operator is Hermitian):

$$\left[K_0 + V_1 + V_{GP}^{\theta} \right] | \theta \rangle = \mu \, | \theta \rangle \tag{43}$$

We have just shown that the optimal value $|\varphi\rangle$ of $|\theta\rangle$ is the solution of the Gross-Pitaevskii equation:

$$\left[K_0 + V_1 + V_{GP}^{\varphi} \right] | \varphi \rangle = \mu \, | \varphi \rangle \tag{44}$$

which is a generalization of (28) to particles with spin, and is valid for one- or two-body arbitrary potentials. For each particle, the operator V_{GP}^{φ} represents the mean field created by all the others in the same state $|\varphi\rangle$.

Comment:

The Gross-Pitaevskii operator V_{GP}^{θ} is simply a partial trace over the second particle:

$$V_{GP}^{\theta}(1) = (N-1) \, \mathrm{Tr}_2 \left\{ P^{\theta}(2) W_2(1,2) \right\} \tag{45}$$

where $P^{\theta}(2)$ is the projection operator $P^{\theta}(2)$ of the state of particle 2 onto $|\theta\rangle$:

$$P^{\theta}(2) = \sum_k |1 : u_k\rangle \langle 1 : u_k| \otimes |2 : \theta\rangle \langle 2 : \theta| = \sum_k |1 : u_k; 2 : \theta\rangle \langle 1 : u_k; 2 : \theta| \tag{46}$$

To show this, let us compute the partial trace on the right-hand side of (45). To obtain this trace (Complement E_{III}, § 5-b), we choose for particle 2 a set of basis states $\{|\theta_n\rangle\}$ whose first vector $|\theta_1\rangle$ coincides with $|\theta\rangle$:

$$\langle u_i|\,\mathrm{Tr}_2\left\{P^\theta(2)W_2(1,2)\right\}|u_j\rangle = \sum_n \langle 1:u_i;2:\theta_n|\,P^\theta(2)W_2(1,2)\,|1:u_j;2:\theta_n\rangle \tag{47}$$

Replacing $P^\theta(2)$ by its value (46) yields the product of δ_{ik} (for the scalar product associated with particle 1) and δ_{n1} (for the one associated with particle 2). This leads to:

$$\langle u_i|\,\mathrm{Tr}_2\left\{P^\theta(2)W_2(1,2)\right\}|u_j\rangle = \langle 1:u_i;2:\theta|\,W_2(1,2)\,|1:u_j;2:\theta\rangle \tag{48}$$

which is simply the initial definition (40) of V^θ_{GP}. Relation (45) is therefore another possible definition for the Gross-Pitaevskii potential.

4. Physical discussion

We have established which conditions the variational wave function must obey to make the energy stationary, but we have yet to study the actual value of this energy. This will allow us to show that the parameter μ is in fact the chemical potential associated with the system of interacting bosons. We shall then introduce the concept of a relaxation (or "healing") length, and discuss the effect, on the final energy, of the fragmentation of a single condensate into several condensates, associated with distinct individual quantum states.

4-a. Energy and chemical potential

Since the ket $|\varphi\rangle$ is normalized, multiplying (44) by the bra $\langle\varphi|$ and by N, we get:

$$N\,\langle\varphi|\,[K_0+V_1+V^\varphi_{GP}]\,|\varphi\rangle = N\mu \tag{49}$$

We recognize the first two terms of the left-hand side as the average values of the kinetic energy and the external potential. As for the last term, using definition (41) for V^φ_{GP}, we can write it as:

$$N\,\langle\varphi|\,V^\varphi_{GP}\,|\varphi\rangle = N\,(N-1)\,\langle 1:\varphi;2:\varphi|\,W_2(1,2)\,|1:\varphi;2:\varphi\rangle \tag{50}$$

which is simply twice the potential interaction energy given in (33) when $|\theta_1\rangle = |\varphi\rangle$. This leads to:

$$N\mu = \langle H_0\rangle + \langle V_1\rangle + 2\left\langle\widehat{W}_2\right\rangle = \widetilde{E} + \left\langle\widehat{W}_2\right\rangle \tag{51}$$

To find the energy $\widetilde{E}$, note that $N\mu/2$ is the sum of $\left\langle\widehat{W}_2\right\rangle$ and of half the kinetic and external potential energies. Adding the missing halves, we finally get for $\widetilde{E}$:

$$\widetilde{E} = \frac{N}{2}\left[\mu + \langle\varphi|\,[K_0+V_1]\,|\varphi\rangle\right] \tag{52}$$

An advantage of this formula is to involve only one- (and not two-) particle operators, which simplifies the computations. The interaction energy is implicitly contained in the factor μ.

The quantity μ does not yield directly the average energy, but it is related to it, as we now show. Taking the derivative, with respect to N, of equation (34) written for $|\theta\rangle = |\varphi\rangle$, we get:

$$\frac{\mathrm{d}\widetilde{E}}{\mathrm{d}N} = \langle\varphi|\,[K_0 + V_1]\,|\varphi\rangle + \left(N - \frac{1}{2}\right)\langle\varphi,\varphi|\,W_2(1,2)\,|\varphi,\varphi\rangle \tag{53}$$

For large N, one can safely replace in this equation $(N - 1/2)$ by $(N - 1)$; after multiplication by N, we obtain a sum of average energies:

$$N\frac{\mathrm{d}\widetilde{E}}{\mathrm{d}N} = \langle H_0\rangle + \langle V_1\rangle + 2\left\langle\widehat{W}_2\right\rangle \tag{54}$$

Taking relation (51) into account, this leads to:

$$\frac{\mathrm{d}\widetilde{E}}{\mathrm{d}N} = \mu \tag{55}$$

We know (Appendix VI, § 2-b) that in the grand canonical ensemble, and at zero temperature, the derivative of the energy with respect to the particle number (for a fixed volume) is equal to the chemical potential. The quantity μ, introduced mathematically as a Lagrange multiplier, can therefore be simply interpreted as this chemical potential.

4-b. Healing length

The "healing length" is an important concept that characterizes the way a solution of the time-independent Gross-Pitaevskii equation reacts to a spatial constraint (for example, the solution can be forced to be zero along a wall, or along the line of a vortex core). We now calculate an approximate order of magnitude for this length.

Assuming the potential $V_1(\mathbf{r})$ to be zero in the region of interest, we divide equation (28) by $\varphi(\mathbf{r})$ and get:

$$-\frac{\hbar^2}{2m}\frac{\Delta\varphi(\mathbf{r})}{\varphi(\mathbf{r})} + g\,|\varphi(\mathbf{r})|^2 = \mu \tag{56}$$

Consequently, the left-hand side of this equation must be independent of $\mathbf{r}$. Let us assume $\varphi(\mathbf{r})$ is constant in an entire region of space where the density is n_0, independent of $\mathbf{r}$:

$$n_0 = |\varphi(\mathbf{r})|^2 \tag{57}$$

but constrained by the boundary conditions to be zero along its border. For the sake of simplicity, we shall treat the problem in one dimension, and assume $\varphi(\mathbf{r})$ only depends on the first coordinate x of $\mathbf{r}$; the wave function must then be zero along a plane (supposed to be at $x = 0$). We are looking for an order of magnitude of the distance ξ over which the wave function goes from a practically constant value to zero, i.e. for the spatial range of the wave function transition regime. In the region where $\varphi(\mathbf{r})$ is constant, relation (56) yields:

$$\mu = gn_0 \tag{58}$$

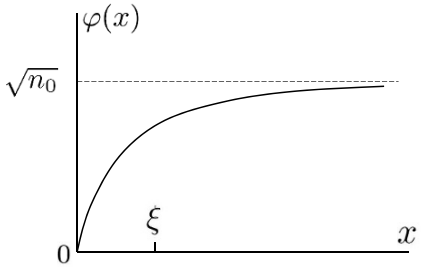


Figure 1: Variation as a function of the position x of the wave function $\varphi(x)$ in the vicinity of a wall (at $x = 0$) where it is forced to be zero. This variation occurs over a distance of the order of the healing length ξ defined in (61); the stronger the particle interactions, the shorter that length. As x increases, the wave function tends towards a constant plateau, of coordinate $\sqrt{n_0}$, represented as a dashed line.

On the other hand, in the whole region where $\varphi(\mathbf{r})$ has significantly decreased, and in particular close to the origin, we have:

$$-\frac{\hbar^2}{2m}\frac{\Delta\varphi(\mathbf{r})}{\varphi(\mathbf{r})} \sim \mu = gn_0 \tag{59}$$

In one dimension[4], we then get the differential equation:

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\varphi(x) \sim gn_0\varphi(x) \tag{60}$$

whose solutions are sums of exponential functions $e^{\pm ix/\xi}$, with:

$$\xi = \sqrt{\frac{\hbar^2}{2mgn_0}} \tag{61}$$

The solution that is zero for $x = 0$ is the difference between these two exponentials; it is proportional to $\sin(x/\xi)$, a function that starts from zero and increases over a characteristic length ξ. Figure 1 shows the wave function variation in the vicinity of the wall where it is forced to be zero.

The stronger the interactions, the shorter this "healing length" ξ; it varies as the inverse of the square root of the product of the coupling constant g and the density n_0. From a physical point of view, the healing length results from a compromise between the repulsive interaction forces, which try to keep the wave function as constant as possible in space, and the kinetic energy, which tends to minimize its spatial derivative (while the wave function is forced to be zero at $x = 0$); ξ is equal (except for a 2π coefficient) to the de Broglie wavelength of a free particle having a kinetic energy comparable to the repulsion energy gn_0 in the boson system.

[4] A more precise derivation can be given by verifying that $\varphi(x) = \sqrt{n_0}\,\tanh x/\xi\sqrt{2}$ is a solution of the one-dimensional equation (56).

4-c. Another trial ket: fragmentation of the condensate

We now show that repulsive interactions do stabilize a boson "condensate" where all the particles occupy the same individual state, as opposed to a "fragmented" state where some particles occupy a different state, which can be very close in energy. Instead of using a trial ket (7), where all the particles form a perfect Bose-Einstein condensate in a single quantum state $|\theta\rangle$, we can "fragment" this condensate by distributing the N particles in two distinct individual states. Consequently, we take a trial ket where N_a particles are in the state $|\theta_a\rangle$ and $N_b = N - N_a$ in the orthogonal state $|\theta_b\rangle$:

$$\widetilde{|\Psi\rangle} = \frac{1}{\sqrt{N_a! N_b!}} \left[a^\dagger_{\theta_a}\right]^{N_a} \left[a^\dagger_{\theta_b}\right]^{N_b} |0\rangle \tag{62}$$

We now compute the change in the average variational energy. In formula (29) giving the average kinetic energy, for the operator $a^\dagger_k a_l$ to yield a Fock state identical to $\widetilde{|\Psi\rangle}$, we must have either $k = l = a$, or $k = l = b$. This leads to:

$$\left\langle \widehat{H}_0 \right\rangle = N_a \ \langle\theta_a| K_0 |\theta_a\rangle + N_b \ \langle\theta_b| K_0 |\theta_b\rangle \tag{63}$$

The computation of the one-body potential energy is similar and leads to:

$$\left\langle \widehat{V}_{\text{ext}} \right\rangle = N_a \ \langle\theta_a| V_1 |\theta_a\rangle + N_b \ \langle\theta_b| V_1 |\theta_b\rangle \tag{64}$$

In both cases, the contributions of two populated states are proportional to their respective populations, as expected for energies involving a single particle.

As for the two-body interaction energy, we use again relation (32). It contains the operator $a^\dagger_k a^\dagger_l a_n a_m$, which will reconstruct the Fock state $\widetilde{|\Psi\rangle}$ in the following three cases:

- $k = l = m = n = a$ or b yields the contribution:

$$\frac{N_a (N_a - 1)}{2} \langle 1:\theta_a; 2:\theta_a| W_2(1,2) |1:\theta_a; 2:\theta_a\rangle + \frac{N_b (N_b - 1)}{2} \langle 1:\theta_b; 2:\theta_b| W_2(1,2) |1:\theta_b; 2:\theta_b\rangle \tag{65}$$

- $k = m = a$ and $l = n = b$, or $k = m = b$ and $l = n = a$; these two possibilities yield the same contribution (since the W_2 operator is symmetric), and the $1/2$ factor disappears, leading to the direct term:

$$N_a N_b \langle 1:\theta_a; 2:\theta_b| W_2(1,2) |1:\theta_a; 2:\theta_b\rangle \tag{66}$$

- Finally, $k = n = a$ and $l = m = b$, or $k = n = b$ and $l = m = a$, yield two contributions whose sum introduces the exchange term (here again without the factor $1/2$):

$$N_a N_b \langle 1:\theta_a; 2:\theta_b| W_2(1,2) |1:\theta_b; 2:\theta_a\rangle \tag{67}$$

The direct and exchange terms have been schematized in Figure 3 in Chapter XV (replacing $|u_i\rangle$ by $|\theta_a\rangle$, and $|u_j\rangle$ by $|\theta_b\rangle$), with the direct term on the left, and the exchange term on the right.

The variational energy can thus be written as:

$$\begin{aligned}\widetilde{E} &= N_a\ \left[\langle\theta_a|\left[K_0+V_1\right]|\theta_a\rangle\right] + N_b\ \left[\langle\theta_b|\left[K_0+V_1\right]|\theta_b\rangle\right] \\ &+ \frac{N_a\left(N_a-1\right)}{2}\langle 1:\theta_a;2:\theta_a|\,W_2(1,2)\,|1:\theta_a;2:\theta_a\rangle \\ &+ \frac{N_b\left(N_b-1\right)}{2}\langle 1:\theta_b;2:\theta_b|\,W_2(1,2)\,|1:\theta_b;2:\theta_b\rangle \\ &+ N_aN_b\langle 1:\theta_a;2:\theta_b|\,W_2(1,2)\,|1:\theta_a;2:\theta_b\rangle \\ &+ N_aN_b\langle 1:\theta_a;2:\theta_b|\,W_2(1,2)\,|1:\theta_b;2:\theta_a\rangle \end{aligned} \tag{68}$$

As above, the interaction between particles in the same state $|\theta_a\rangle$ contributes a term proportional to $N_a\left(N_a-1\right)/2$, the number of pairs of particles in that state; the same is true for the interaction term between particles in the same state $|\theta_b\rangle$. The direct term associated with the interaction between two particles in distinct states is proportional to N_aN_b, the number of such pairs. But to this direct term we must add an exchange term, also proportional to N_aN_b, corresponding to an additional interaction. This increased interaction is due to the bunching effect of two bosons in different quantum states, that will be discussed in more detail in § 3-b of Complement $\mathrm{A_{XVI}}$. As they are indistinguishable, two bosons occupying individual orthogonal states show correlations in their positions; this increases the probability of finding them at the same point in space. This increase does not occur when the two bosons occupy the same individual quantum state.

We now assume the diagonal matrix elements of $[K_0+V_1]$ between the two states $|\theta_a\rangle$ and $|\theta_b\rangle$ to be practically the same. For example, if these two states are the lowest energy levels of spinless particles in a cubic box of edge L, the corresponding energy difference is proportional to $1/L^2$ – hence very small in the limit of large L. We also assume all the matrix elements of $W_2(1,2)$ to be equal, which is the case if the (microscopic) range of the particle interaction potential is very small compared to the distances over which the wave functions of the two states vary. We can therefore replace in all the matrix elements the kets $|\theta_a\rangle$ and $|\theta_b\rangle$ by the same ket $|\theta\rangle$. Since $N_a+N_b=N$, we obtain:

$$\begin{aligned}\widetilde{E} &= N\ \langle\theta|\left[K_0+V_1\right]|\theta\rangle \\ &+ \frac{1}{2}\left[N_a\left(N_a-1\right)+N_b\left(N_b-1\right)+2N_aN_b\right]\langle 1:\theta;2:\theta|\,W_2(1,2)\,|1:\theta;2:\theta\rangle \\ &+ N_aN_b\langle 1:\theta;2:\theta|\,W_2(1,2)\,|1:\theta;2:\theta\rangle \end{aligned} \tag{69}$$

However:

$$N\left(N-1\right)=\left(N_a+N_b\right)\left(N_a+N_b-1\right)=N_a\left(N_a-1\right)+N_b\left(N_b-1\right)+2N_aN_b \tag{70}$$

so that:

$$\widetilde{E}=N\ \langle\theta|\left[K_0+V_1\right]|\theta\rangle+\frac{N\left(N-1\right)}{2}\langle 1:\theta;2:\theta|\,W_2(1,2)\,|1:\theta;2:\theta\rangle+\Delta\widetilde{E} \tag{71}$$

with:

$$\Delta\widetilde{E}=N_aN_b\ \langle 1:\theta;2:\theta|\,W_2(1,2)\,|1:\theta;2:\theta\rangle \tag{72}$$

We find again result (34), but with an additional term $\Delta\widetilde{E}$, the exchange term. Two cases are then possible, depending on whether the particle interactions are attractive or repulsive. In the first case, the fragmentation of the condensate lowers the energy and leads to a more stable state. Consequently, when the particle interactions are attractive, a condensate where only one individual state is occupied tends to split into two condensates, which might each split again, and so on. This means that the initial single condensate is unstable (we will come back and discuss this instability in § 2-b of Complement F_{XV} for the more general case of thermal equilibrium at non-zero temperature). On the contrary, for repulsive interactions the fragmentation increases the energy and leads to a less stable state: repulsive interactions therefore tend to stabilize the condensate in a single individual quantum state[5]. This result will be interpreted in § 3-b of Complement A_{XVI} in terms of changes of the particle position correlation function (bunching effect of bosons). As for the ideal gas, an intermediate case between the two previous ones, it is a marginal borderline case: adding any infinitesimal attractive interaction, no matter how small, destabilizes any condensate.

[5]We are discussing here the simple case of spinless bosons, contained in a box. When the bosons have several internal quantum states, and in other geometries, more complex situations may arise where the ground state is fragmented [4].

Complement D_{XV}

Time-dependent Gross-Pitaevskii equation

1 Time evolution 1657
1-a Functional variation 1658
1-b Variational computation: the time-dependent Gross-Pitaevskii equation 1659
1-c Phonons and Bogolubov spectrum 1660
2 Hydrodynamic analogy 1664
2-a Probability current 1664
2-b Velocity evolution 1665
3 Metastable currents, superfluidity 1667
3-a Toroidal geometry, quantization of the circulation, vortex . . 1667
3-b Repulsive potential barrier between states of different l . . . 1669
3-c Critical velocity, metastable flow 1671
3-d Generalization; topological aspects 1674

In this complement, we return to the calculations of Complement C_{XV}, concerning a system of bosons all in the same individual state. We now consider the more general case where that state is time-dependent. Using a variational method similar to the one we used in Complement C_{XV}, we shall study the time variations of the N-particle state vector. This amounts to using a time-dependent mean field approximation. We shall establish in § 1 a time-dependent version of the Gross-Pitaevskii equation, and explore some of its predictions such as the small oscillations associated with Bogolubov phonons. In § 2, we shall study local conservation laws derived from this equation for which we will give a hydrodynamic analogy, introducing a characteristic relaxation length. Finally, we will show in § 3 how the Gross-Pitaevskii equation predicts the existence of metastable flows and superfluidity.

1. Time evolution

We assume that the ket describing the physical system of N bosons can be written using relation (7) of Complement C_{XV}:

$$\widetilde{|\Psi(t)\rangle} = \frac{1}{\sqrt{N!}} \left[a_\theta^\dagger(t)\right]^N |0\rangle \tag{1}$$

but we now suppose that the individual ket $|\theta\rangle$ is a function of time $|\theta(t)\rangle$. The creation operator $a_\theta^\dagger(t)$ in the corresponding individual state is then time-dependent:

$$a_\theta^\dagger(t)\,|0\rangle = |\theta(t)\rangle \tag{2}$$

We will let the ket $|\theta(t)\rangle$ vary arbitrarily, as long as it remains normalized at all times:

$$\langle\theta(t)\,|\theta(t)\rangle = 1 \tag{3}$$

We are looking for the time variations of $|\theta(t)\rangle$ that will yield for $\widetilde{|\Psi(t)\rangle}$ variations as close as possible to those predicted by the exact N-particle Schrödinger equation. As the one-particle potential V_1 may also be time-dependent, it will be written as $V_1(t)$.

1-a. Functional variation

Let us introduce the functional of $|\Psi(t)\rangle$:

$$S\left[|\Psi(t)\rangle\right] = \left\{ \int_{t_0}^{t_1} \mathrm{d}t \, \langle \Psi(t)| \left[i\hbar \frac{d}{dt} - H(t) \right] |\Psi(t)\rangle \right\} + \frac{i\hbar}{2} \left[\langle \Psi(t_0) \, |\Psi(t_0)\rangle - \langle \Psi(t_1) \, |\Psi(t_1)\rangle \right] \tag{4}$$

It can be shown that this functional is stationary when $|\Psi(t)\rangle$ is solution of the exact Schrödinger equation (an explicit demonstration of this property is given in § 2 of Complement F_{XV}. If $|\Psi(t)\rangle$ belongs to a variational family, imposing the stationarity of this functional allows selecting, among all the family kets, the one closest to the exact solution of the Schrödinger equation. We shall therefore try and make this functional stationary, choosing as the variational family the set of kets $\widetilde{|\Psi(t)\rangle}$ written as in (1) where the individual ket $|\theta(t)\rangle$ is time-dependent.

As condition (3) means that the norm of $\widetilde{|\Psi(t)\rangle}$ remains constant, the second bracket in expression (4) must be zero. We now have to evaluate the average value of the Hamiltonian $H(t)$ that, actually, has been already computed in (34) of Complement C_{XV}:

$$\widetilde{\langle \Psi(t)|} \left[H(t) \right] \widetilde{|\Psi(t)\rangle} = N \, \langle \theta(t)| \left[K_0 + V_1(t) \right] |\theta(t)\rangle + \frac{N(N-1)}{2} \langle 1:\theta(t); 2:\theta(t)| \, W_2(1,2) \, |1:\theta(t); 2:\theta(t)\rangle \tag{5}$$

The only term left to be computed in (4) contains the time derivative.

This term includes the diagonal matrix element:

$$\widetilde{\langle \Psi(t)|} \left[i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \right] \widetilde{|\Psi(t)\rangle} = \frac{i\hbar}{N!} \langle 0| \left[a_\theta(t) \right]^N \sum_{k=0}^{N-1} \left[a_\theta^\dagger(t) \right]^k \frac{\mathrm{d}a_\theta^\dagger}{\mathrm{d}t} \left[a_\theta^\dagger(t) \right]^{N-k-1} |0\rangle \tag{6}$$

For an infinitesimal time dt, the operator $da_\theta^\dagger/dt$ is proportional to the difference $a_\theta^\dagger(t+dt) - a_\theta^\dagger(t)$, hence to the difference between two creation operators associated with two slightly different orthonormal bases. Now, for bosons, all the creation operators commute with each other, regardless of their associated basis. Therefore, in each term of the summation over k, we can move the derivative of the operator to the far right, and obtain the same result, whatever the value of k. The summation is therefore equal to N times the expression:

$$\frac{1}{N!} \langle 0| \left[a_\theta(t) \right]^N \left[a_\theta^\dagger(t) \right]^{N-1} \frac{\mathrm{d}a_\theta^\dagger}{\mathrm{d}t} |0\rangle \tag{7}$$

Now, we know that:

$$\left[a_\theta(t) \right]^{N-1} \left[a_\theta^\dagger(t) \right]^N |0\rangle = N! \, |1:\theta(t)\rangle = N! \left[a_\theta^\dagger(t) \right] |0\rangle \tag{8}$$

Using in (6) the bra associated with that expression, multiplied by N, we get:

$$\widetilde{\langle\Psi(t)|}\left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\right]\widetilde{|\Psi(t)\rangle} = i\hbar\ N\langle 0|\,a_\theta(t)\frac{\mathrm{d}a_\theta^\dagger}{\mathrm{d}t}\,|0\rangle = N\langle\theta(t)|\left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\right]|\theta(t)\rangle \tag{9}$$

Regrouping all these results, we finally obtain:

$$S\left[\widetilde{|\Psi(t)\rangle}\right] = -\int_{t_0}^{t_1}\mathrm{d}t\left\{N\ \langle\theta(t)|\left[[K_0+V_1(t)]-i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\right]|\theta(t)\rangle + \frac{N(N-1)}{2}\langle 1:\theta(t);2:\theta(t)|\,W_2(1,2)\,|1:\theta(t);2:\theta(t)\rangle\right\} \tag{10}$$

1-b. Variational computation: the time-dependent Gross-Pitaevskii equation

We now make an infinitesimal variation of $|\theta(t)\rangle$:

$$|\theta(t)\rangle \Rightarrow |\theta(t)\rangle + e^{i\chi}\,|\delta\theta(t)\rangle \tag{11}$$

in order to find the kets $|\theta(t)\rangle$ for which the previous expression will be stationary. As in the search for a stationary state in Complement $\mathrm{C_{XV}}$, we get variations coming from the infinitesimal ket $e^{i\chi}\,|\delta\theta(t)\rangle$ and others from the infinitesimal bra $e^{-i\chi}\,\langle\delta\theta(t)|$; as χ is chosen arbitrarily, the same argument as before leads us to conclude that each of these variations must be zero. Writing only the variation associated with the infinitesimal bra, we see that the stationarity condition requires $|\theta(t)\rangle$ to be a solution of the following equation, written for $|\varphi(t)\rangle$:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\,|\varphi(t)\rangle = [K_0+V_1(t)+V_{GP}^{\varphi}(t)]\,|\varphi(t)\rangle \tag{12}$$

The mean field operator $V_{GP}^{\varphi}(t)$ is defined as in relations (45) and (46) of Complement $\mathrm{C_{XV}}$ by a partial trace:

$$V_{GP}^{\varphi}(1,t) = (N-1)\,\mathrm{Tr}_2\left\{P^{\varphi(t)}(2)\ W_2(1,2)\right\} \tag{13}$$

where $P^{\varphi(t)}$ is the projector onto the ket $|\varphi(t)\rangle$:

$$P^{\varphi(t)} = |\varphi(t)\rangle\langle\varphi(t)| \tag{14}$$

As we take the trace over particle 2 whose state is time-dependent, the mean field is also time-dependent. Relation (12) is the general form of the time-dependent Gross-Pitaevskii equation.

Let us return, as in § 2 of Complement $\mathrm{C_{XV}}$, to the simple case of spinless bosons, interacting through a contact potential:

$$W_2(\mathbf{r},\mathbf{r}') = g\ \delta(\mathbf{r}-\mathbf{r}') \tag{15}$$

Using definition (13) of the Gross-Pitaevskii potential, we can compute its effect in the position representation, as in Complement $\mathrm{C_{XV}}$. The same calculations as in §§ 2-b-β and 2-b-γ of that complement allow showing that relation (12) becomes the Gross-Pitaevskii

time-dependent equation (N is supposed to be large enough to permit replacing $N-1$ by N):

$$i\hbar\frac{\partial}{\partial t}\varphi(\mathbf{r},t)=\left[-\frac{\hbar^2}{2m}\Delta+V_1(\mathbf{r},t)+Ng\left|\varphi(\mathbf{r},t)\right|^2\right]\varphi(\mathbf{r},t) \tag{16}$$

Normalizing the wave function $\varphi(\mathbf{r},t)$ to N:

$$\int \mathrm{d}^3r\ \left|\varphi(\mathbf{r},t)\right|^2=N \tag{17}$$

equation (16) simply becomes:

$$i\hbar\frac{\partial}{\partial t}\varphi(\mathbf{r},t)=\left[-\frac{\hbar^2}{2m}\Delta+V_1(\mathbf{r},t)+g\left|\varphi(\mathbf{r},t)\right|^2\right]\varphi(\mathbf{r},t) \tag{18}$$

Comment:

It can be shown that this time evolution does conserve the norm of $|\varphi(t)\rangle$, as required by (3). Without the nonlinear term of (16), it would be obvious since the usual Schrödinger equation conserves the norm. With the nonlinear term present, it will be shown in § 2-a that the norm is still conserved.

1-c. Phonons and Bogolubov spectrum

Still dealing with spinless bosons, we consider a uniform system, at rest, of particles contained in a cubic box of edge length L. The external potential $V_1(\mathbf{r})$ is therefore zero inside the box and infinite outside. This potential may be accounted for by forcing the wave function to be zero at the walls. In many cases, it is however more convenient to use periodic boundary conditions (Complement C_{XIV}, § 1-c), for which the wave function of the individual lowest energy state is simply a constant in the box. We thus consider a system in its ground state, whose Gross-Pitaevskii wave function is independent of $\mathbf{r}$:

$$\varphi(\mathbf{r},t)=\varphi_0(t)=\frac{1}{L^{3/2}}e^{-i\mu t/\hbar} \tag{19}$$

with a μ value that satisfies equation (16):

$$\mu=\frac{Ng}{L^3}=gn_0 \tag{20}$$

where $n_0=N/L^3$ is the system density. Comparing this expression with relation (58) of Complement C_{XV} allows us to identify μ with the ground state chemical potential. We assume in this section that the interactions between the particles are repulsive (see the comment at the end of the section):

$$g\geq 0 \tag{21}$$

α. Excitation propagation

Let us see which excitations can propagate in this physical system, whose wave function is no longer the function (19), uniform in space. We assume:

$$\varphi(\mathbf{r},t) = \varphi_0(t) + \delta\varphi(\mathbf{r},t) \tag{22}$$

where $\delta\varphi(\mathbf{r},t)$ is sufficiently small to be treated to first order. Inserting this expression in the right-hand side of (16), and keeping only the first-order terms, we find in the interaction term the first-order expression:

$$\begin{aligned}\delta\left[Ng\ \varphi^2(\mathbf{r},t)\varphi^*(\mathbf{r},t)\right] &= Ng\left[(2\varphi_0\delta\varphi)\,\varphi_0^* + \varphi_0^2\delta\varphi^*\right] \\ &= gn_0\left[2\delta\varphi + e^{-2i\mu t/\hbar}\delta\varphi^*\right]\end{aligned} \tag{23}$$

We therefore get, to first-order:

$$i\hbar\frac{\partial}{\partial t}\delta\varphi(\mathbf{r},t) = \left[-\frac{\hbar^2}{2m}\Delta + 2gn_0\right]\delta\varphi(\mathbf{r},t) + gn_0 e^{-2i\mu t/\hbar}\delta\varphi^*(\mathbf{r},t) \tag{24}$$

which shows that the evolution of $\delta\varphi(\mathbf{r},t)$ is coupled to that of $\delta\varphi^*(\mathbf{r},t)$. The complex conjugate equation can be written as:

$$i\hbar\frac{\partial}{\partial t}\delta\varphi^*(\mathbf{r},t) = \left[\frac{\hbar^2}{2m}\Delta - 2gn_0\right]\delta\varphi^*(\mathbf{r},t) - gn_0 e^{2i\mu t/\hbar}\delta\varphi(\mathbf{r},t) \tag{25}$$

We can make the time-dependent exponentials on the right-hand side disappear by defining:

$$\begin{aligned}\delta\varphi(\mathbf{r},t) &= \overline{\delta\varphi}(\mathbf{r},t)\ e^{-i\mu t/\hbar} \\ \delta\varphi^*(\mathbf{r},t) &= \overline{\delta\varphi}^*(\mathbf{r},t)\ e^{i\mu t/\hbar}\end{aligned} \tag{26}$$

This leads us to a differential equation with constant coefficients, which can be simply expressed in a matrix form:

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\overline{\delta\varphi}(\mathbf{r},t)\\ \overline{\delta\varphi}^*(\mathbf{r},t)\end{pmatrix} = \begin{pmatrix}\left[-\frac{\hbar^2}{2m}\Delta + gn_0\right] & gn_0\\ -gn_0 & \left[\frac{\hbar^2}{2m}\Delta - gn_0\right]\end{pmatrix}\begin{pmatrix}\overline{\delta\varphi}(\mathbf{r},t)\\ \overline{\delta\varphi}^*(\mathbf{r},t)\end{pmatrix} \tag{27}$$

where we have used definition (20) for μ to replace $2gn_0 - \mu$ by gn_0. If we now look for solutions having a plane wave spatial dependence:

$$\begin{aligned}\overline{\delta\varphi}(\mathbf{r},t) &= \widetilde{\delta\varphi}(\mathbf{k},t)\ e^{i\mathbf{k}\cdot\mathbf{r}} \\ \overline{\delta\varphi}^*(\mathbf{r},t) &= \widetilde{\delta\varphi}^*(\mathbf{k},t)\ e^{-i\mathbf{k}\cdot\mathbf{r}}\end{aligned} \tag{28}$$

the differential equation can be written as:

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\widetilde{\delta\varphi}(\mathbf{k},t)\\ \widetilde{\delta\varphi}^*(\mathbf{k},t)\end{pmatrix} = \begin{pmatrix}\left[\frac{\hbar^2k^2}{2m} + gn_0\right] & gn_0\\ -gn_0 & \left[-\frac{\hbar^2k^2}{2m} - gn_0\right]\end{pmatrix}\begin{pmatrix}\widetilde{\delta\varphi}(\mathbf{k},t)\\ \widetilde{\delta\varphi}^*(\mathbf{k},t)\end{pmatrix} \tag{29}$$

The eigenvalues $\hbar\omega(\mathbf{k})$ of this matrix satisfy the equation:

$$\left[\frac{\hbar^2k^2}{2m}+gn_0-\hbar\omega(\mathbf{k})\right]\left[-\frac{\hbar^2k^2}{2m}-gn_0-\hbar\omega(\mathbf{k})\right]+(gn_0)^2=0 \tag{30}$$

that is:

$$[\hbar\omega(\mathbf{k})]^2-\left[\frac{\hbar^2k^2}{2m}+gn_0\right]^2+(gn_0)^2=0 \tag{31}$$

The solution of this equation is:

$$\hbar\omega(\mathbf{k})=\sqrt{\left[\frac{\hbar^2k^2}{2m}+gn_0\right]^2-(gn_0)^2}=\sqrt{\frac{\hbar^2k^2}{2m}\left[\frac{\hbar^2k^2}{2m}+2gn_0\right]} \tag{32}$$

(the opposite value is also a solution, as expected since we calculate at the same time the evolution of $\widetilde{\delta\varphi}(t)$ and of its complex conjugate; we only use here the positive value). Setting:

$$k_0=\frac{2}{\hbar}\sqrt{gn_0\ m} \tag{33}$$

relation (32) can be written:

$$\omega(\mathbf{k})=\frac{\hbar}{2m}\sqrt{k^2\left(k^2+k_0^2\right)} \tag{34}$$

The spectrum given by (32) is plotted in Figure 1, where one sees the intermediate regime between the linear region at low energy, and the quadratic region at higher energy. It is called the "Bogolubov spectrum" of the boson system.

β. Discussion

Let us compute the spatial and time evolution of the particle density $n(\mathbf{r},\mathbf{t})$ when $\delta\varphi(\mathbf{r},t)$ obeys relation (28). The particle density at each point $\mathbf{r}$ of space is the sum of the densities associated with each particle, that is N times the squared modulus of the wave function $\varphi(\mathbf{r},t)$. To first-order in $\delta\varphi(\mathbf{r},t)$, we obtain:

$$\delta n(\mathbf{r},\mathbf{t})=N\left[\varphi_0(t)e^{-i\mu t/\hbar}\right]^*[\delta\varphi(\mathbf{r},t)]+\text{c.c.} \tag{35}$$

(where c.c. stands for "complex conjugate"). Using (26) and (28), we can finally write:

$$\begin{aligned}\delta n(\mathbf{r},\mathbf{t}) &= N\left[\varphi_0^*(t)e^{i\mu t/\hbar}\right]\left[e^{-i\mu t/\hbar}\ \widetilde{\delta\varphi}(\mathbf{k},0)\ e^{i[\mathbf{k}\cdot\mathbf{r}-\omega(\mathbf{k})t]}\right]+\text{c.c.}\\ &= \frac{N}{L^{3/2}}\left[\widetilde{\delta\varphi}(\mathbf{k},0)\ e^{i[\mathbf{k}\cdot\mathbf{r}-\omega(\mathbf{k})t]}+\text{c.c.}\right]\end{aligned} \tag{36}$$

Consequently, the excitation spectrum we have calculated corresponds to density waves propagating in the system with a phase velocity $\omega(\mathbf{k})/k$.

In the absence of interactions, $(g=k_0=0)$, this spectrum becomes:

$$\hbar\omega(\mathbf{k})=\frac{\hbar^2k^2}{2m} \tag{37}$$

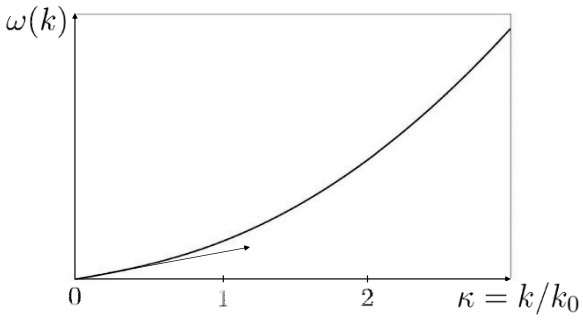


Figure 1: Bogolubov spectrum: variations of the function $\omega(\mathbf{k})$ given by equation (32) as a function of the dimensionless variable $\kappa = k/k_0$. When $\kappa \ll 1$, we get a linear spectrum (the arrow in the figure shows the tangent to the curve at the origin), whose slope is equal to the sound velocity c; when $\kappa \gg 1$, the spectrum becomes quadratic, as for a free particle.

which simply yields the usual quadratic relation for a free particle. Physically, this means that the boson system can be excited by transferring a particle from the individual ground state, with wave function $\varphi_0(\mathbf{r})$ and zero kinetic energy, to any state $\varphi_{\mathbf{k}}(\mathbf{r})$ having an energy $\hbar^2 k^2/2m$.

In the presence of interactions, it is no longer possible to limit the excitation to a single particle, which immediately transmits it to the others. The system's excitations become what we call "elementary excitations", involving a collective motion of all the particles, and hence oscillations in the density of the boson system. If $k \ll k_0$, we see from (34) that:

$$\omega(\mathbf{k}) \sim kc \tag{38}$$

where c is defined as:

$$c = \frac{\hbar}{2m} k_0 = \sqrt{\frac{gn_0}{m}} \tag{39}$$

For small values of k, the interactions have the effect of replacing the quadratic spectrum (37) by a linear spectrum. The phase velocity of all the excitations in this k value domain is a constant c. It is called the "sound velocity " in the interacting boson system, by analogy with a classical fluid where the sound wave dispersion relation is linear, as predicted by the Helmholtz equation. We shall see in § 3 that the quantity c plays a fundamental role in the computations related to superfluidity, especially for the critical velocity determination. If, on the other hand, $k \gg k_0$, the spectrum becomes:

$$\hbar\omega(\mathbf{k}) \sim \frac{\hbar^2 k^2}{2m} + gn_0 + .. \tag{40}$$

(the following corrections being in k_0^2/k^2, k_0^4/k^4, etc.). We find again, within a small correction, the free particle spectrum: exciting the system with enough energy allows

exciting individual particles almost as if they were independent. Figure 1 shows the complete variation of the spectrum (32), with the transition from the linear region at low energies, to the quadratic region at high energies.

Comment:

As we assumed the interactions to be repulsive in (21), the square roots in (32) and (39) are well defined. If the coupling constant g becomes negative, the sound velocity c will become imaginary, and, as seen from (32), so will the frequencies $\omega(k)$ (at least for small values of k). This will lead, for the evolution equation (29), to solutions that are exponentially increasing or decreasing in time, instead of oscillating. An exponentially increasing solution corresponds to an instability of the system. As already encountered in § 4-c of Complement C_{XV}, we see that a boson system becomes unstable in the presence of attractive interactions, however small they might be. In § 4-b of Complement H_{XV}, we shall see that this instability persists even for non-zero temperature. In a general way, an attractive condensate occupying a large region in space tends to collapse onto itself, concentrating into an ever smaller region. However, when it is confined in a finite region (as is the case for experiments where cold atoms are placed in a magneto-optical trap), any change in the wave function that brings the system closer to the instability also increases the gas energy; this results in an energy barrier, which allows the system of condensed attractive bosons to remain in a metastable state.

2. Hydrodynamic analogy

Let us return to the study of the time evolution of the Gross-Pitaevskii wave function and of the density variations $n(\mathbf{r},t)$, without assuming as in § 1-c that the boson system stays very close to uniform equilibrium. We will show that the Gross-Pitaevskii equation can take a form similar to the hydrodynamic equation describing a fluid's evolution. In this discussion, it is useful to normalize the Gross-Pitaevskii wave function to the particle number, as in equation (17). Equation (16) can then be written as:

$$i\hbar\frac{\partial}{\partial t}\varphi(\mathbf{r},t)=\left[-\frac{\hbar^2}{2m}\Delta+V_1(\mathbf{r},t)+gn(\mathbf{r},t)\right]\varphi(\mathbf{r},t) \tag{41}$$

where the local particle density $n(\mathbf{r},t)$ is given by:

$$n(\mathbf{r},t)=|\varphi(\mathbf{r},t)|^2 \tag{42}$$

2-a. Probability current

Since:

$$\frac{\partial}{\partial t}n(\mathbf{r},t)=\varphi^*(\mathbf{r},t)\frac{\partial}{\partial t}\varphi(\mathbf{r},t)+\varphi(\mathbf{r},t)\frac{\partial}{\partial t}\varphi^*(\mathbf{r},t) \tag{43}$$

the time variation of the density may be obtained by first multiplying (41) by $\varphi^*(\mathbf{r},t)$, then its complex conjugate by $\varphi(\mathbf{r},t)$, and then adding the two results; the potential terms in $V_1(\mathbf{r},t)$ and $g\ n(\mathbf{r},t)$ cancel out, and we get:

$$\frac{\partial}{\partial t}n(\mathbf{r},t)=-\frac{\hbar}{2mi}\left[\varphi^*(\mathbf{r},t)\Delta\varphi(\mathbf{r},t)-\varphi(\mathbf{r},t)\Delta\varphi^*(\mathbf{r},t)\right] \tag{44}$$

Let us now define a vector $\mathbf{J}(\mathbf{r},t)$ by:

$$\mathbf{J}(\mathbf{r},t) = \frac{\hbar}{2mi}\left[\varphi^*(\mathbf{r},t)\boldsymbol{\nabla}\varphi(\mathbf{r},t) - \varphi(\mathbf{r},t)\boldsymbol{\nabla}\varphi^*(\mathbf{r},t)\right] \tag{45}$$

If we compute the divergence of this vector, the terms in $\boldsymbol{\nabla}\varphi^* \cdot \boldsymbol{\nabla}\varphi$ cancel out and we are left with terms identical to the right-hand side of (44), with the opposite sign. This leads to the conservation equation:

$$\frac{\partial}{\partial t} n(\mathbf{r},t) + \boldsymbol{\nabla} \cdot \mathbf{J}(\mathbf{r},t) = 0 \tag{46}$$

$\mathbf{J}(\mathbf{r},t)$ is thus the probability current associated with our boson system. Integrating over all space, using the divergence theorem, and assuming $\varphi(\mathbf{r},t)$ (hence the current) goes to zero at infinity, we obtain:

$$\frac{\partial}{\partial t}\int \mathrm{d}^3r\ n(\mathbf{r},t) = \frac{\partial}{\partial t}\int \mathrm{d}^3r\ \ |\varphi(\mathbf{r},t)|^2 = 0 \tag{47}$$

This shows, as announced earlier, that the Gross-Pitaevskii equation conserves the norm of the wave function describing the particle system.

We now set:

$$\varphi(\mathbf{r},t) = \sqrt{n(\mathbf{r},\mathbf{t})}\ e^{i\alpha(\mathbf{r},t)} \tag{48}$$

The gradient of this function is written as:

$$\boldsymbol{\nabla}\varphi(\mathbf{r},t) = e^{i\alpha(\mathbf{r},\mathbf{t})}\left[\boldsymbol{\nabla}\sqrt{n(\mathbf{r},t)} + i\sqrt{n(\mathbf{r},t)}\boldsymbol{\nabla}\alpha(\mathbf{r},t)\right] \tag{49}$$

Inserting this result in (45), we get:

$$\mathbf{J}(\mathbf{r},t) = \frac{\hbar}{m} n(\mathbf{r},t)\boldsymbol{\nabla}\alpha(\mathbf{r},t) \tag{50}$$

or, defining the particle local velocity $\mathbf{v}(\mathbf{r},t)$ as the ratio of the current to the density:

$$\mathbf{v}(\mathbf{r},t) = \frac{\mathbf{J}(\mathbf{r},t)}{n(\mathbf{r},t)} = \frac{\hbar}{m}\boldsymbol{\nabla}\alpha(\mathbf{r},t) \tag{51}$$

We have defined a velocity field, similar to the velocity field of a fluid in motion in a certain region of space; this field velocity is irrotational (zero curl everywhere).

2-b. Velocity evolution

We now compute the time derivative of this velocity. Taking the derivative of (48), we get:

$$i\hbar\frac{\partial}{\partial t}\varphi(\mathbf{r},t) = e^{i\alpha(\mathbf{r},\mathbf{t})} i\hbar\frac{\partial}{\partial t}\sqrt{n(\mathbf{r},t)} - \hbar\sqrt{n(\mathbf{r},t)}e^{i\alpha(\mathbf{r},t)}\frac{\partial}{\partial t}\alpha(\mathbf{r},t) \tag{52}$$

so that we can isolate the time derivative of $\alpha(\mathbf{r},\mathbf{t})$ by the following combination:

$$i\hbar\left[\varphi^*(\mathbf{r},t)\frac{\partial}{\partial t}\varphi(\mathbf{r},t) - \varphi(\mathbf{r},t)\frac{\partial}{\partial t}\varphi^*(\mathbf{r},t)\right] = -2\hbar\ n(\mathbf{r},t)\frac{\partial}{\partial t}\alpha(\mathbf{r},t) \tag{53}$$

The left-hand side of this relation can be computed with the Gross-Pitaevskii equation (18) and its complex conjugate, as we now show. We first take the divergence of the gradient (49) to obtain the Laplacian:

$$\Delta\varphi(\mathbf{r},t) = \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\varphi(\mathbf{r},t) = e^{i\alpha(\mathbf{r},t)}\left[\Delta\sqrt{n(\mathbf{r},t)} + 2i\ \left(\boldsymbol{\nabla}\sqrt{n(\mathbf{r},t)}\right)\cdot\boldsymbol{\nabla}\alpha(\mathbf{r},t)\right.$$
$$\left. +i\ \sqrt{n(\mathbf{r},t)}\Delta\alpha(\mathbf{r},t) - \sqrt{n(\mathbf{r},t)}\,(\boldsymbol{\nabla}\alpha(\mathbf{r},t))^2\right] \tag{54}$$

We then insert the time derivative of $\varphi(\mathbf{r},t)$ given by the Gross-Pitaevskii equation (18) in the left-hand side of relation (53), which becomes:

$$-\frac{\hbar^2}{2m}\left[\varphi^*(\mathbf{r},t)\Delta\varphi(\mathbf{r},t) + \varphi(\mathbf{r},t)\Delta\varphi^*(\mathbf{r},t)\right] + 2\left[V_1(\mathbf{r},t) + g\ n(\mathbf{r},t)\right]|\varphi(\mathbf{r},t)|^2$$
$$= -\frac{\hbar^2}{2m}\left[2\sqrt{n(\mathbf{r},t)}\ \Delta\left(\sqrt{n(\mathbf{r},t)}\right) - 2n(\mathbf{r},t)\,(\boldsymbol{\nabla}\alpha(\mathbf{r},t))^2\right]$$
$$+ 2\Big[V_1(\mathbf{r},t) + g\ n(\mathbf{r},t)\Big]n(\mathbf{r},t) \tag{55}$$

This result must be equal to the right-hand side of (53). We therefore get, after dividing both sides by $-2n(\mathbf{r},t)$:

$$\hbar\frac{\partial}{\partial t}\alpha(\mathbf{r},\mathbf{t}) = \frac{\hbar^2}{2m}\left[\frac{1}{\sqrt{n(\mathbf{r},t)}}\Delta\left(\sqrt{n(\mathbf{r},t)}\right) - (\boldsymbol{\nabla}\alpha(\mathbf{r},t))^2\right] - \left[V_1(\mathbf{r},t) + g\ n(\mathbf{r},t)\right] \tag{56}$$

Using (51), we finally obtain the evolution equation for the velocity $\mathbf{v}(\mathbf{r},t)$:

$$m\frac{\partial}{\partial t}\mathbf{v}(\mathbf{r},t) = -\boldsymbol{\nabla}\left[V_1(\mathbf{r},t) + g\ n(\mathbf{r},t) + \frac{m\mathbf{v}^2(\mathbf{r},t)}{2} + \frac{\hbar^2}{2m}\frac{1}{\sqrt{n(\mathbf{r},t)}}\Delta\left(\sqrt{n(\mathbf{r},t)}\right)\right] \tag{57}$$

This equation looks like the classical Newton equation. Its right-hand side includes the sum of the forces corresponding to the external potential $V_1(\mathbf{r},t)$, and to the mean interaction potential with the other particles $gn(\mathbf{r},t)$; the third term in the gradient is the classical kinetic energy gradient[1] (as in Bernoulli's equation of classical hydrodynamics). The only purely quantum term is the last one, as shown by its explicit dependence on $\hbar^2$. It involves spatial derivatives of $n(\mathbf{r},t)$, and is only important if the relative variations $\delta n/n$ of the density occur over small enough distances (for example, this term is zero for a uniform density). This term is sometimes called "quantum potential", or "quantum pressure term" or, in other contexts, "Bohm potential". A frequently used approximation is to consider the spatial variations of $n(\mathbf{r},t)$ to be slow, which amounts to ignoring this quantum potential term: this is the so-called Thomas-Fermi approximation.

We have found for a system of N particles a series of properties usually associated with the wave function of a single particle, and in particular a local velocity directly proportional to its phase gradient[2]. The only difference is that, for the N-particle case,

[1] It is a "total derivative" term (the derivative describing, in a fluid, the motion of each particle). As the velocity field has a zero curl according to (51), a simple vector analysis calculation shows this term to be equal to $m\,(\mathbf{v}\cdot\boldsymbol{\nabla})\,\mathbf{v}$; it can therefore be accounted for by replacing on the left-hand side of (57) the partial derivative $\partial/\partial t$ by the total derivative $\mathrm{d}/\mathrm{d}t = \partial/\partial t + \mathbf{v}\cdot\boldsymbol{\nabla}$.

[2] The quantum potential is still present for a single particle, since making $g = 0$ in (57) does not change this potential. For $g = 0$, the Gross-Pitaevskii equation simply reduces to the standard Schrödinger equation, valid for a single particle.

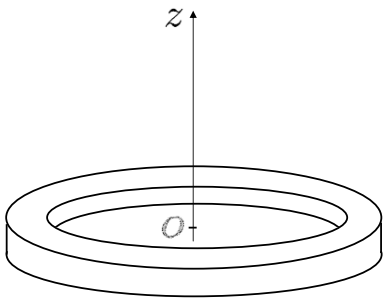


Figure 2: A repulsive boson gas is contained in a toroidal box. All the bosons are supposed to be initially in the same quantum state describing a rotation around the Oz axis. As we explain in the text, this rotation can only slow down if the system overcomes a potential energy barrier that comes from the repulsive interactions between the particles. This prevents any observable damping of the rotation over any accessible time scale; the fluid rotates indefinitely, and is said to be superfluid.

we must add to the external potential $V_1(\mathbf{r},t)$ a local interaction potential $gn(\mathbf{r},t)$, which does not significantly change the form of the equations but introduces some nonlinearity that can lead to completely new physical effects.

3. Metastable currents, superfluidity

Consider now a system of repulsive bosons contained in a toroidal box with a rotational axis Oz (Figure 2); the shape of the torus cross-section (circular, rectangular or other) is irrelevant for our argument and we shall use cylindrical coordinates r, φ and z. We first introduce solutions of the Gross-Pitaevskii equation that correspond to the system rotating inside the toroidal box, around the Oz axis. We will then show that these rotational states are metastable, as they can only relax towards lower energy rotational states by overcoming a macroscopic energy barrier: this is the physical origin of superfluidity.

3-a. Toroidal geometry, quantization of the circulation, vortex

To prevent any confusion with the azimuthal angle φ, we now call χ the Gross-Pitaevskii wave function. The time-independent Gross-Pitaevskii equation then becomes (in the absence of any potential except the wall potentials of the box):

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\chi}{\partial r}\right)+\frac{1}{r^2}\frac{\partial^2\chi}{\partial\varphi^2}+\frac{\partial^2\chi}{\partial z^2}\right]+g\left|\chi(\mathbf{r})\right|^2\chi(\mathbf{r})=\mu\chi(\mathbf{r}) \tag{58}$$

We look for solutions of the form:

$$\chi_l(\mathbf{r})=u_l(r,z)e^{il\varphi} \tag{59}$$

where l is necessarily an integer (otherwise the wave function would be multi-valued). Such a solution has an angular momentum with a well defined component along Oz, equal to $l\hbar$ per atom. Inserting this expression in (58), we obtain the equation for $u_l(r,z)$:

$$-\frac{\hbar^2}{2m}\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial u_l(r,z)}{\partial r}\right)+\frac{\partial^2 u_l(r,z)}{\partial z^2}\right]+\left[g\left|u_l(r,z)\right|^2+\frac{l^2\hbar^2}{2mr^2}\right]u_l(r,z) = \mu\ u_l(r,z) \tag{60}$$

which must be solved with the boundary conditions imposed by the torus shape to obtain the ground state (associated with the lowest value of μ). The term in $l^2\hbar^2/2mr^2$ is simply the rotational kinetic energy around Oz. If the tore radius R is very large compared to the size of its cross-section, the term l^2/r^2 may, to a good approximation, be replaced by the constant l^2/R^2. It follows that the same solution of (60) is valid for any value of l as long as the chemical potential is increased accordingly. Each value of the angular momentum thus yields a ground state and the larger l, the higher the corresponding chemical potential. All the coefficients of the equation being real, we shall assume, from now on, the functions $u_l(r,z)$ to be real.

As the wave function is of the form (59), its phase only depends on φ, and expression (51) for the fluid velocity is written as:

$$\mathbf{v}=\frac{1}{r}\frac{l\hbar}{m}\mathbf{e}_\varphi \tag{61}$$

where $\mathbf{e}_\varphi$ is the tangential unit vector (perpendicular both to $\mathbf{r}$ and the Oz axis). Consequently, the fluid rotates along the toroidal tube, with a velocity proportional to l. As $\mathbf{v}$ is a gradient, its circulation along a closed loop "equivalent to zero" (i.e. which can be contracted continuously to a point) is zero. If the closed loop goes around the tore, the path is no longer equivalent to zero and its circulation may be computed along a circle where r and z remain constant, and φ varies from 0 to 2π; as the path length equals $2\pi r$, we get:

$$\int \mathbf{v}\cdot \mathrm{d}\mathbf{s}=\pm l\frac{2\pi\hbar}{m} \tag{62}$$

(with a + sign if the rotation is counterclockwise and a − sign in the opposite case). As l is an integer, the velocity circulation around the center of the tore is quantized in units of h/m. This is obviously a pure quantum property (for a classical fluid, this circulation can take on a continuous set of values).

To simplify the calculations, we have assumed until now that the fluid rotates as a whole inside the toroidal ring. More complex fluid motions, with different geometries, are obviously possible. An important case, which we will return to later, concerns the rotation around an axis still parallel to Oz, but located inside the fluid. The Gross-Pitaevskii wave function must then be zero along a line inside the fluid itself, which thus contains a singular line. This means that the phase may change by 2π as one rotates around this line. This situation corresponds to what is called a "vortex", a little swirl of fluid rotating around the singular line, called the "vortex core line". As the circulation of the velocity only depends on the phase change along the path going around the vortex core, the quantization relation (62) remains valid. Actually, from a historical point of view, the Gross-Pitaevskii equation was first introduced for the study of superfluidity and the quantization of the vortices circulation.

3-b. Repulsive potential barrier between states of different l

A classical rotating fluid will always come to rest after a certain time, due to the viscous dissipation at the walls. In such a process, the macroscopic rotational kinetic energy of the whole fluid is progressively degraded into numerous smaller scale excitations, which end up simply heating the fluid. Will a rotating quantum fluid of repulsive bosons, described by a wave function $\chi_l(\mathbf{r})$, behave in the same way? Will it successively evolve towards the state $\chi_{l-1}(\mathbf{r})$, then $\chi_{l-2}(\mathbf{r})$, etc., until it comes to rest in the state $\chi_0(\mathbf{r})$?

We have seen in § 4-c of Complement $\mathrm{C_{XV}}$ that, to avoid the energy cost of fragmentation, the system always remains in a state where all the particles occupy the same quantum state. This is why we can use the Gross-Pitaevskii equation (18).

α. *A simple geometry*

Let us first assume that the wave function $\chi(\mathbf{r},t)$ changes smoothly from $\chi_l(\mathbf{r})$ to $\chi_{l'}(\mathbf{r})$ according to:

$$\chi(\mathbf{r},t) = c_l(t)\,\chi_l(\mathbf{r}) + c_{l'}(t)\,\chi_{l'}(\mathbf{r}) \tag{63}$$

where the modulus of $c_l(t)$ decreases with time from 1 to 0, whereas $c_{l'}(t)$ does the opposite. Normalization imposes that at all times t:

$$|c_l(t)|^2 + |c_{l'}(t)|^2 = 1 \tag{64}$$

In such a state, let us show that the numerical density $n(r,\varphi,z;t)$ now depends on φ (this was not the case for either states l or l' separately). The transverse dependence of the density as a function of the variables r and z, is barely affected[3]. The variations of $n(r,\varphi,z;t)$ are given by:

$$\begin{aligned} n(r,\varphi,z;t) &= \left| c_l(t)\,u_l(r,z)e^{il\varphi} + c_{l'}(t)\,u_{l'}(r,z)e^{il'\varphi} \right|^2 \\ &= |c_l(t)|^2\,|u_l(r,z)|^2 + |c_{l'}(t)|^2\,|u_{l'}(r,z)|^2 \\ &\quad + c_l(t)\,c_{l'}^*(t)\,u_l(r,z)u_{l'}(r,z)e^{i(l-l')\varphi} + \text{c.c.} \end{aligned} \tag{65}$$

where c.c. stands for the complex conjugate of the preceding factor. The first two terms are independent of φ, and are just a weighted average of the densities associated with each of the states l and l'. The last term oscillates as a function of φ with an amplitude $|c_l(t)| \times |c_{l'}(t)|$, which is only zero if one of the two coefficients $c_l(t)$ or $c_{l'}(t)$ is zero. Calling φ_l the phase of the coefficient $c_l(t)$, this last term is proportional to:

$$c_l(t)\,c_{l'}^*(t)\,e^{i(l-l')\varphi} + \text{c.c.} = 2\,|c_l(t)| \times |c_{l'}(t)| \cos\left[(l-l')\,\varphi + \varphi_l - \varphi_{l'}\right] \tag{66}$$

Whatever the phases of the two coefficients $c_l(t)$ and $c_{l'}(t)$, the cosine will always oscillate between -1 and 1 as a function of φ. Adjusting those phases, one can deliberately change the value of φ for which the density is maximum (or minimum), but this will always occur somewhere on the circle. Superposing two states necessarily modulates the density.

Let us evaluate the consequences of this density modulation on the internal repulsive interaction energy of the fluid. As we did in relation (15), we use for the interaction

[3] or not at all, if we suppose the functions $u_l(r,z)$ and $u_{l'}(r,z)$ to be equal.

energy the zero range potential approximation, and insert it in expression (15) of Complement C_{XV}. Taking into account the normalization (17) of the wave function, we get:

$$\left\langle \widehat{W}_2 \right\rangle_\chi = \frac{g}{2} \int \mathrm{d}^3 r \ \left|\chi\left(\mathbf{r},t\right)\right|^4 = \frac{g}{2} \int_0^\infty \mathrm{d}r \int_0^{2\pi} \mathrm{d}\varphi \int_{-\infty}^{+\infty} \mathrm{d}z \ \left[n(r,\varphi,z;t)\right]^2 \tag{67}$$

We must now include the square of (65) in this expression, which will yield several terms. The first one, in $\left|c_l\left(t\right)\right|^4$, leads to the contribution:

$$\left|c_l\left(t\right)\right|^4 \left\langle \widehat{W}_2 \right\rangle_l \tag{68}$$

where $\left\langle \widehat{W}_2 \right\rangle_l$ is the interaction energy for the state $\chi_l(\mathbf{r})$. The second contribution is the similar term for the state l', and the third one, a cross term in $2\left|c_l\left(t\right)\right|^2 \left|c_{l'}\left(t\right)\right|^2$. Assuming, to keep things simple, that the densities associated with the states l and l' are practically the same, the sum of these three terms is just:

$$\left[\left|c_l\left(t\right)\right|^2 + \left|c_{l'}\left(t\right)\right|^2\right]^2 \left\langle \widehat{W}_2 \right\rangle_l = \left\langle \widehat{W}_2 \right\rangle_l \tag{69}$$

Up to now, the superposition has had no effect on the repulsive internal interaction energy. As for the cross terms between the terms independent of φ in (65) and the terms in $e^{\pm i\left(l-l'\right)\varphi}$, they will cancel out when integrated over φ. We are then left with the cross terms in $e^{\pm i\left(l-l'\right)\varphi} \times e^{\mp i\left(l-l'\right)\varphi}$, whose integral over φ yields:

$$2\left|c_l\left(t\right)\right|^2 \left|c_{l'}\left(t\right)\right|^2 \left|u_l(r,z)\right|^2 \left|u_{l'}(r,z)\right|^2 \tag{70}$$

Assuming as before that the densities associated with the states l and l' are practically the same, we obtain, after integration over r and z:

$$2\left|c_l\left(t\right)\right|^2 \left|c_{l'}\left(t\right)\right|^2 \left\langle \widehat{W}_2 \right\rangle_l \tag{71}$$

Adding (69), we finally obtain:

$$\left\langle \widehat{W}_2 \right\rangle_\chi = \left[1 + 2\left|c_l\left(t\right)\right|^2 \left|c_{l'}\left(t\right)\right|^2\right] \left\langle \widehat{W}_2 \right\rangle_l \tag{72}$$

We have shown that the density modulation associated with the superposition of states always increases the internal repulsion energy: this modulation does lower the energy in the low density region, but the increase in the high energy region outweighs the decrease (since the repulsive energy is a quadratic function of the density). The internal energy therefore varies between $\left\langle \widehat{W}_2 \right\rangle_l$ and the maximum $(3/2)\left\langle \widehat{W}_2 \right\rangle_l$, reached when the moduli of $c_l\left(t\right)$ and $c_{l'}\left(t\right)$ are both equal to $1/\sqrt{2}$.

β. Other geometries, different relaxation channels

There are many other ways for the Gross-Pitaevskii wave function to go from one rotational state to another. We have limited ourselves to the simplest geometry to introduce the concept of energy barriers with minimal mathematics. The fluid could

transit, however, through more complex geometries, such as the frequently observed creation of a vortex on the wall, the little swirl we briefly talked about at the end of § 3-a. A vortex introduces a 2π phase shift around a singular line along which the wave function is zero. Once the vortex is created, and contrary to what was the case in (62), the velocity circulation along a loop going around the torus is no longer independent of its path: it will change by $2\pi\hbar/m$ depending on whether the vortex is included in the loop or not. Furthermore, as the vortex moves in the fluid from one wall to another, it can be shown that the proportion of fluid conserving the initial circulation decreases while the proportion having a circulation where the quantum number l differs by one unit increases. Consequently, this vortex motion changes progressively the rotational angular momentum. Once the vortex has vanished on the other wall, the final result is a decrease by one unit of the quantum number l associated with the fluid rotation.

The continuous passage of vortices from one wall to another therefore yields another mechanism that allows the angular moment of the fluid to decrease. The creation of a vortex, however, is necessarily accompanied by a non-uniform fluid density, described by the Gross-Pitaevskii equation (this density must be zero along the vortex core). As we have seen above, this leads to an increase in the average repulsive energy between the particles (the fluid elastic energy). This process thus also encounters an energy barrier (discussed in more detail in the conclusion). In other words, the creation and motion of vortices provide another "relaxation channel" for the fluid velocity, with its own energy barrier, and associated relaxation time.

Many other geometries can be imagined for changing the fluid flow. Each of them is associated with a potential barrier, and therefore a certain lifetime. The relaxation channel with the shortest lifetime will mainly determine the damping of the fluid velocity, which may take, in certain cases, an extraordinarily long time (dozens of years or more), hence the name of "superfluid".

3-c. Critical velocity, metastable flow

For the sake of simplicity, we will use in our discussion the simple geometry of § 3-a. The transposition to other geometries involving, for example, the creation of vortices in the fluid would be straightforward. The main change would concern the height of the energy barrier[4].

With this simple geometry, the potential to be used in (60) is the sum of a repulsive potential $g\left|u_l(r,z)\right|^2$ and a kinetic energy of rotation around Oz, equal to $l^2\hbar^2/2mr^2$. We now show that, in a given l state, these two contributions can be expressed as a function of two velocities. First, relation (61) yields the rotation velocity v_l associated with state l:

$$v_l = \frac{1}{r}\frac{l\hbar}{m} \tag{73}$$

and the rotational energy is simply written as:

$$\langle E_{rot}\rangle_l = \frac{l^2\hbar^2}{2mr^2} = \frac{1}{2}m\left(v_l\right)^2 \tag{74}$$

[4]When several relaxation channels are present, the one associated with the lowest barrier mainly determines the time evolution.

As for the interaction term (term in g on the left-hand side), we can express it in a more convenient way, defining as before the numerical density n_0:

$$n_0 = |u_l(r,z)|^2 \tag{75}$$

and using the definition (39) for the sound velocity c. It can then be written in a form similar to (74):

$$gn_0 = mc^2 \tag{76}$$

The two velocities v_l and c allow an easy comparison of the respective importance of the kinetic and potential energies in a state l.

We now compare the contributions of these two terms either for states with a given l, or for a superposition of states (63). To clarify the discussion and be able to draw a figure, we will use a continuous variable defined as the average $\langle J_z \rangle$ of the component along Oz of the angular momentum:

$$\langle J_z \rangle = l\hbar \ |c_l(t)|^2 + l'\hbar \ |c_{l'}(t)|^2 \tag{77}$$

This expression varies continuously between $l\hbar$ and $l'\hbar$ when the relative weights of $|c_l(t)|^2$ and $|c_{l'}(t)|^2$ are changed while imposing relation (64); the continuous variable:

$$x = \langle J_z \rangle / \hbar \tag{78}$$

allows making interpolations between the discrete integer values of l.

Using the normalization relation (64) of the wave function (63), we can express x as a function of $|c_l(t)|^2$:

$$x = (l - l')\,|c_l(t)|^2 + l' \tag{79}$$

The variable x characterizes the modulus of each of the two components of the variational function (63). A second variable is needed to define the relative phase between these two components, which comes into play for example in (66). Instead of studying the time evolution of the fluid state vector inside this variational family, we shall simply give a qualitative argument, for several reasons. First of all, it is not easy to characterize precisely the coupling between the fluid and the environment by a Hamiltonian that can change the fluid rotational angular momentum (for example, the wall's irregularities may transfer energy and angular momentum from the fluid to the container). Furthermore, as the time-dependent Gross-Pitaevskii equation is nonlinear, its precise solutions are generally found numerically. This is why we shall only qualitatively discuss the effects of the potential barrier found in §3-b. The higher this barrier, the more difficult it is for x to go from l to l'. Let us evaluate the variation of the average energy as a function of x.

For integer values of x, relation (74) shows that the average rotational kinetic energy varies as the square of x; in between, its value can be found by interpolation as in (77). As for the potential energy, we saw that a continuous variation of $c_l(t)$ and $c_{l'}(t)$ necessarily involves a coherent superposition, which has an energy cost and increases the repulsive potential interaction. In particular, this interaction energy is multiplied by the factor 3/2 when the moduli of $c_l(t)$ and $c_{l'}(t)$ are equal (i.e. when x is an integer plus 1/2). As a result, to the quadratic variation of the rotational kinetic energy, we must

add an oscillating variation of the potential energy, minimum for all the integer values of x, and maximum half-way between. The oscillation amplitude is given by:

$$\frac{gn_0}{2} = \frac{1}{2}mc^2 \tag{80}$$

Figure 3 shows three plots of the variation of the system energy as a function of the average value $\langle J_z \rangle$. The lowest one, shown as a dotted line, corresponds to a superposition of the state l with the state $l' = l-1$, for a very small value of the coupling constant g (weak interactions, gas almost ideal). In this case and according to (39), the sound velocity is also very small and we are in the case $c \ll v_l$. Comparing (74) and (80) then shows that the potential energy contribution is negligible compared to the variation of the rotational kinetic energy between the two states. As a result, the modulation on this dotted line is barely perceptible, and this curve presents a single minimum at $x = 0$: whatever the initial rotational state, no potential barrier prevents the fluid rotational velocity from returning to zero (for example under the effect of the interactions with the irregularities of the walls containing the fluid).

The other two curves in Figure 3 correspond to a much larger value of g, hence, according to (39), to a much higher value of c. There are now several values of l for which v_l is small compared to c. The dashed line corresponds, as for the previous curve, to a superposition of the two states $l = 1$ and $l' = l-1$; the solid line (for the same value of g) to a superposition of $l = 3$ and $l' = 0$, corresponding to the case where the system goes directly from the state $l = 3$ to the rotational ground state in the torus, with $l' = 0$. It is obviously this last curve that presents the lowest energy barrier starting from $l = 3$ (shown with a circle in the figure). This is normal since this is the curve that involves the largest variation in the kinetic energy, in a sense opposite to that of the potential energy variation. It is thus the direct transition from $l = 3$ to $l' = 0$ that will determine the possibility for the system to relax towards a state of slower rotation. Let us again use (74) and (80) to compare the kinetic energy variation and the height of the repulsive potential barrier. All the states l, with velocities v_l much larger than c, have a kinetic energy much bigger than the maximum value of the potential energy: no energy barrier can be formed. On the other hand, all the states l with velocities v_l much smaller than c cannot lower their rotational state without going over a potential barrier.

In between these two extreme cases, there exists (for a given g) a "critical" value l_c corresponding to the onset of the barrier. It is associated with a "critical velocity" $v_c = l_c\hbar/mr$, of the order of the sound velocity c, fixing the maximum value of v_l for which this potential barrier exists. If the fluid rotational velocity in the torus is greater than v_c, the liquid can slow down its rotation without going over an energy barrier, and dissipation occurs as in an ordinary viscous liquid – the fluid is said to be "normal". If, however, the fluid velocity is less than the critical velocity, the physical system must necessarily go over a potential barrier (or more) to continuously tend towards $l = 0$. As this barrier results from the repulsion between all the particles and their neighbors, it has a macroscopic value. In principle, any barrier can be overcome, be it by thermal excitation, or by the quantum tunnel effect. However the time needed for this passage may take a gigantic value. First of all, it is extremely unlikely for a thermal fluctuation to reach a macroscopic energy value. As for the tunnel effect, its transition probability decreases exponentially with the barrier height and becomes extremely low for a macroscopic object. Consequently, the relaxation times of the fluid velocity may

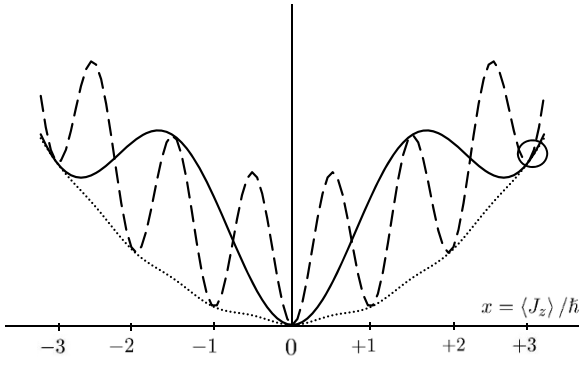


Figure 3: Plots of the energy of a rotating repulsive boson system, in a coherent superposition of the state l and the state l', as a function of its average angular momentum $\langle J_z \rangle$, expressed in units of $\hbar$. The lower dotted curve corresponds to the case where $l' = l - 1$ and the interaction constant g is small (almost ideal gas). The potential energy is then negligible and the total energy presents a single minimum in $\langle J_z \rangle = 0$. Consequently, whatever the initial rotational state of the fluid, it will relax to a motionless state $l = 0$ without having to go over any energy barrier, and its rotational kinetic energy will dissipate: it behaves as a normal fluid. The other two curves correspond to a much larger value of g – therefore, according to (39) to a much higher value of c. The dashed curve still corresponds to a superposition of the rotational states l and $l' = l - 1$, and the solid line to the direct superposition of the state $l = 3$ (shown with a circle in the figure) and the ground state $l' = 0$. The solid line curve presents the smallest barrier, hence determining the metastability of the current.
The higher the coupling constant g, the more l states presenting a minimum in the potential energy appear. They correspond to flow velocities in the torus that are smaller than the critical velocity. To go from the rotational state $l = 1$ to the motionless state $l = 0$, the system must go over a macroscopic energy barrier, which only occurs with a probability so small it can be considered equal to zero. The rotational current is therefore permanent, lasting for years, and the system is said to be superfluid. On the other hand, the states with higher values of l, for which the curve presents no minima, correspond to a normal fluid, whose rotation can slow down because of the viscosity (dissipation of the kinetic energy into heat).

become extraordinarily large, and, on the human scale, the rotation can be considered to last indefinitely. This phenomenon is called "superfluidity".

3-d. Generalization; topological aspects

Our argument remained qualitative for several reasons. To begin with, we showed the existence for the fluid of a critical velocity v_c, of the order of c, without giving its precise value. It would require a more detailed study of the potential curves such as the ones plotted in Figure 3, to obtain the precise values of the parameters for which the potential barrier appears or disappears. We also limited ourselves to simple geometries that could be described by a single variable φ, not taking into account other possible

deformations of the wave function. Various situations could occur, such as the creation of vortices or more complex processes, which would require a more elaborate mathematical treatment. In other words, we would have to take into account the existence of other relaxation channels for the moving fluid to come to rest, and look for the one leading to the lowest potential barrier, thereby determining the lifetime of the superfluid current.

There is, however, a more general way to address the problem, which shows that our basic conclusions are not limited to the particular case we have studied. It is based on the topological aspects of the wave function phase. When this phase varies by $2l\pi$ as we go around the torus, it expresses a topological property characterized by the winding number l, which is an integer and cannot vary continuously. This is why, as long as the phase is well defined everywhere – i.e. as long as the wave function does not go to zero – we cannot go continuously from l to $l \pm 1$. We already saw this in the particular example of the wave function (63): when the modulus of $c_l\,(t)$ varies in time from 1 to 0, while the modulus of $c_{l'}\,(t)$ does the opposite, we necessarily went through a situation where the wave function went to zero through interference, in a plane corresponding to a certain value of φ; but the phase of the wave function is undetermined in this plane, and as we cross it, the phase undergoes a discontinuous jump. Now the canceling of the wave function of a great number of condensed bosons means the density must also be zero at that point, hence larger in other points of space. This spatial density variation introduces an energy increase, due to the finite compressibility of the fluid (as we saw in § 3-b, the energy increase in the high density regions is larger than the energy decrease in low density regions). This means there is an energy barrier opposing the change in the number of turns l of the phase. The height of this barrier must now be compared with the kinetic energy variation. As seen above, there is a drastic change in the flow regime, depending on whether the fluid velocity is smaller or larger than a certain critical velocity v_c. In the first case, superfluidity allows a current to flow without dissipation, lasting practically indefinitely. In the second, no energy consideration opposes dissipation, and the rotation slows down progressively, as in an ordinary liquid.

The essential idea to remember is that superfluidity comes from the repulsive interactions, and for two reasons. First of all, they explain the presence of the energy barrier, responsible for the metastability. The second reason, even more essential, is that the repulsion between bosons constantly tends to put all the fluid particles in the same quantum state - see § 4-c of Complement C_{XV}; thanks to this property, we were able to characterize the intermediate rotational states by a very simple wave function (63). This implies that the quantum fluid can only occupy a very limited number of states, compared to a situation where the particles would be distinguishable. Consequently, it has a hard time dissipating its kinetic energy into heat, as a classical fluid would do, and it therefore maintains its rotation over such long times that a slowing down is practically impossible to observe.

Complement E_{XV}

Fermion system, Hartree-Fock approximation

1 Foundation of the method 1678
1-a Trial family and Hamiltonian 1678
1-b Energy average value 1680
1-c Optimization of the variational wave function 1682
1-d Equivalent formulation for the average energy stationarity . . 1684
1-e Variational energy 1685
1-f Hartree-Fock equations 1686
2 Generalization: operator method 1688
2-a Average energy 1689
2-b Optimization of the one-particle density operator 1692
2-c Mean field operator 1693
2-d Hartree-Fock equations for electrons 1695
2-e Discussion 1698

Introduction

Computing the energy levels of a system of N electrons, interacting with each other through the Coulomb force, and placed in an external potential $V_1(\mathbf{r})$ is a very important problem in physics and chemistry. It is encountered in the determination of the energy levels of atoms (in which case the external potential for the electrons[1] is the Coulomb potential created by the nucleus $-Zq^2/4\pi\varepsilon_0 r$), or of molecules as well, or of electrons in a solid (submitted to a periodic potential), or in an aggregate or a nanocristal, etc. It is a problem where two ingredients simultaneously play an essential role: the fermionic character of the electrons, which forbids them to occupy the same individual state, and the effects of their mutual interactions. Ignoring the Coulomb repulsion between electrons would make the calculation fairly simple, and similar to that of § 1 in Complement C_{XIV}, concerning free fermions in a box; the free plane wave individual states would have to be replaced by the energy eigenstates of a single particle placed in the potential $V_1(\mathbf{r})$. This would lead to a 3-dimensional Schrödinger equation, which can be solved with very good precision, although not necessarily analytically.

However, be it in atoms or in solids, the repulsion between electrons plays an essential role. Neglecting it would lead us to conclude, for example, that, as Z increases, the size of atoms decreases due to the attractive effect of the nucleus, whereas the opposite occurs[2]! For N interacting particles, even without taking the spin into account, an exact

[1] We assume the nucleus mass to be infinitely larger than the electron mass. The electronic system can then be studied assuming the nucleus fixed and placed at the origin.

[2] The Pauli exclusion principle is not sufficient to explain why an atom's size increases with its atomic number Z. One can evaluate the approximate size of a hypothetical atom with non-interacting electrons (we consider the atom's size to be given by the size of the outermost occupied orbit). The Bohr radius

computation would require solving a Schrödinger equation in a $3N$-dimensional space; this is clearly impossible when N becomes large, even with the most powerful computer. Hence, approximation methods are needed, and the most common one is the Hartree-Fock method, which reduces the problem to solving a series of 3-dimensional equations. It will be explained in this complement for fermionic particles.

The Hartree-Fock method is based on the variational approximation (Complement E_{XI}), where we choose a trial family of state vectors, and look for the one that minimizes the average energy. The chosen family is the set of all possible Fock states describing the system of N fermions. We will introduce and compute the "self-consistent" mean field in which each electron moves; this mean field takes into account the repulsion due to the other electrons, hence justifying the central field method discussed in Complement A_{XIV}. This method applies not only to the atom's ground state but also to all its stationary states. It can also be generalized to many other systems such as molecules, for example, or to the study of the ground level and excited states of nuclei, which are protons and neutrons in bound systems.

This complement presents the Hartree-Fock method in two steps, starting in § 1 with a simple approach in terms of wave functions, which is then generalized in § 2 by using Dirac notation and projector operators. The reader may choose to go through both steps or go directly to the second. In § 1, we deal with spinless particles, which allows discussing the basic physical ideas and introducing the mean field concept keeping the formalism simple. A more general point of view is exposed in § 2, to clarify a number of points and to introduce the concept of a one-particle (with or without spin) effective Hartree-Fock Hamiltonian. This Hamiltonian reduces the interactions with all the other particles to a mean field operator. More details on the Hartree-Fock methods, and in particular their relations with the Wick theorem, can be found in Chapters 7 and 8 of reference [5].

1. Foundation of the method

Let us first expose the foundation of the Hartree-Fock method in a simple case where the particles have no spin (or are all in the same individual spin state) so that no spin quantum number is needed to define their individual states, specified by their wave functions. We introduce the notation and define the trial family of the N-particle state vectors.

1-a. Trial family and Hamiltonian

We choose as the trial family for the state of the N-fermion system all the states that can be written as:

$$\widetilde{|\Psi\rangle} = a^\dagger_{\theta_1} a^\dagger_{\theta_2} ... \; a^\dagger_{\theta_N} |0\rangle \tag{1}$$

where $a^\dagger_{\theta_1}$, $a^\dagger_{\theta_2}$,..., $a^\dagger_{\theta_N}$ are the creation operators associated with a set of normalized individual states $|\theta_1\rangle$, $|\theta_2\rangle$, ...$|\theta_N\rangle$, all orthogonal to each other (and hence distinct). The state $\widetilde{|\Psi\rangle}$ is therefore normalized to 1. This set of individual states is, at the moment, arbitrary; it will be determined by the following variational calculation.

a_0 varies as $1/Z$, whereas the highest value of the principal quantum number n of the occupied states varies approximately as $Z^{1/3}$. The size $n^2 a_0$ we are looking for varies approximately as $Z^{-1/3}$.

For spinless particles, the corresponding wave function $\widetilde{\Psi}(\mathbf{r}_1, \mathbf{r}_2, ..\mathbf{r}_N)$ can be written in the form of a Slater determinant (Chapter XIV, § C-3-c-β):

$$\widetilde{\Psi}(\mathbf{r}_1, \mathbf{r}_2, ..\mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \theta_1(\mathbf{r}_1) & \theta_2(\mathbf{r}_1) & .. & \theta_N(\mathbf{r}_1) \\ \theta_1(\mathbf{r}_2) & \theta_2(\mathbf{r}_2) & .. & \theta_N(\mathbf{r}_2) \\ .. & .. & .. & .. \\ \theta_1(\mathbf{r}_N) & \theta_2(\mathbf{r}_N) & .. & \theta_N(\mathbf{r}_N) \end{vmatrix} \tag{2}$$

The system Hamiltonian is the sum of the kinetic energy, the one-body potential energy and the interaction energy:

$$\widehat{H} = \widehat{H}_0 + \widehat{V}_{\text{ext}} + \widehat{W}_{\text{int}} \tag{3}$$

The first term, $\widehat{H}_0$, is the operator associated with the fermion kinetic energy, sum of the individual kinetic energies:

$$\widehat{H}_0 = \sum_{q=1}^{N} \frac{(\mathbf{P}_q)^2}{2m} \tag{4}$$

where m is the particle mass and $\mathbf{P}_q$, the momentum operator of particle q. The second term, $\widehat{V}_{\text{ext}}$, is the operator associated with their energy in an applied external potential $V_{1.}$:

$$\widehat{V}_{\text{ext}} = \sum_{q=1}^{N} V_1(\mathbf{R}_q) \tag{5}$$

where $\mathbf{R}_q$ is the position operator of particle q. For electrons with charge q_e placed in the attractive Coulomb potential of a nucleus of charge $-Zq_e$ positioned at the origin (Z is the nucleus atomic number), this potential is attractive and equal to:

$$V_1(\mathbf{r}) = -\frac{Zq_e^2}{4\pi\varepsilon_0} \frac{1}{|\mathbf{r}|} \tag{6}$$

where ε_0 is the vacuum permittivity. Finally, the term $\widehat{W}_{\text{int}}$ corresponds to their mutual interaction energy:

$$\widehat{W}_{\text{int}} = \frac{1}{2} \sum_{q \neq q'} W_2(\mathbf{R}_q, \mathbf{R}_{q'}) \tag{7}$$

For electrons, the function W_2 is given by the Coulomb repulsive interaction:

$$W_2(\mathbf{r}, \mathbf{r}') = \frac{q_e^2}{4\pi\varepsilon_0} \frac{1}{|\mathbf{r} - \mathbf{r}'|} \tag{8}$$

The expressions given above are just examples; as mentioned earlier, the Hartree-Fock method is not limited to the computation of the electronic energy levels in an atom.

1-b. Energy average value

Since state (1) is normalized, the average energy in this state is given by:

$$\widetilde{E} = \widetilde{\langle\Psi|}\widehat{H}\widetilde{|\Psi\rangle} \tag{9}$$

Let us evaluate successively the contributions of the three terms of (3), to obtain an expression which we will eventually vary.

α. Kinetic energy

Let us introduce a complete orthonormal basis $\{|\theta_s\rangle\}$ of the one-particle state space by adding to the set of states $|\theta_i\rangle$ ($i = 1, 2, ..., N$) other orthonormal states; the subscript s now ranges from 1 to D, dimension of this space (D may be infinite). We can then expand $\widehat{H}_0$ as in relation (B-12) of Chapter XV:

$$\widehat{H}_0 = \sum_{r,s} \langle\theta_r| \frac{\mathbf{P}^2}{2m} |\theta_s\rangle \; a^\dagger_{\theta_r} a_{\theta_s} \tag{10}$$

where the two summations over r and s range from 1 to D. The average value in $\widetilde{|\Psi\rangle}$ of the kinetic energy can then be written:

$$\left\langle \widehat{H}_0 \right\rangle = \sum_{r,s} \langle\theta_r| \frac{\mathbf{P}^2}{2m} |\theta_s\rangle \;\; \langle 0| \, a_{\theta_N} ... \; a_{\theta_2} a_{\theta_1} \left(a^\dagger_{\theta_r} a_{\theta_s} \right) a^\dagger_{\theta_1} a^\dagger_{\theta_2} ... \; a^\dagger_{\theta_N} |0\rangle \tag{11}$$

which contains the scalar product of the ket:

$$\left(a_{\theta_s} \right) \; a^\dagger_{\theta_1} a^\dagger_{\theta_2} ... \; a^\dagger_{\theta_N} |0\rangle = \left(a_{\theta_s} \right) \;\; |\theta_1, \theta_2, \; ... \; , \theta_N\rangle \tag{12}$$

by the bra:

$$\langle 0| \, a_{\theta_N} ... \; a_{\theta_2} a_{\theta_1} \;\; \left(a^\dagger_{\theta_r} \right) = \langle \theta_1, \theta_2, \; ... \; , \theta_N| \;\; \left(a^\dagger_{\theta_r} \right) \tag{13}$$

Note however that in the ket, the action of the annihilation operator a_{θ_s} yields zero unless it acts on a ket where the individual state is already occupied; consequently, the result will be different from zero only if the state $|\theta_s\rangle$ is included in the list of the N states $|\theta_1\rangle$, $|\theta_2\rangle$, ...$|\theta_N\rangle$. Taking the Hermitian conjugate of (13), we see that the same must be true for the state $|\theta_r\rangle$, which must be included in the same list. Furthermore, if $r \neq s$ the resulting kets have different occupation numbers, and are thus orthogonal. The scalar product will therefore only differ from zero if $r = s$, in which case it is simply equal to 1. This can be shown by moving to the front the state $|\theta_r\rangle$ both in the bra and in the ket; this will require two transpositions with two sign changes which cancel out, or none if the state $|\theta_r\rangle$ was already in the front. Once the operators have acted, the bra and the ket correspond to exactly the same occupied states and their scalar product is 1. We finally get:

$$\left\langle \widehat{H}_0 \right\rangle = \sum_{i=1}^{N} \langle\theta_i| \frac{\mathbf{P}^2}{2m} |\theta_i\rangle \tag{14}$$

Consequently, the average value of the kinetic energy is simply the sum of the average kinetic energy in each of the occupied states $|\theta_i\rangle$.

For spinless particles, the kinetic energy operator is actually a differential operator $-\hbar^2\Delta/2m$ acting on the individual wave functions. We therefore get:

$$\left\langle \widehat{H}_0 \right\rangle = -\frac{\hbar^2}{2m} \int \mathrm{d}^3 r \ \sum_{i=1}^{N} \theta_i^*(\mathbf{r}) \ \Delta\theta_i(\mathbf{r}) \tag{15}$$

β. Potential energy

As the potential energy $\widehat{V}_1$ is also a one-particle operator, its average value can be computed in a similar way. We obtain:

$$\left\langle \widehat{V}_{\mathrm{ext}} \right\rangle = \sum_{i=1}^{N} \langle\theta_i| \, V_1(\mathbf{R}) \, |\theta_i\rangle \tag{16}$$

that is, for spinless particles:

$$\left\langle \widehat{V}_{\mathrm{ext}} \right\rangle = \int \mathrm{d}^3 r \ V_1(\mathbf{r}) \sum_{i=1}^{N} \ |\theta_i(\mathbf{r})|^2 \tag{17}$$

As before, the result is simply the sum of the average values associated with the individual occupied states.

γ. Interaction energy

The average value of the interaction energy $\widehat{W}_2$ in the state $\widetilde{|\Psi\rangle}$ has already been computed in § C-5 of Chapter XV. We just have to replace, in the relations (C-28) as well as (C-32) to (C-34) of that chapter, the n_i by 1 for all the occupied states $|\theta_i\rangle$, by zero for the others, and to rename the wave functions $u_i(\mathbf{r})$ as $\theta_i(\mathbf{r})$. We then get:

$$\begin{aligned} \left\langle \widehat{W}_{\mathrm{int}} \right\rangle = \widetilde{\langle\Psi|} \widehat{W}_{\mathrm{int}} \widetilde{|\Psi\rangle} = \frac{1}{2} \ \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \ W_2(\mathbf{r}, \mathbf{r}') \\ \times \sum_{i,j=1}^{N} \left[|\theta_i(\mathbf{r})|^2 \, |\theta_j(\mathbf{r}')|^2 - \theta_i^*(\mathbf{r}) \theta_j^*(\mathbf{r}') \theta_i(\mathbf{r}') \theta_j(\mathbf{r}) \right] \end{aligned} \tag{18}$$

We have left out the condition $i \neq j$, no longer useful since the $i = j$ terms are zero. The second line of this equation contains the sum of the direct and the exchange terms.

The result can be written in a more concise way by introducing the projector P_N over the subspace spanned by the N kets $|\theta_i\rangle$:

$$P_N = \sum_{i=1}^{N} |\theta_i\rangle \, \langle\theta_i| \tag{19}$$

Its matrix elements are:

$$\langle\mathbf{r}| \, P_N \, |\mathbf{r}'\rangle = \sum_{i=1}^{N} \theta_i(\mathbf{r}) \theta_i^*(\mathbf{r}') \tag{20}$$

This leads to:

$$\left\langle \widehat{W}_{\text{int}} \right\rangle = \frac{1}{2} \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \; W_2(\mathbf{r},\mathbf{r}') \left[\langle \mathbf{r}| \, P_N \, |\mathbf{r}\rangle \, \langle \mathbf{r}'| \, P_N \, |\mathbf{r}'\rangle - \langle \mathbf{r}| \, P_N \, |\mathbf{r}'\rangle \, \langle \mathbf{r}'| \, P_N \, |\mathbf{r}\rangle \right] \tag{21}$$

Comment:

The matrix elements of P_N are actually equal to the spatial non-diagonal correlation function $G_1(\mathbf{r},\mathbf{r}')$, which will be defined in Chapter XVI (§ B-3-a). This correlation function can be expressed as the average value of the product of field operators $\Psi(\mathbf{r})$:

$$G_1(\mathbf{r},\mathbf{r}') = \left\langle \Psi^\dagger(\mathbf{r}) \Psi(\mathbf{r}') \right\rangle \tag{22}$$

For a system of N fermions in the states $|\theta_1\rangle$, $|\theta_2\rangle$, .., $|\theta_N\rangle$, we can write:

$$\begin{array}{l} G_1(\mathbf{r},\mathbf{r}') = \langle \theta_1, \theta_2, .., \; \theta_N| \, \Psi^\dagger(\mathbf{r}) \Psi(\mathbf{r}') \, |\theta_1, \theta_2, .., \; \theta_N\rangle \\ = \sum_{i,j} \langle \theta_1, \theta_2, .., \; \theta_N| \left[\theta_i^*(\mathbf{r}) a_i^\dagger \right] \left[\theta_j(\mathbf{r}') a_j \right] |\theta_1, \theta_2, .., \; \theta_N\rangle = \sum_{i=1}^N \theta_i^*(\mathbf{r}) \theta_i(\mathbf{r}') \end{array} \tag{23}$$

Inserting this relation in (18) we get:

$$\left\langle \widehat{W}_{\text{int}} \right\rangle = \frac{1}{2} \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \; W_2(\mathbf{r},\mathbf{r}') \left[G_1(\mathbf{r},\mathbf{r}) G_1(\mathbf{r}',\mathbf{r}') - G_1(\mathbf{r},\mathbf{r}') G_1(\mathbf{r}',\mathbf{r}) \right] \tag{24}$$

Comparison with relation (C-28) of Chapter XV, which gives the same average value, shows that the right-hand side bracket contains the two-particle correlation function $G_2(\mathbf{r},\mathbf{r}')$. For a Fock state, this function can therefore be simply expressed as two products of one-particle correlation functions at two points:

$$G_2(\mathbf{r},\mathbf{r}') = G_1(\mathbf{r},\mathbf{r}) G_1(\mathbf{r}',\mathbf{r}') - G_1(\mathbf{r},\mathbf{r}') G_1(\mathbf{r}',\mathbf{r}) \tag{25}$$

1-c. Optimization of the variational wave function

We now vary $\widetilde{|\Psi\rangle}$ to determine the conditions leading to a stationary value of the total energy $\widetilde{E}$:

$$\widetilde{E} = \left\langle \widehat{H}_0 \right\rangle + \left\langle \widehat{V}_{\text{ext}} \right\rangle + \left\langle \widehat{W}_{\text{int}} \right\rangle \tag{26}$$

where the three terms in this summation are given by (15), (16) and (18). Let us vary one of the kets $|\theta_k\rangle$, k being arbitrarily chosen between 1 and N:

$$|\theta_k\rangle \Rightarrow |\theta_k\rangle + |\delta\theta_k\rangle \tag{27}$$

or, in terms of an individual wave function:

$$\theta_k(\mathbf{r}) \Rightarrow \theta_k(\mathbf{r}) + \delta\theta_k(\mathbf{r}) \tag{28}$$

This will yield the following variations:

$$\delta \left\langle \widehat{H}_0 \right\rangle = \frac{-\hbar^2}{2m} \int \mathrm{d}^3 r \; \left[\delta\theta_k^*(\mathbf{r}) \; \Delta\theta_k(\mathbf{r}) + \theta_k^*(\mathbf{r}) \; \Delta\delta\theta_k(\mathbf{r}) \right] \tag{29}$$

and:

$$\delta \left\langle \widehat{V}_{\text{ext}} \right\rangle = \int \mathrm{d}^3 r \; V_1(\mathbf{r}) \left[\delta\theta_k^*(\mathbf{r}) \; \theta_k(\mathbf{r}) + \theta_k^*(\mathbf{r}) \; \delta\theta_k(\mathbf{r}) \right] \tag{30}$$

As for the variation of $\left\langle \widehat{W}_{\text{int}} \right\rangle$, we must take from (18) two contributions: the first one from the terms $i = k$, and the other from the terms $j = k$. These contributions are actually equal as they only differ by the choice of a dummy subscript. The factor 1/2 disappears and we get:

$$\begin{aligned} \delta \left\langle \widehat{W}_{\text{int}} \right\rangle = \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \; W_2(\mathbf{r}, \mathbf{r}') \sum_{j=1}^{N} \Big[\Big[\delta\theta_k^*(\mathbf{r}) \theta_k(\mathbf{r}) + \theta_k^*(\mathbf{r}) \delta\theta_k(\mathbf{r}) \Big] \left| \theta_j(\mathbf{r}') \right|^2 \\ - \delta\theta_k^*(\mathbf{r}) \theta_j^*(\mathbf{r}') \theta_k(\mathbf{r}') \theta_j(\mathbf{r}) - \theta_k^*(\mathbf{r}) \theta_j^*(\mathbf{r}') \delta\theta_k(\mathbf{r}') \theta_j(\mathbf{r}) \Big] \end{aligned} \tag{31}$$

The variation of $\widetilde{E}$ is simply the sum of (29), (30) and (31).

We now consider variations $\delta\theta_k$, which can be written as:

$$\delta\theta_k(\mathbf{r}) = \delta\varepsilon \; e^{i\chi} \theta_l(\mathbf{r}) \qquad \text{with } l > N \tag{32}$$

(where $\delta\varepsilon$ is a first order infinitely small parameter). These variations are proportional to the wave function of one of the non-occupied states, which was added to the occupied states to form a complete orthonormal basis; the phase χ is an arbitrary parameter. Such a variation does not change, to first order, either the norm of $|\theta_k\rangle$, or its scalar product with all the occupied states $l \leq N$; it therefore leaves unchanged our assumption that the occupied states basis is orthonormal. The first order variation of the energy $\delta\widetilde{E}$ is obtained by inserting $\delta\theta_k$ and its complex conjugate $\delta\theta_k^*$ into (29), (30) and (31); we then get terms in $e^{i\chi}$ in the first case, and terms in $e^{-i\chi}$ in the second. For $\widetilde{E}$ to be stationary, its variation must be zero to first order for any value of χ; now the sum of a term in $e^{i\chi}$ and another in $e^{-i\chi}$ will be zero for any value of χ only if both terms are zero. It follows that we can impose $\delta\widetilde{E}$ to be zero (stationary condition) considering the variations of $\delta\theta_k$ and $\delta\theta_k^*$ to be independent. Keeping only the terms in $\delta\theta_k^*$, we obtain the stationary condition of the variational energy:

$$\begin{aligned} \int \mathrm{d}^3 r \; \delta\theta_k^*(\mathbf{r}) \; \Bigg\{ & \frac{-\hbar^2}{2m} \Delta\theta_k(\mathbf{r}) + V_1(\mathbf{r}) \theta_k(\mathbf{r}) + \\ & + \int \mathrm{d}^3 r' \; W_2(\mathbf{r}, \mathbf{r}') \sum_{j=1}^{N} \Big[\theta_k(\mathbf{r}) \left| \theta_j(\mathbf{r}') \right|^2 - \theta_j^*(\mathbf{r}') \theta_k(\mathbf{r}') \theta_j(\mathbf{r}) \Big] \Bigg\} = 0 \end{aligned} \tag{33}$$

or, taking (20) into account:

$$\begin{aligned} \int \mathrm{d}^3 r \; \delta\theta_k^*(\mathbf{r}) \; \Bigg\{ & \frac{-\hbar^2}{2m} \Delta\theta_k(\mathbf{r}) + V_1(\mathbf{r}) \theta_k(\mathbf{r}) + \\ & + \int \mathrm{d}^3 r' \; W_2(\mathbf{r}, \mathbf{r}') \left[\langle \mathbf{r}' | \, P_N \, | \mathbf{r}' \rangle \, \theta_k(\mathbf{r}) - \langle \mathbf{r} | \, P_N \, | \mathbf{r}' \rangle \, \theta_k(\mathbf{r}') \right] \Bigg\} = 0 \end{aligned} \tag{34}$$

This relation can also be written as::

$$\int \mathrm{d}^3 r \; \delta\theta_k^*(\mathbf{r}) \; \mathcal{D} \left[\theta_k(\mathbf{r}) \right] = 0 \tag{35}$$

where the integro-differential operator $\mathcal{D}$ is defined by its action on an arbitrary function $\theta(\mathbf{r})$:

$$\mathcal{D}\left[\theta(\mathbf{r})\right] = \left\{ \frac{-\hbar^2}{2m}\Delta + V_1(\mathbf{r}) + \int \mathrm{d}^3 r' \; W_2(\mathbf{r},\mathbf{r}') \left\langle \mathbf{r}' \right| P_N \left| \mathbf{r}' \right\rangle \right\} \theta(\mathbf{r}) - \int \mathrm{d}^3 r' \; W_2(\mathbf{r},\mathbf{r}') \; \left\langle \mathbf{r} \right| P_N \left| \mathbf{r}' \right\rangle \theta(\mathbf{r}') \tag{36}$$

This operator depends on the diagonal $\left\langle \mathbf{r}' \right| P_N \left| \mathbf{r}' \right\rangle$ and non-diagonal $\left\langle \mathbf{r} \right| P_N \left| \mathbf{r}' \right\rangle$ spatial correlation functions associated with the set of states occupied by the N fermions.

Relation (35) thus shows that the action of the differential operator $\mathcal{D}$ on the function $\theta_k(\mathbf{r})$ yields a function orthogonal to all the functions $\theta_l(\mathbf{r})$ for $l > N$. This means that the function $\mathcal{D}\left[\theta_k(\mathbf{r})\right]$ only has components on the wave functions of the occupied states: it is a linear combination of these functions. Consequently, for the energy $\widetilde{E}$ to be stationary there is a simple condition: the invariance under the action of the integro-differential operator $\mathcal{D}$ of the N-dimensional vector space $\mathscr{F}_N$, spanned by all the linear combinations of the functions $\theta_i(\mathbf{r})$ with $i = 1, 2, ..N$.

Comment:

One could wonder why we limited ourselves to the variations $\delta\theta_k$ written in (32), proportional to non-occupied individual states. The reason will become clearer in § 2, where we use a more general method that shows directly which variations of each individual states are really useful to consider (see in particular the discussion at the end of § 2-a). For now, it can be noted that choosing a variation $\delta\theta_k$ proportional to the same wave function $\theta_k(\mathbf{r})$ would simply change its norm or phase, and therefore have no impact on the associated quantum state (in addition, a change of norm would not be compatible with our hypotheses, as in the computation of the average values we always assumed the individual states to remain normalized). If the state does not change, the energy $\widetilde{E}$ must remain constant and writing a stationary condition is pointless. Similarly, to give $\theta_k(\mathbf{r})$ a variation proportional to another occupied wave function $\theta_l(\mathbf{r})$ (where l is included between 1 and N) is just as useless, as we now show. In this operation, the creation operator $a_k^\dagger$ acquires a component on $a_l^\dagger$ (Chapter XV, § A-6), but the state vector expression (1) remains unchanged. The state vector thus acquire a component including the square of a creation operator, which is zero for fermions. Consequently, the stationarity of the energy is automatically ensured in this case.

1-d. Equivalent formulation for the average energy stationarity

Operator $\mathcal{D}$ can be diagonalized in the subspace $\mathscr{F}_N$, as can be shown[3] from its definition (36) – a more direct demonstration will be given in § 2. We call $\varphi_n(\mathbf{r})$ its

[3] As any Hermitian operator can be diagonalized, we simply show that (36) leads to matrix elements obeying the Hermitian conjugation relation. Let us verify that the two integrals $\int d^3r \; \theta_1^*(\mathbf{r}) \; \mathcal{D}\left[\theta_2(\mathbf{r})\right]$ and $\int d^3r \; \theta_2^*(\mathbf{r}) \; \mathcal{D}\left[\theta_1(\mathbf{r})\right]$ are complex conjugates of each other. For the contributions to these matrix elements of the kinetic and potential (in V_1) energy, we simply find the usual relations insuring the corresponding operators are Hermitian. As for the interaction term, the complex conjugation is obvious for the direct term; for the exchange term, a simple inversion of the integral variables d^3r and d^3r', plus the fact that $W_2(\mathbf{r},\mathbf{r}')$ is equal to $W_2(\mathbf{r}',\mathbf{r})$ allows verifying the conjugation.

eigenfunctions. These functions $\varphi_n(\mathbf{r})$ are linear combinations of the $\theta_j(\mathbf{r})$ corresponding to the states appearing in the trial ket (1), and therefore lead to the same N-particle state, because of the antisymmetrization[4]. The basis change from the $\theta_i(\mathbf{r})$ to the $\varphi_n(\mathbf{r})$ has no effect on the projector P_N onto to the subspace $\mathcal{F}_N$, whose matrix elements appearing in (36) can be expressed in a way similar to those in (20):

$$\langle \mathbf{r}| \, P_N \, |\mathbf{r}'\rangle = \sum_{n=1}^{N} \varphi_n(\mathbf{r}) \varphi_n^*(\mathbf{r}') \tag{37}$$

Consequently, the eigenfunctions of the operator $\mathcal{D}$ obey the equations:

$$\boxed{\begin{aligned} &\left[\frac{-\hbar^2}{2m} \Delta + V_1(\mathbf{r}) + \int \mathrm{d}^3 r' \; W_2(\mathbf{r}, \mathbf{r}') \sum_{p=1}^{N} |\varphi_p(\mathbf{r}')|^2 \right] \varphi_n(\mathbf{r}) \\ &\qquad - \left[\int \mathrm{d}^3 r' \; W_2(\mathbf{r}, \mathbf{r}') \sum_{p=1}^{N} \varphi_p^*(\mathbf{r}') \varphi_p(\mathbf{r}) \varphi_n(\mathbf{r}') \right] = \widetilde{e}_n \varphi_n(\mathbf{r}) \end{aligned}} \tag{38}$$

where $\widetilde{e}_n$ are the associated eigenvalues. These relations are called the "Hartree-Fock equations".

For the average total energy associated with a state such as (1) to be stationary, it is therefore necessary for this state to be built from N individual states whose orthogonal wave functions φ_1, φ_2, .. , φ_N are solutions of the Hartree-Fock equations (38) with $n = 1$, 2, .. , N. Conversely, this condition is sufficient since, replacing the $\theta_k(\mathbf{r})$ by solutions $\varphi_n(\mathbf{r})$ of the Hartree-Fock equations in the energy variation (34) yields the result:

$$\widetilde{e}_n \int \mathrm{d}^3 r \; \delta\varphi_k^*(\mathbf{r}) \; \varphi_n(\mathbf{r}) \tag{39}$$

which is zero for all $\delta\varphi_k(\mathbf{r})$ variations, since, according to (32), they must be orthogonal to the N solutions $\varphi_n(\mathbf{r})$. Conditions (38) are thus equivalent to energy stationarity.

1-e. Variational energy

Assume we found a series of solutions for the Hartree-Fock equations, i.e. a set of N eigenfunctions $\varphi_n(\mathbf{r})$ with the associated eigenvalues $\widetilde{e}_i$. We still have to compute the minimal variational energy of the N-particle system. This energy is given by the sum (26) of the three terms of kinetic, potential and interaction energies obtained by replacing in (15), (16) and (18) the $\theta_i(\mathbf{r})$ by the eigenfunctions $\varphi_n(\mathbf{r})$:

$$\widetilde{E}_{HF} = \left\langle \widehat{H}_0 \right\rangle_{HF} + \left\langle \widehat{V}_{\mathrm{ext}} \right\rangle_{HF} + \left\langle \widehat{W}_{\mathrm{int}} \right\rangle_{HF} \tag{40}$$

[4] A determinant value does not change if one adds to one of its column a linear combination of the others. Hence we can add to the first column of the Slater determinant (2) the linear combination of the $\theta_2(\mathbf{r})$, $\theta_3(\mathbf{r})$, ... that makes it proportional to $\varphi_1(\mathbf{r})$. One can then add to the second column the combination that makes it proportional to $\varphi_2(\mathbf{r})$, etc. Step by step, we end up with a new expression for the original wave function $\widetilde{\Psi}(\mathbf{r}_1, \mathbf{r}_2, ...\mathbf{r}_N)$, which now involves the Slater determinant of the $\varphi_i(\mathbf{r})$. It is thus proportional to this determinant. A demonstration of the strict equality (within a phase factor) will be given in § 2.

(the subscripts HF indicate we are dealing with the average energies after the Hartree-Fock optimization, which minimizes the variational energy). Intuitively, one could expect this total energy to be simply the sum of the energies $\widetilde{e}_n$, but, as we are going to show, this is not the case. Multiplying the left-hand side of equation (38) by $\varphi_n^*(\mathbf{r})$ and after integration over d^3r, we get:

$$\widetilde{e}_n = \int \mathrm{d}^3r\ \varphi_n^*(\mathbf{r}) \left\{ \left[\frac{-\hbar^2}{2m}\Delta + V_1(\mathbf{r}) \right] \varphi_n(\mathbf{r}) + \int \mathrm{d}^3r'\ W_2(\mathbf{r},\mathbf{r}') \sum_{p=1}^{N} \left[\left|\varphi_p(\mathbf{r}')\right|^2 \varphi_n(\mathbf{r}) - \varphi_p^*(\mathbf{r}')\varphi_n(\mathbf{r}')\varphi_p(\mathbf{r}) \right] \right\} \tag{41}$$

We then take a summation over the subscript n, and use (15), (16) and (18), the θ being replaced by the φ:

$$\sum_{n=1}^{N} \widetilde{e}_n = \left\langle \widehat{H}_0 \right\rangle_{HF} + \left\langle \widehat{V}_{\mathrm{ext}} \right\rangle_{HF} + 2 \left\langle \widehat{W}_{\mathrm{int}} \right\rangle_{HF} \tag{42}$$

This expression does not yield the stationary value of the total energy, but rather a sum where the particle interaction energy is counted twice. From a physical point of view, it is clear that if each particle energy is computed taking into account its interaction with all the others, and if we then add all these energies, we get an expression that includes twice the interaction energy associated with each pair of particles.

The sum of the $\widetilde{e}_n$ does contain, however, useful information that enables us to avoid computing the interaction energy contribution to the variational energy. Eliminating $\left\langle \widehat{W}_{\mathrm{int}} \right\rangle$ between (40) and (42), we get:

$$\widetilde{E}_{HF} = \frac{1}{2} \left[\sum_{n=1}^{N} \widetilde{e}_n + \left\langle \widehat{H}_0 \right\rangle_{HF} + \left\langle \widehat{V}_{\mathrm{ext}} \right\rangle_{HF} \right] \tag{43}$$

where the interaction energy is no longer present. One can then compute $\left\langle \widehat{H}_0 \right\rangle_{HF}$ and $\left\langle \widehat{V}_{\mathrm{ext}} \right\rangle_{HF}$ using the solutions of the Hartree-Fock equations (38), without worrying about the interaction energy. Using (15) and (17) in this relation, we can write the total energy $\widetilde{E}$ as:

$$\widetilde{E}_{HF} = \frac{1}{2} \left[\sum_{n=1}^{N} \widetilde{e}_n + \sum_{n=1}^{N} \int \mathrm{d}^3r'\ \left|\varphi_n(\mathbf{r}')\right|^2 V_1(\mathbf{r}') - \frac{\hbar^2}{2m} \sum_{n=1}^{N} \int \mathrm{d}^3r' \varphi_n^*(\mathbf{r}')\ \Delta\varphi_n(\mathbf{r}') \right] \tag{44}$$

The total energy is thus half the sum of the $\widetilde{e}_i$, of the average kinetic energy, and finally of the one-body average potential energy.

1-f. Hartree-Fock equations

Equation (38) may be written as:

$$\boxed{\left[-\frac{\hbar^2}{2m}\Delta + V_1(\mathbf{r}) + V^{\mathrm{dir}}(\mathbf{r}) \right] \varphi_n(\mathbf{r}) - \int \mathrm{d}^3r'\ V^{\mathrm{ex}}(\mathbf{r},\mathbf{r}')\varphi_n(\mathbf{r}') = \widetilde{e}_n\ \varphi_n(\mathbf{r})} \tag{45}$$

where the direct $V^{\rm dir}(\mathbf{r})$ and exchange $V^{\rm ex}(\mathbf{r},\mathbf{r}')$ potentials are defined as:

$$\boxed{\begin{aligned} V^{\rm dir}(\mathbf{r}) &= \sum_{p=1}^{N} \int \mathrm{d}^3 r' \ \left|\varphi_p(\mathbf{r}')\right|^2 \ W_2(\mathbf{r},\mathbf{r}') \\ V^{\rm ex}(\mathbf{r},\mathbf{r}') &= \sum_{p=1}^{N} \varphi_p^*(\mathbf{r}')\varphi_p(\mathbf{r}) \ W_2(\mathbf{r},\mathbf{r}') \end{aligned}} \tag{46}$$

Note that the terms $p = n$ coming from the two potentials cancel each other; hence they can be eliminated from the two summations, without changing the final result. The contribution of the direct potential is sometimes called the "Hartree term", and the contribution of the exchange potential, the "Fock term". The first is easy to understand: with the exception of the term $p = n$, it corresponds to the interaction of a particle at point $\mathbf{r}$ with all the others at points $\mathbf{r}'$, averaged for each of them by its density distribution $|\varphi_p(\mathbf{r}')|^2$. As for the exchange potential, and in spite of its name, this term is not, strictly speaking, a potential; it is not diagonal in the position representation, even though it basically comes from a particle interaction which is diagonal in that representation. This peculiar non-diagonal form actually comes from the combination of the fermion antisymmetrization and the variational approximation. This exchange potential is homogeneous to a potential divided by the cube of a length. It is obviously a Hermitian operator as it is derived from a potential $W_2(\mathbf{r},\mathbf{r}')$ which is real and symmetric with respect to $\mathbf{r}$ and $\mathbf{r}'$.

A more intuitive and simplified version of these equations was suggested by Hartree, in which the exchange potentials are ignored in (45). Without the integral term, these equations become very similar to a series of Schrödinger equations for independent particles, each of them moving in the mean potential created by all the others (still with the exception of the term $p = n$ in the summation). Including the Fock term should, however, lead to more precise calculations.

Using for the potentials their expressions (46), the Hartree-Fock equations (45) become a set of N coupled equations. They are nonlinear, since the direct and exchange potentials depend on the functions $\varphi_p(\mathbf{r})$. Even though they look like linear eigenvalue equations with eigenfunctions $\varphi_n(\mathbf{r})$ as solutions, a linear resolution would actually require knowing in advance the solutions, since these functions also appear in the potentials (46). The term "self-consistent" is used to characterize this type of situation and the solutions $\varphi_n(\mathbf{r})$ it leads to.

There are no general analytical methods to solve nonlinear self-consistent equations of this type, even in their simplified Hartree version, and numerical methods using successive approximations are commonly used. We start from a series of plausible functions $\varphi_n^{(0)}(\mathbf{r})$, and compute with (46) the associated potentials. Considering them to be fixed, we obtain linear eigenvalue equations which can be solved quite readily with computers (the single very complicated equation in a $3N$-dimensional space has been replaced by N independent 3-dimensional equations); we have to diagonalize a Hermitian operator to get a new series of orthonormal functions, resulting from the first iteration, and called $\varphi_n^{(1)}(\mathbf{r})$ and $\widetilde{e}_n^{(1)}$. The second iteration starts from these $\varphi_n^{(1)}(\mathbf{r})$, to compute the new potential values, and get new linear differential equations. Solving these equations yields the next order $\varphi_n^{(2)}(\mathbf{r})$, $\widetilde{e}_n^{(2)}$, etc. After a few iterations, one expects the $\varphi_n^{(n)}(\mathbf{r})$ and $\widetilde{e}_n^{(n)}$ to vary only slightly with the iteration order (n), in which case the Hartree-Fock

equations have been solved to a good approximation. Using (44) we can then compute the energy we were looking for. It is also possible that physical arguments can help us choose directly adequate trial functions $\varphi_n(\mathbf{r})$ without any iteration. Inserting them in (44) then directly provides the energy.

Comments:

(i) The solutions of the Hartree-Fock equations may not be unique. Using the iteration process described above, one can easily wind up with different solutions, depending on the initial choice for the $\varphi_n^{(0)}(\mathbf{r})$ functions. This multiplicity of solutions is actually one of the method's advantages, as it can help us find not only the ground level but also the excited levels.

(ii) As we shall see in § 2, taking into account the 1/2 spin of the electrons in an atom does not bring major complications to the Hartree-Fock equations. It is generally assumed that the one-body potential is diagonal in a basis of the two spin states, labeled + and −, and that the interaction potential does not act on the spins. We then simply assemble N_+ equations, for N_+ wave functions $\varphi_n^+(\mathbf{r})$ associated with the spin + particles, with N_- other equations, for N_- wave functions $\varphi_n^-(\mathbf{r})$ associated with spin − particles. These two sets of equations are not independent, since they contain the same direct potential (computed using (46), whose first line includes a summation over p of all the $N = N_+ + N_-$ wave functions). As for the exchange potential, it does not lead to any coupling between the two sets of equations: in the second line of (46), the summation over p only includes particles in the same spin state for the following reason. If the particles have opposite spins, they can be recognized by the direction of their spin (the interaction does not act on the spins), and they no longer behave as indistinguishable particles. The exchange effects only arise for particles having the same spin.

2. Generalization: operator method

We now describe the method in a more general way, using an operator method that leads to more concise expressions, while taking into account explicitly the possible existence of a spin – which plays an essential role in the atomic structure. We will identify more precisely the mathematical object, actually a projector, which we vary to optimize the energy. Physically, this projector is simply the one-particle density operator defined in § B-4 of Chapter XV. This will lead to expressions both more compact and general for the Hartree-Fock equations. They contain a Hartree-Fock operator acting on a single particle, as if it were alone, but which includes a potential operator defined by a partial trace which reflects the interactions with the other particles in the mean field approximation. Thanks to this operator we can get an approximate value of the entire system energy, computing only individual energies; these energies are obtained with calculations similar to the one used for a single particle placed in a mean field. With this approach, we have a better understanding of the way the mean field approximately represents the interaction with all the other particles; this approach can also suggest ways to make the approximations more precise.

We assume as before that the N-particle variational ket $\widetilde{|\Psi\rangle}$ is written as:

$$\widetilde{|\Psi\rangle} = a_{\theta_1}^\dagger a_{\theta_2}^\dagger \ldots a_{\theta_N}^\dagger |0\rangle \tag{47}$$

This ket is derived from N individual orthonormal kets $|\theta_k\rangle$, but these kets can now describe particles having an arbitrary spin. Consider the orthonormal basis $\{|\theta_k\rangle\}$ of the one-particle state space, in which the set of $|\theta_i\rangle$ ($i = 1, 2, ...N$) was completed by other orthonormal states. The projector P_N onto the subspace $\mathcal{F}_N$ is the sum of the projections onto the first N kets $|\theta_i\rangle$:

$$P_N = \sum_{i=1}^{N} |\theta_i\rangle \langle\theta_i| \tag{48}$$

This is simply the one-particle density operator defined in § B-4 of Chapter XV (normalized by a trace equal to the particle number N and not to one), as we now show. Relation (B-24) of that chapter can be written in the $|\theta_k\rangle$ basis:

$$\langle\theta_l| \widehat{\rho}_1 |\theta_k\rangle = \left\langle a_k^\dagger a_l \right\rangle \tag{49}$$

where the average value $\left\langle a_k^\dagger a_l \right\rangle$ is taken in the quantum state (47). In this kind of Fock state, the average value is different from zero only when the creation operator reconstructs the population destroyed by the annihilation operator, hence if $k = l$, in which case it is equal to the population n_k of the individual states $|\theta_k\rangle$. In the variational ket (47), all the populations are zero except for the first N states $|\theta_i\rangle$ ($i = 1, 2, ...N$), where they are equal to one. Consequently, the one-particle density operator is represented by a matrix, diagonal in the basis $|\theta_i\rangle$, and whose N first elements on the diagonal are all equal to one. It is indeed the matrix associated with the projector P_N, and we can write:

$$\widehat{\rho}_1 = P_N \tag{50}$$

As we shall see, all the average values useful in our calculation can be simply expressed as a function of this operator.

2-a. Average energy

We now evaluate the different terms included in the average energy, starting with the terms containing one-particle operators.

α. Kinetic and external potential energy

Using relation (B-12) of Chapter XV, we obtain for the average kinetic energy $\left\langle \widehat{H}_0 \right\rangle$:

$$\left\langle \widehat{H}_0 \right\rangle = \sum_{r,s} \langle\theta_r| \frac{\mathbf{P}^2}{2m} |\theta_s\rangle \ \ \left\langle a_r^\dagger a_s \right\rangle \tag{51}$$

The same argument as that for the evaluation of the matrix elements (49) shows that the average value $\left\langle a_r^\dagger a_s \right\rangle$ in the state (47) is only different from zero if $r = s$; in that case, it is equal to one when $r \leq N$, and to zero otherwise. This leads to:

$$\left\langle \widehat{H}_0 \right\rangle = \sum_{i=1}^{N} \langle\theta_i| \frac{\mathbf{P}^2}{2m} |\theta_i\rangle = \mathrm{Tr}_1 \left\{ P_N \ \frac{\mathbf{P}^2}{2m} \right\} \tag{52}$$

The subscript 1 was added to the trace to underline the fact that this trace is taken in the one-particle state space and not in the Fock space. The two operators included in the trace only act on that same particle, numbered arbitrarily 1; the subscript 1 could obviously be replaced by the subscript of any other particle, since they all play the same role. The average potential energy coming from the external potential is computed in a similar way and can be written as:

$$\left\langle \widehat{V}_{\text{ext}} \right\rangle = \sum_{i=1}^{N} \langle \theta_i | \, V_1 \, | \theta_i \rangle = \text{Tr}_1 \left\{ P_N \; V_1 \right\} \tag{53}$$

β. Average interaction energy, Hartree-Fock potential operator

The average interaction energy $\widehat{W}_{\text{int}}$ can be computed using the general expression (C-16) of Chapter XV for any two-particle operator, which yields:

$$\left\langle \widehat{W}_{\text{int}} \right\rangle = \frac{1}{2} \sum_{r,s,n,q} \langle 1:\theta_r; 2:\theta_s | \, W_2(1,2) \, | 1:\theta_n; 2:\theta_q \rangle \;\; \left\langle a_r^\dagger a_s^\dagger a_q a_n \right\rangle \tag{54}$$

For the average value $\left\langle a_r^\dagger a_s^\dagger a_q a_n \right\rangle$ in the Fock state $\widetilde{|\Psi\rangle}$ to be different from zero, the operator must leave unchanged the populations of the individual states $|\theta_n\rangle$ and $|\theta_q\rangle$. As in § C-5-b of Chapter XV, two possibilities may occur: either $r = n$ and $s = q$ (the direct term), or $r = q$ and $s = n$ (the exchange term). Commuting some of the operators, we can write:

$$\begin{aligned} \left\langle a_r^\dagger a_s^\dagger a_q a_n \right\rangle &= \delta_{rn} \; \delta_{sq} \left\langle a_r^\dagger a_s^\dagger a_s a_r \right\rangle + \delta_{rq} \; \delta_{sn} \left\langle a_r^\dagger a_s^\dagger a_r a_s \right\rangle \\ &= \left[\delta_{rn} \; \delta_{sq} - \delta_{rq} \; \delta_{sn} \right] n_r n_s \end{aligned} \tag{55}$$

where n_r and n_s are the respective populations of the states $|\theta_r\rangle$ and $|\theta_s\rangle$. Now these populations are different from zero only if the subscripts r and s are between 1 and N, in which case they are equal to 1 (note also that we must have $r \neq s$ to avoid a zero result). We finally get[5]:

$$\begin{aligned} \left\langle \widehat{W}_{\text{int}} \right\rangle = \frac{1}{2} \sum_{i \neq j}^{N} \Big[& \langle 1:\theta_i; 2:\theta_j | \, W_2(1,2) \, | 1:\theta_i; 2:\theta_j \rangle \\ & - \langle 1:\theta_i; 2:\theta_j | \, W_2(1,2) \, | 1:\theta_j; 2:\theta_i \rangle \Big] \end{aligned} \tag{56}$$

(the constraint $i \neq j$ may be ignored since the right-hand side is equal to zero in this case). Here again, the subscripts 1 and 2 label two arbitrary, but different particles, that could have been labeled arbitrarily. We can therefore write:

$$\left\langle \widehat{W}_{\text{int}} \right\rangle = \frac{1}{2} \sum_{i,j=1}^{N} \langle 1:\theta_i; 2:\theta_j | \, W_2(1,2) \left[1 - P_{\text{ex}}(1,2) \right] | 1:\theta_i; 2:\theta_j \rangle \tag{57}$$

where $P_{\text{ex}}(1,2)$ is the exchange operator between particle 1 and 2 (the transposition which permutes them). This result can be written in a way similar to (53) by introducing a

[5] As in the previous complement, we have replaced $W_2(\mathbf{R}_1, \mathbf{R}_2)$ by $W_2(1,2)$ to simplify the notation

"Hartree-Fock potential" W_{HF}, similar to an external potential acting in the space of particle 1; this potential is defined as the operator having the matrix elements:

$$\langle\theta_i|\, W_{HF}(1)\,|\theta_k\rangle = \sum_{j=1}^{N} \langle 1:\theta_i; 2:\theta_j|\, W_2(1,2)\,[1-P_{\rm ex}(1,2)]\,|1:\theta_k; 2:\theta_j\rangle \tag{58}$$

This operator is Hermitian, since, as the two operators $P_{\rm ex}$ and W_2 are Hermitian and commute, we can write:

$$\begin{aligned}\langle\theta_k|\, W_{HF}(1)\,|\theta_i\rangle &= \sum_{j=1}^{N} \langle 1:\theta_k; 2:\theta_j|\, W_2(1,2)\,[1-P_{\rm ex}(1,2)]\,|1:\theta_i; 2:\theta_j\rangle \\ &= \sum_{j=1}^{N} \langle 1:\theta_i; 2:\theta_j|\,[1-P_{\rm ex}(1,2)]\, W_2(1,2)\,|1:\theta_k; 2:\theta_j\rangle^* \\ &= \langle\theta_i|\, W_{HF}(1)\,|\theta_k\rangle^* \end{aligned} \tag{59}$$

Furthermore, we recognize in (58) the matrix element of a partial trace on particle 2 (Complement $E_{\rm III}$, § 5-b):

$$\boxed{W_{HF}(1) = {\rm Tr}_2\left\{P_N(2)\; W_2(1,2)\,[1-P_{\rm ex}(1,2)]\right\}} \tag{60}$$

where the projector P_N has been introduced inside the trace to limit the sum over j to its first N terms, as in (57). The one-particle operator $W_{HF}(1)$ is thus the partial trace over a second particle (with the arbitrary label 2) of a product of operators acting on both particles. As the summation over j is now taken into account, we are left in (57) with a summation over i, which introduces a trace over the remaining particle 1, and we get:

$$\left\langle \widehat{W}_{\rm int} \right\rangle = \frac{1}{2}{\rm Tr}_1\left\{P_N(1)\; W_{HF}(1)\right\} \tag{61}$$

This average value depends on the subspace chosen with the variational ket $\widetilde{|\Psi\rangle}$ in two ways: explicitly as above, via the projector $P_N(1)$ that shows up in the average value (61), but also implicitly via the definition of the Hartree-Fock potential in (60).

γ. Role of the one-particle reduced density operator

All the average values can be expressed in terms of the projector P_N onto the subspace $\mathcal{F}_N$ of the space the individual states spanned by the N individual states $|\theta_1\rangle$, $|\theta_2\rangle$, ...$|\theta_N\rangle$, which means, according to (50), in terms of the one-particle reduced density operator $\widehat{\rho}_1 = P_N$. Hence it is this operator that is the pertinent variable to optimize rather than the set of individual states: certain variations of those states do not change P_N, and are meaningless for our purpose.

Furthermore, the choice of the trial ket $\widetilde{|\Psi\rangle}$ is equivalent to that of P_N. In other words, the variational ket $\widetilde{|\Psi\rangle}$ built in (1) does not depend on the basis chosen in the subspace $\mathcal{F}_N$: if we choose in this subspace any orthonormal basis $\{|u_j\rangle\}$ other than the $\{|\theta_j\rangle\}$ basis, and if we replace in (1) the $a^\dagger_{u_j}$ by the $a^\dagger_{\theta_i}$, the ket will remain the same

(to within a non-relevant phase factor) as we now show. As seen in § A-6 of Chapter XV, each operator $a^{\dagger}_{u_j}$ is a linear combination of the $a^{\dagger}_{\theta_i}$, so that in the product of all the $a^{\dagger}_{u_j}$ ($j = 1, 2, ..N$) we will find products of N operators $a^{\dagger}_{\theta_i}$. Relation (A-43) of Chapter XV however indicates that the squares of any creation operators are zero, which means that the only non-zero products are those including once and only once each of the N different operators $a^{\dagger}_{\theta_i}$. Each term is then proportional to the ket $\widetilde{|\Psi\rangle}$ built from the $a^{\dagger}_{\theta_i}$. Consequently, the two variational kets built from the two bases are necessarily proportional. As definition (1) ensures they are also normalized, they can only differ by a phase factor, which means they are equivalent from a physical point of view. It is thus the operator $P_N = \widehat{\rho}_1$ that best embodies the trial ket $\widetilde{|\Psi\rangle}$.

2-b. Optimization of the one-particle density operator

We now vary $P_N = \widehat{\rho}_1$ to look for the stationary conditions for the total energy:

$$\widetilde{E} = \langle H_0 \rangle + \langle V_1 \rangle + \left\langle \widehat{W}_{\text{int}} \right\rangle = \text{Tr}_1 \left\{ P_N \left[\frac{\mathbf{P}^2}{2m} + V_1 + \frac{1}{2} W_{HF} \right] \right\} \tag{62}$$

We therefore consider the variation:

$$P_N \Rightarrow P_N + \delta P_N \tag{63}$$

which leads to the following variations for the average values of the one-particle operators:

$$\delta \langle H_0 \rangle + \delta \langle V_1 \rangle = \text{Tr}_1 \left\{ \delta P_N \left[\frac{\mathbf{P}^2}{2m} + V_1 \right] \right\} \tag{64}$$

As for the interaction energy, we get two terms:

$$\delta \left\langle \widehat{W}_{\text{int}} \right\rangle = \frac{1}{2} \text{Tr}_1 \left\{ \delta P_N \; W_{HF} \right\} + \frac{1}{2} \text{Tr}_1 \left\{ P_N \; \delta W_{HF} \right\} \tag{65}$$

which are actually equal since:

$$\text{Tr}_1 \left\{ P_N(1) \; \delta W_{HF}(1) \right\} = \text{Tr}_{1,2} \left\{ P_N(1) \; \delta P_N(2) \; \widehat{W}_2(1,2) \left[1 - P_{\text{ex}}(1,2) \right] \right\} \tag{66}$$

and we recognize in the right-hand side of this expression the trace:

$$\text{Tr}_2 \left\{ \delta P_N(2) \; W_{HF}(2) \right\} \tag{67}$$

As we can change the label of the particle from 2 to 1 without changing the trace, the two terms of the interaction energy are equal. As a result, we end up with the energy variation:

$$\delta \widetilde{E} = \text{Tr}_1 \left\{ \left[\frac{\mathbf{P}^2}{2m} + V_1 + W_{HF} \right] \delta P_N \right\} \tag{68}$$

To vary the projector P_N, we choose a value j_0 of j and make the change:

$$|\theta_{j_0}\rangle \Rightarrow |\theta_{j_0}\rangle + e^{i\chi} \, |\delta\theta\rangle \tag{69}$$

where $|\delta\theta\rangle$ is any ket from the space of individual states, and χ any real number; no other individual state vector varies except for $|\theta_{j_0}\rangle$. The variation of P_N is then written as:

$$\delta P_N = e^{i\chi} |\delta\theta\rangle \langle\theta_{j_0}| + e^{-i\chi} |\theta_{j_0}\rangle \langle\delta\theta| \tag{70}$$

We assume $|\delta\theta\rangle$ has no components on any $|\theta_i\rangle$, that is no components in $\mathcal{F}_N$, since this would change neither $\mathcal{E}_N$, nor the corresponding projector P_N. We therefore impose:

$$P_N |\delta\theta\rangle = 0 \tag{71}$$

which also implies that the norm of $|\theta_{j_0}\rangle$ remains constant[6] to first order in $|\delta\theta\rangle$. Inserting (70) into (68), we obtain:

$$\delta\widetilde{E} = e^{i\chi}\mathrm{Tr}_1 \left\{ \left[\frac{\mathbf{P}^2}{2m} + V_1 + W_{HF} \right] |\delta\theta\rangle \langle\theta_{j_0}| \right\} + e^{-i\chi}\mathrm{Tr}_1 \left\{ \left[\frac{\mathbf{P}^2}{2m} + V_1 + W_{HF} \right] |\theta_{j_0}\rangle \langle\delta\theta| \right\} \tag{72}$$

For the energy to be stationary, this variation must remain zero whatever the choice of the arbitrary number χ. Now the linear combination of two exponentials $e^{i\chi}$ and $e^{-i\chi}$ will remain zero for any value of χ only if the two factors in front of the exponentials are zero themselves. As each term can be made equal to zero separately, we obtain:

$$0 = \mathrm{Tr}_1 \left\{ \left[\frac{\mathbf{P}^2}{2m} + V_1 + W_{HF} \right] |\theta_{j_0}\rangle \langle\delta\theta| \right\} = \langle\delta\theta| \left[\frac{\mathbf{P}^2}{2m} + V_1 + W_{HF} \right] |\theta_{j_0}\rangle \tag{73}$$

This relation must be satisfied for any ket $|\delta\theta\rangle$ orthogonal to the subspace $\mathcal{F}_N$. This means that if we define the one-particle Hartree-Fock operator as:

$$\boxed{H_{HF} = \frac{\mathbf{P}^2}{2m} + V_1 + W_{HF}} \tag{74}$$

the stationary condition for the total energy is simply that the ket $H_{HF} |\theta_{j_0}\rangle$ must belong to $\mathcal{F}_N$:

$$H_{HF} |\theta_{j_0}\rangle \in \mathcal{F}_N \tag{75}$$

As this relation must hold for any $|\theta_{j_0}\rangle$ chosen among the $|\theta_1\rangle$, $|\theta_2\rangle$, ...$|\theta_N\rangle$, it follows that the subspace $\mathcal{F}_N$ is stable under the action of the operator (74).

2-c. Mean field operator

We can then restrict the operator H_{HF} to that subspace:

$$\bar{H}_{HF} = P_N(1) \left[\frac{\mathbf{P}^2}{2m} + V_1(1) + W_{HF}(1) \right] P_N(1) \tag{76}$$

[6]Since (71) shows that $|\delta\theta\rangle$ is orthogonal to any linear combinations of the $|\theta_i\rangle$, we can write $(\langle\theta_{j_0}| + \langle\delta\theta|)\,(|\theta_{j_0}\rangle + |\delta\theta\rangle) = \langle\theta_{j_0}|\,\theta_{j_0}\rangle + \langle\delta\theta|\,\delta\theta\rangle = 1+$ second order terms.

This operator, acting in the subspace $\mathcal{F}_N$ spanned by the N kets $|\theta_j\rangle$, is a Hermitian linear operator, hence it can be diagonalized. We call $|\varphi_n\rangle$ its eigenvectors (n = 1, 2, ..N), which are linear combinations of the kets $|\theta_j\rangle$. The stationary condition for the energy (75) amounts to imposing the $|\varphi_n\rangle$ to be not only eigenvectors of $\bar{H}_{HF}$, but also of the operator H_{HF} defined by (74) in the entire one-particle state space (without the restriction to $\mathcal{F}_N$); consequently, the $|\varphi_n\rangle$ must obey:

$$\boxed{H_{HF}\,|\varphi_n\rangle = \widetilde{e}_n\ |\varphi_n\rangle} \tag{77}$$

Operator H_{HF} is defined in (74), where the operator W_{HF} is given by (60) and depends on the projector P_N. This last operator may be expressed as a function of the $|\varphi_n\rangle$ in the same way as with the $|\theta_i\rangle$, and relation (48) may be replaced by:

$$P_N = \sum_{n=1}^{N} |\varphi_n\rangle\,\langle\varphi_n| \tag{78}$$

Relations (77), together with definition (60) where (78) has been inserted, are a set of equations allowing the self-consistent determination of the $|\varphi_n\rangle$; they are called the Hartree-Fock equations. This operator form (77) is simpler than the one obtained in § 1-c; it emphasizes the similarity with the usual eigenvalue equation for a single particle moving in an external potential, illustrating the concept of a self-consistent mean field. One must keep in mind, however, that via the projector (78) included in W_{HF}, this particle moves in a potential depending on the whole set of states occupied by all the particles. Remember also that we did not carry out an exact computation, but merely presented an approximate theory (variational method).

The discussion in § 1-f is still relevant. As the operator W_{HF} depends on the $|\varphi_n\rangle$, the Hartree-Fock equations have an intrinsic nonlinear character, which generally requires a resolution by successive approximations. We start from a set of N individual states $\left|\varphi_n^0\right\rangle$ to build a first value of P_N and the operator W_{HF}, which are used to compute the Hamiltonian (74). Considering this Hamiltonian now fixed, the Hartree-Fock equations (77) become linear, and can be solved as usual eigenvalue equations. This leads to new values $\left|\varphi_n^1\right\rangle$ for the $|\varphi_n\rangle$, and finishes the first iteration. In the second iteration, we use the $\left|\varphi_n^1\right\rangle$ in (78) to compute a new value of the mean field operator W_{HF}; considering again this operator as fixed, we solve the eigenvalue equation and obtain the second iteration values $\left|\varphi_n^2\right\rangle$ for the $|\varphi_n\rangle$, and so on. If the initial values $\left|\varphi_n^0\right\rangle$ are physically reasonable, one can hope for a rapid convergence towards the expected solution of the nonlinear Hartree-Fock equations.

The variational energy can be computed in the same way as in § 1-e. Multiplying on the left equation (77) by the bra $\langle\varphi_n|$, we get:

$$\widetilde{e}_n = \langle\varphi_n|\left[\frac{\mathbf{P}^2}{2m} + V_1 + W_{HF}\right]|\varphi_n\rangle \tag{79}$$

After summing over the subscript n, we obtain:

$$\sum_{n=1}^{N}\widetilde{e}_n = \sum_{n=1}^{N}\langle\varphi_n|\left[\frac{\mathbf{P}^2}{2m} + V_1 + W_{HF}\right]|\varphi_n\rangle = \mathrm{Tr}_1\left\{P_N(1)\left[\frac{\mathbf{P}^2}{2m} + V_1 + W_{HF}\right]\right\} \tag{80}$$

Taking into account (51), (53), and (61), we get:

$$\sum_{n=1}^{N} \widetilde{e}_n = \langle H_0 \rangle_{HF} + \langle V_1 \rangle_{HF} + 2 \left\langle \widehat{W}_{\rm int} \right\rangle_{HF} \tag{81}$$

where the particle interaction energy is counted twice. To compute the energy $\widetilde{E}$, we can eliminate $\left\langle \widehat{W}_{\rm int} \right\rangle$ between (26) and this relation and we finally obtain:

$$\widetilde{E} = \frac{1}{2} \left[\sum_{n=1}^{N} \widetilde{e}_n + \langle H_0 \rangle_{HF} + \langle V_1 \rangle_{HF} \right] \tag{82}$$

2-d. Hartree-Fock equations for electrons

Assume the fermions we are studying are particles with spin 1/2, electrons for example. The basis $\{|\mathbf{r}\rangle\}$ of the individual states used in § 1 must be replaced by the basis formed with the kets $\{|\mathbf{r}, \nu\rangle\}$, where ν is the spin index, which can take 2 distinct values noted $\pm 1/2$, or more simply $\pm$. To the summation over $\mathrm{d}^3 r$ we must now add a summation over the 2 values of the index spin ν. A vector $|\varphi\rangle$ in the individual state space is now written:

$$|\varphi\rangle = \sum_{\nu = \pm 1/2} \int \mathrm{d}^3 r \; \varphi(\mathbf{r}, \nu) \, |\mathbf{r}, \nu\rangle \tag{83}$$

with:

$$\varphi(\mathbf{r}, \nu) = \langle \mathbf{r}, \nu \, | \varphi \rangle \tag{84}$$

The variables $\mathbf{r}$ and ν play a similar role but the first one is continuous whereas the second is discrete. Writing them in the same parenthesis might hide this difference, and we often prefer noting the discrete index as a superscript of the function φ, and write:

$$\varphi^{\nu}(\mathbf{r}) = \langle \mathbf{r}, \nu \, | \varphi \rangle \tag{85}$$

Let us build an N particle variational state $\widetilde{|\Psi\rangle}$ from N orthonormal states $|\varphi_n^{\nu_n}\rangle$, with $n = 1, 2, .. , N$. Each of the $|\varphi_n^{\nu_n}\rangle$ describes an individual state including the spin and position variables; the first N_+ values of ν_n ($n = 1, 2, ..N_+$) are equal to $+1/2$, the last N_- are equal to $-1/2$, with $N_+ + N_- = N$ (we assume N_+ and N_- are fixed for the moment but we may allow them to vary later to enlarge the variational family). In the space of the individual states, we introduce a complete basis $\{|\varphi_k^{\nu_k}\rangle\}$ whose first N kets are the $|\varphi_n^{\nu_n}\rangle$, but where the subscript k varies from 1 to infinity[7].

We assume the matrix elements of the external potential V_1 to be diagonal for ν; these two diagonal matrix elements can however take different values $V_1^{\pm}(\mathbf{r})$, which allows including the eventual presence of a magnetic field coupled with the spins. We also assume the particle interaction $W_2(1,2)$ to be independent of the spins, and diagonal in the position representation of the two particles, as is the case, for example, for the Coulomb

[7] The subscript k determines both the orbital and the spin state of the particle; the index ν is not independent since it is fixed for each value of k.

interaction between electrons. With these assumptions, the Hamiltonian cannot couple states having different particle numbers N_+ and N_-.

Let us see what the general Hartree-Fock equations become in the $\{|\mathbf{r},\nu\rangle\}$ representation. In this representation, the effect of the kinetic and potential operators are well known. We just have to compute the effect of the Hartree-Fock potential W_{HF}. To obtain its matrix elements, we use the basis $\{|1:\mathbf{r},\nu;\ 2:\varphi_k^{\nu_k}\rangle\}$ to write the trace in (60):

$$\langle \mathbf{r},\nu|\, W_{HF}(1)\,|\mathbf{r}',\nu'\rangle = \sum_{k=1}^{\infty}\ \langle 1:\mathbf{r},\nu;2:\varphi_k^{\nu_k}|\left[\sum_{p=1}^{N}\left|2:\varphi_p^{\nu_p}\right\rangle\left\langle 2:\varphi_p^{\nu_p}\right|\right] W_2(1,2)\,[1-P_{\rm ex}(1,2)]\,|1:\mathbf{r}',\nu';2:\varphi_k^{\nu_k}\rangle \tag{86}$$

As the right-hand side includes the scalar product $\langle 2:\varphi_k^{\nu_k}\ |2:\varphi_p^{\nu_p}\rangle$ which is equal to δ_{kp}, the sum over k disappears and we get:

$$\langle \mathbf{r},\nu|W_{HF}(1)\,|\mathbf{r}',\nu'\rangle = \sum_{p=1}^{N}\left\langle 1:\mathbf{r},\nu;2:\varphi_p^{\nu_p}\right| W_2(1,2)\,[1-P_{\rm ex}(1,2)]\left|1:\mathbf{r}',\nu';2:\varphi_p^{\nu_p}\right\rangle \tag{87}$$

(i) We first deal with the direct term contribution, hence ignoring in the bracket the term in $P_{\rm ex}(1,2)$. We can replace the ket $\left|2:\varphi_p^{\nu_p}\right\rangle$ by its expression:

$$\left|2:\varphi_p^{\nu_p}\right\rangle = \int \mathrm{d}^3 r_2\ \varphi_p^{\nu_p}(\mathbf{r}_2)\,|2:\mathbf{r}_2,\nu_p\rangle \tag{88}$$

As the operator is diagonal in the position representation, we can write:

$$\begin{aligned} W_2(1,2)\left|1:\mathbf{r}',\nu';2:\varphi_p^{\nu_p}\right\rangle &= W_2(1,2)\int \mathrm{d}^3 r_2\ \varphi_p^{\nu_p}(\mathbf{r}_2)\,|1:\mathbf{r}',\nu';2:\mathbf{r}_2,\nu_p\rangle \\ &= \int \mathrm{d}^3 r_2\ \varphi_p^{\nu_p}(\mathbf{r}_2)\ W_2(\mathbf{r}',\mathbf{r}_2)\,|1:\mathbf{r}',\nu';2:\mathbf{r}_2,\nu_p\rangle \end{aligned} \tag{89}$$

The direct term of (87) is then written:

$$\int \mathrm{d}^3 r_2 \sum_{p=1}^{N} W_2(\mathbf{r}',\mathbf{r}_2)\ \varphi_p^{\nu_p}(\mathbf{r}_2)\ \left\langle 1:\mathbf{r},\nu;2:\varphi_p^{\nu_p}\ |1:\mathbf{r}',\nu';2:\mathbf{r}_2,\nu_p\right\rangle \tag{90}$$

where the scalar product of the bra and the ket is equal to $\delta_{\nu\nu'}\delta(\mathbf{r}-\mathbf{r}')\left[\varphi_p^{\nu_p}(\mathbf{r}_2)\right]^*$. We finally obtain:

$$\delta_{\nu\nu'}\ \delta(\mathbf{r}-\mathbf{r}')\int \mathrm{d}^3 r_2\ W_2(\mathbf{r},\mathbf{r}_2)\sum_{p=1}^{N}\left|\varphi_p^{\nu_p}(\mathbf{r}_2)\right|^2 = \delta_{\nu\nu'}\ \delta(\mathbf{r}-\mathbf{r}')V_{\rm dir}(\mathbf{r}) \tag{91}$$

with:

$$V_{\rm dir}(\mathbf{r}) = \sum_{p=1}^{N}\int \mathrm{d}^3 r'\ W_2(\mathbf{r},\mathbf{r}')\ \left|\varphi_p^{\nu_p}(\mathbf{r}')\right|^2 \tag{92}$$

This component of the mean field (Hartree term) contains a sum over all occupied states, whatever their spin is; it is spin independent.

(ii) We now turn to the exchange term, which contains the operator $P_{\rm ex}(1,2)$ in the bracket of (87). To deal with it, we can for example commute in (87) the two operators $W_2(1,2)$ and $P_{\rm ex}(1,2)$; this last operator will then permute the two particles in the bra. Performing this operation in (90), we get, with the minus sign of the exchange term:

$$-\int d^3r_2 \sum_{p=1}^{N} W_2(\mathbf{r}',\mathbf{r}_2)\ \varphi_p^{\nu_p}(\mathbf{r}_2)\ \left\langle 1:\varphi_p^{\nu_p};2:\mathbf{r},\nu\ |1:\mathbf{r}',\nu';2:\mathbf{r}_2,\nu_p\right\rangle \tag{93}$$

The scalar product will yield the products of $\delta_{\nu\nu_p}\delta_{\nu_p\nu'}\delta(\mathbf{r}-\mathbf{r}_2)$, making the integral over d^3r_2 disappear; this term is zero if $\nu\neq\nu'$, hence the factor $\delta_{\nu\nu'}$. Since $W_2(\mathbf{r}',\mathbf{r}) = W_2(\mathbf{r},\mathbf{r}')$, we are left with:

$$-\delta_{\nu\nu'}\sum_{\nu_p=\nu} W_2(\mathbf{r},\mathbf{r}')\left[\varphi_p^{\nu_p}(\mathbf{r})\right]\left[\varphi_p^{\nu_p}(\mathbf{r}')\right]^* = -\delta_{\nu\nu'}V_{ex}^{\nu}(\mathbf{r},\mathbf{r}') \tag{94}$$

where the sum is over the values of p for which $\nu_p=\nu=\nu'$ (hence, limited to the first N_+ values of p, or the last N_-, depending on the case); the exchange potential $V_{\rm ex}^{\nu}$ has been defined as:

$$V_{\rm ex}^{\nu}(\mathbf{r},\mathbf{r}') = \sum_{\nu_p=\nu} W_2(\mathbf{r},\mathbf{r}')\ \left[\varphi_p^{\nu_p}(\mathbf{r}')\right]^*\left[\varphi_p^{\nu_p}(\mathbf{r})\right] \tag{95}$$

As is the case for the direct term, the exchange term does not act on the spin. There are however two differences. To begin with, the summation over p is limited to the states having the same spin ν; second, it introduces a contribution which is non-diagonal in the positions (but without an integral), and which cannot be reduced to an ordinary potential (the term "non-local potential" is sometimes used to emphasize this property).

We have shown that the scalar product of equation (77) with $\langle\mathbf{r},\nu|$ introduces three potentials (in addition to the the one-body potential $V_1^{\nu_n}$), a direct potential $V_{\rm dir}(\mathbf{r})$ and two exchange potentials $V_{\rm ex}^{\nu}(\mathbf{r},\mathbf{r}')$ with $\nu=\pm1/2$. Equation (77) then becomes, in the $\{|\mathbf{r},\nu\rangle\}$ representation, a pair of equations:

$$\boxed{\left[-\frac{\hbar^2}{2m}\Delta + V_1^{\nu_n}(\mathbf{r}) + V_{\rm dir.}(\mathbf{r})\right]\varphi_n^{\nu_n}(\mathbf{r}) - \int \mathrm{d}^3r'\ V_{\rm ex}^{\nu_n}(\mathbf{r},\mathbf{r}')\varphi_n^{\nu_n}(\mathbf{r}') = \widetilde{e}_n\ \varphi_n^{\nu_n}(\mathbf{r})} \tag{96}$$

These are the Hartree-Fock equations with spin and in the position representation, widely used in quantum physics and chemistry. It is not necessary to worry, in these equations, about the term in which the subscript p in the summation appearing in (92) and (95) is the same as the subscript n (of the wave function we are looking for); the contributions $n=p$ cancel each other exactly in the direct and exchange potentials.

Both the "Hartree term" giving the direct potential contribution, and the "Fock term " giving the exchange potential, can be interpreted in the same way as above (§ 1-f). The Hartree term contains the contributions of all the other electrons to the mean potential felt by one electron. The exchange potential, on the other hand, only involves electrons in the same spin state, and this can be simply interpreted: the exchange effect only occurs for two totally indistinguishable particles. Now if these particles are in

orthogonal spin states, and as the interactions do not act on the spins, one can in principle determine which is which and the particles become distinguishable: the quantum exchange effects cancel out. As we already pointed out, the exchange potential is not a potential stricto sensu. It is not diagonal in the position representation, even though it basically comes from a particle interaction that is diagonal in position. It is the antisymmetrization of the fermions, together with the chosen variational approximation, which led to this peculiar non-diagonal form. It is however a Hermitian operator, as can be shown using the fact that the initial potential $W_2(\mathbf{r}, \mathbf{r}')$ is real and symmetric with respect to $\mathbf{r}$ and $\mathbf{r}'$.

2-e. Discussion

The resolution of the nonlinear Hartree-Fock equations is generally done by the successive iteration approximate method discussed in § 1-f. There is no particular reason for the solution of the Hartree-Fock equations to be unique[8]; on the contrary, they can yield solutions that depend on the states chosen to begin the nonlinear iterations. They can actually lead to a whole spectrum of possible energies for the system. This is how the ground state and excited state energies of the atom are generally computed. The atomic orbitals discussed in Complement E_{VII}, the central field approximation and the electronic "configurations" discussed in Complement B_{XIV} can now be discussed in a more precise and quantitative way. We note that the exchange energy, introduced in this complement for a two-electron system, is a particular case of the exchange energy term of the Hartree-Fock potential. There exist however many other physical systems where the same ideas can be applied: nuclei (the Coulomb force is then replaced by the nuclear interaction force between the nucleons), atomic aggregates (with an interatomic potential having both repulsive and attractive components, see Complements C_{XI} and G_{XI}), and many others.

Once a Hartree-Fock solution for a complex problem has been found, we can go further. One can use the basis of the eigenfunctions just obtained as a starting point for more precise perturbation calculations, including for example correlations between particles (Chapter XI). In atomic spectra, we sometimes find cases where two configurations yield very close mean field energies. The effects of the interaction terms beyond the mean field approximation will then be more important. Perturbation calculations limited to the space of the configurations in question permits obtaining better approximations for the energy levels and their wave functions; one then speaks of "mixtures", or of "interactions between configurations".

Comment:

The variational method based on the Fock states is not the only one that leads to the Hartree-Fock equations. One could also start from an approximation of the two-particle density operator ρ_{II} by a function of the one-particle density operator ρ_I and write:

$$\rho_{II}(1,2) \propto \frac{1}{2}\left[1 - P_{\text{ex}}(1,2)\right] \rho_I(1) \otimes \rho_I(2) \tag{97}$$

Expressing the energy of the N-particle system as a function of ρ_I, we minimize it by varying this operator, and find the same results as above. This method amounts to a closure of the hierarchy of the N-body equations (§ C-4 of Chapter XVI). We have in

[8] They all yield, however, an upper limit for the ground state energy

fact already seen with equation (21) and in § 2-a that the Hartree-Fock approximation amounts to expressing the two-particle correlation functions as a function of the one-particle correlation functions. In terms of correlation functions (Complement A_{XVI}), this amounts to replacing the two-particle function (four-point function) by a product of one-particle functions (two-point function), including an exchange term. Finally, another method is to use the diagram perturbation theory; the Hartree-Fock approximation corresponds to retaining only a certain class of diagrams (class of connected diagrams).

Finally note that the Hartree-Fock method is not the only one yielding approximate solutions of Schrödinger's equation for a system of interacting fermions; in particular, one can use the "electronic density functional" theory (a functional is a function of another function, as for instance the action S in classical lagrangian mechanics). The method is used to obtain the electronic structure of molecules or condensed phases in physics, chemistry, and materials science. Its study nevertheless lies outside the scope of this book, and the reader is referred to [6], which summarizes the method and gives a number of references.

Complement F_{XV}

Fermions, time-dependent Hartree-Fock approximation

1 Variational ket and notation 1701
2 Variational method 1702
2-a Definition of a functional 1702
2-b Stationarity 1703
2-c Particular case of a time-independent Hamiltonian 1705
3 Computing the optimizer 1705
3-a Average energy 1705
3-b Hartree-Fock potential 1706
3-c Time derivative 1707
3-d Functional value 1707
4 Equations of motion 1707
4-a Time-dependent Hartree-Fock equations 1708
4-b Particles in a single spin state 1709
4-c Discussion 1709

The Hartree-Fock mean field method was introduced in Complement E_{XV} for a time-independent problem: the search for the stationary states of a system of interacting fermions (the search for its thermal equilibrium will be discussed in Complement G_{XV}. In this complement, we show how this method can be used for time-dependent problems. We start, in § 1, by including a time dependence in the Hartree-Fock variational ket (time-dependent Fock state). We then introduce in § 2 a general variational principle that can be used for solving the time-dependent Schrödinger equation. We then compute, in § 3, the function to be optimized for a Fock state; the same mean field operator as the one introduced in Complement E_{XV} will here again play a very useful role. Finally, the time-dependent Hartree-Fock equations will be obtained and discussed in § 4. More details on the Hartree-Fock methods in general can be found, for example, in Chapter 7 of reference [5], and especially in its Chapter 9 for time-dependent problems.

1. Variational ket and notation

We assume the N-particle state vector $\widetilde{|\Psi(t)\rangle}$ to be of the form:

$$\widetilde{|\Psi(t)\rangle} = a^\dagger_{\theta_1(t)} a^\dagger_{\theta_2(t)} \ ... \ a^\dagger_{\theta_N(t)} |0\rangle \tag{1}$$

where the $a^\dagger_{\theta_1(t)}$, $a^\dagger_{\theta_2(t)}$, ..., $a^\dagger_{\theta_N(t)}$ are the creation operators associated with an arbitrary series of orthonormal individual states $|\theta_1(t)\rangle$, $|\theta_2(t)\rangle$, ..., $|\theta_N(t)\rangle$ which depend on time t. This series is, for the moment, arbitrary, but the aim of the following variational calculation is to determine its time dependence.

As in the previous complements, we assume that the Hamiltonian $\widehat{H}$ is the sum of three terms: a kinetic energy Hamiltonian, an external potential Hamiltonian, and a particle interaction term:

$$\widehat{H} = \widehat{H}_0 + \widehat{V}_{\text{ext}}(t) + \widehat{W}_{\text{int}} \tag{2}$$

with:

$$\widehat{H}_0 = \sum_{q=1}^{N} \frac{(\mathbf{P}_q)^2}{2m} \tag{3}$$

(m is the particles' mass, $\mathbf{P}_q$ the momentum operator of particle q), and:

$$\widehat{V}_{\text{ext}}(t) = \sum_{q=1}^{N} V_1(\mathbf{R}_q, t) \tag{4}$$

and finally:

$$\widehat{W}_{\text{int}} = \frac{1}{2} \sum_{q \neq q'} W_2(\mathbf{R}_q, \mathbf{R}_{q'}) \tag{5}$$

2. Variational method

Let us introduce a general variational principle; using the stationarity of a functional S of the state vector $|\Psi(t)\rangle$, it will yield the time-dependent Schrödinger equation.

2-a. Definition of a functional

Consider an arbitrarily given Hamiltonian $H(t)$. We assume the state vector $|\Psi(t)\rangle$ to have any time dependence, and we note $\left|\overline{\Psi}(t)\right\rangle$ the ket physically equivalent to $|\Psi(t)\rangle$, but with a constant norm:

$$\left|\overline{\Psi}(t)\right\rangle = \frac{|\Psi(t)\rangle}{\sqrt{\langle\Psi(t)\ |\Psi(t)\rangle}} \tag{6}$$

The functional S of $\left|\overline{\Psi}(t)\right\rangle$ is defined as[1]:

$$\begin{aligned} S\left[\left|\overline{\Psi}(t)\right\rangle\right] &= \int_{t_0}^{t_1} \mathrm{d}t\ \ \mathrm{Re}\ \ \left\langle\overline{\Psi}(t)\right| \left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - H(t)\right] \left|\overline{\Psi}(t)\right\rangle \\ &= \int_{t_0}^{t_1} \mathrm{d}t \left\{ \frac{i\hbar}{2} \left[\left\langle\overline{\Psi}(t)\right| \frac{\mathrm{d}}{\mathrm{d}t}\ \overline{\Psi}(t)\right\rangle - \left\langle \frac{\mathrm{d}}{\mathrm{d}t}\overline{\Psi}(t)\ \left|\overline{\Psi}(t)\right\rangle\right] - \left\langle\overline{\Psi}(t)\right| H(t) \left|\overline{\Psi}(t)\right\rangle \right\} \end{aligned} \tag{7}$$

where t_0 and t_1 are two arbitrary times such that $t_0 < t_1$. In the particular case where the chosen $\left|\overline{\Psi}(t)\right\rangle$ is equal to a solution $\left|\overline{\Psi}_S(t)\right\rangle$ of the Schrödinger equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left|\overline{\Psi}_S(t)\right\rangle = H(t)\left|\overline{\Psi}_S(t)\right\rangle \tag{8}$$

[1]The notation where the differential operator $\mathrm{d}/\mathrm{d}t$ is written between a bra and a ket means that the operator takes the derivative of the ket that follows (and not of the bra just before).

the bracket on the first line of (7) obviously cancels out and we have:

$$S\left[\left|\overline{\Psi}_S(t)\right\rangle\right] = 0 \tag{9}$$

Integrating by parts the second term[2] of the bracket in the second line of (7), we get the same form as the first term in the bracket, plus an already integrated term. The final result is then:

$$\begin{aligned} S\left[\left|\overline{\Psi}(t)\right\rangle\right] &= \left\{\int_{t_0}^{t_1} \mathrm{d}t \ \left\langle\overline{\Psi}(t)\right| \left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - H(t)\right] \left|\overline{\Psi}(t)\right\rangle\right\} \\ &\qquad + \frac{i\hbar}{2}\left[\left\langle\overline{\Psi}(t_0)\right| \left.\overline{\Psi}(t_0)\right\rangle - \left\langle\overline{\Psi}(t_1)\right| \left.\overline{\Psi}(t_1)\right\rangle\right] \\ &= \int_{t_0}^{t_1} \mathrm{d}t \ \left\langle\overline{\Psi}(t)\right| \left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - H(t)\right] \left|\overline{\Psi}(t)\right\rangle \end{aligned} \tag{10}$$

where we have used in the second line the fact that the norm of $\left|\overline{\Psi}(t)\right\rangle$ always remains equal to unity. This expression for S is similar to the initial form (7), but without the real part.

2-b. Stationarity

Suppose now $\left|\overline{\Psi}(t)\right\rangle$ has an arbitrary time dependence between t_0 and t_1, while keeping its norm constant, as imposed by (6); the functional then takes a certain value S, a priori different from zero. Let us see under which conditions S will be stationary when $\left|\overline{\Psi}(t)\right\rangle$ changes by an infinitely small amount $\left|\delta\overline{\Psi}(t)\right\rangle$:

$$\left|\overline{\Psi}(t)\right\rangle \Rightarrow \left|\overline{\Psi}(t)\right\rangle + \left|\delta\overline{\Psi}(t)\right\rangle \tag{11}$$

For what follows, it will be convenient to assume that the variation $\left|\delta\overline{\Psi}(t)\right\rangle$ is free; we therefore have to ensure that the norm of $\left|\overline{\Psi}(t)\right\rangle$ remains constant, equal to unity[3]. We introduce Lagrange multipliers (Appendix V) $\lambda(t)$ to control the square of the norm at every time between t_0 and t_1, and we look for the stationarity of a function where the sum of constraints has been added. This sum introduces an integral, and we the function in question is:

$$\begin{aligned} \overline{S}\left[\left|\overline{\Psi}(t)\right\rangle\right] &= S\left[\left|\overline{\Psi}(t)\right\rangle\right] - \int_{t_0}^{t_1} \mathrm{d}t \ \lambda(t) \left\langle\overline{\Psi}(t) \ \left|\overline{\Psi}(t)\right\rangle\right. \\ &= \int_{t_0}^{t_1} \mathrm{d}t \ \left\langle\overline{\Psi}(t)\right| \left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - H(t) - \lambda(t)\right] \left|\overline{\Psi}(t)\right\rangle \end{aligned} \tag{12}$$

where $\lambda(t)$ is a real function of the time t.

[2] If we integrate by parts the first term rather than the second, we get the complex conjugate of equation (10), which brings no new information.

[3] For the normalization of $\left|\overline{\Psi}(t)\right\rangle$ to be conserved to first order, it is necessary (and sufficient) for the scalar product $\left\langle\overline{\Psi}(t) \ \left|\delta\overline{\Psi}(t)\right\rangle\right.$ to be zero or purely imaginary. If this is the case, the Lagrangian multiplier $\lambda(t)$ is not needed

The variation $\delta\overline{S}$ of $\overline{S}$ to first order is obtained by inserting (11) in (10). It yields the sum of a first term $\delta\overline{S}_1$ containing the ket $\left|\delta\overline{\Psi}(t)\right\rangle$ and of another $\delta\overline{S}_2$ containing the bra $\left\langle\delta\overline{\Psi}(t)\right|$:

$$\begin{aligned}\delta\overline{S}_1 &= \int_{t_0}^{t_1} \mathrm{d}t \ \left\langle\overline{\Psi}(t)\right| \left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - H(t) - \lambda(t)\right] \left|\delta\overline{\Psi}(t)\right\rangle \\ \delta\overline{S}_2 &= \int_{t_0}^{t_1} \mathrm{d}t \ \left\langle\delta\overline{\Psi}(t)\right| \left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - H(t) - \lambda(t)\right] \left|\overline{\Psi}(t)\right\rangle \end{aligned} \tag{13}$$

We now imagine another variation for the ket:

$$\left|\overline{\Psi}(t)\right\rangle \Rightarrow \left|\overline{\Psi}(t)\right\rangle + i\left|\delta\overline{\Psi}(t)\right\rangle \tag{14}$$

which yields a variation $\delta'\overline{S}$ of $\overline{S}$; in this second variation, the term in $\left|\delta\overline{\Psi}(t)\right\rangle$ becomes $\delta'\overline{S}_1 = i\delta\overline{S}_1$, whereas the term in $\left\langle\delta\overline{\Psi}(t)\right|$ becomes $\delta'\overline{S}_2 = -i\delta\overline{S}_2$. Now, if the functional is stationary in the vicinity of $\left|\overline{\Psi}(t)\right\rangle$, the two variations $\delta\overline{S}$ and $\delta'\overline{S}$ are necessarily zero, as are also $\delta\overline{S} - i\delta'\overline{S}$ and $\delta\overline{S} + i\delta'\overline{S}$. In those combinations, only terms in $\delta\overline{S}_1$ appear for the first one, and in $\delta\overline{S}_2$ for the second; consequently they must both be zero. As a result, we can write the stationarity conditions with respect to variations of the bra and the ket separately.

Let us write for example that $\delta\overline{S}_2 = 0$, which means the right-hand side of the second line in (13) must be zero. As the time evolution between t_0 and t_1 of the bra $\left\langle\delta\overline{\Psi}(t)\right|$ is arbitrary, this condition imposes this bra multiplies a zero-value ket, at all times. Consequently, the ket $\left|\overline{\Psi}(t)\right\rangle$ must obey the equation:

$$\left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - H(t) - \lambda(t)\right] \left|\overline{\Psi}(t)\right\rangle = 0 \tag{15}$$

which is none other than the Schrödinger equation associated with the Hamiltonian $H(t) + \lambda(t)$.

Actually, $\lambda(t)$ simply introduces a change in the origin of the energies and this only modifies the total phase[4] of the state vector $\left|\overline{\Psi}(t)\right\rangle$, which has no physical effect. Without loss of generality, this Lagrange factor may therefore be ignored, and we can set:

$$\lambda(t) = 0 \tag{16}$$

A necessary condition[5] for the stationarity of S is that $\left|\overline{\Psi}(t)\right\rangle$ obey the Schrödinger equation (8) – or be physically equivalent (i.e. equal to within a global time-dependent phase factor) to a solution of this equation. Conversely, assume $\left|\overline{\Psi}(t)\right\rangle$ is a solution of the Schrödinger equation, and give this ket a variation as in (11). It is then obvious from the second line of (13) that $\delta\overline{S}_2$ is zero. As for $\delta\overline{S}_1$, an integration by parts over time shows that it is the complex conjugate of $\delta\overline{S}_2$, and therefore also equal to zero. The

[4] If in (15) we set $\left|\overline{\Psi}(t)\right\rangle = e^{-i\alpha(t)}\left|\overline{\Theta}(t)\right\rangle$, we see that $\left|\overline{\Theta}(t)\right\rangle$ obeys the differential equation obtained by replacing $\lambda(t)$ by $\lambda(t) - \hbar\frac{\mathrm{d}\alpha}{\mathrm{d}t}$ in (15). If we simply choose for $\alpha(t)$ the integral over time of the function $\lambda(t)$, this constant will disappear from the differential equation.

[5] The same argument as above, but starting from the variation $\delta S - i\delta' S$, would lead to the complex conjugate of (8), and hence to the same equation.

functional S is thus stationary in the vicinity of any exact solution of the Schrödinger equation.

Suppose we choose any variational family $\mathcal{F}$ of normalized kets $\widetilde{|\Psi_{\mathcal{F}}(t)\rangle}$, but which now includes a ket $\widetilde{|\Psi^{\text{stat}}_{\mathcal{F}}(t)\rangle}$ for which S is stationary. A simple example is the case where $\mathcal{F}$ is a family $\mathcal{F}_0$ that contains the exact solution of the Schrödinger equation; according to what we just saw, this exact solution will make S stationary, and conversely, the ket that makes S stationary is necessarily $\widetilde{\left|\Psi^{\text{stat}}_{\mathcal{F}_0}(t)\right\rangle}$. In this case, imposing the variation of S to be zero allows identifying, inside the family $\mathcal{F}_0$, the exact solution we are looking for. If we now change the family continuously from $\mathcal{F}_0$ to $\mathcal{F}$, in general $\mathcal{F}$ will no longer contain the exact solution of the Schrödinger equation. We can however follow the modifications at all times of the values of the ket $\widetilde{|\Psi^{\text{stat}}_{\mathcal{F}}(t)\rangle}$. Starting from an exact solution of the equation, this ket progressively changes, but, by continuity, will stay in the vicinity of this exact solution if $\mathcal{F}$ stays close to $\mathcal{F}_0$. This is why annulling the variation of S in the family $\mathcal{F}$ is a way of identifying a member of that family whose evolution remains close to that of a solution of the Schrödinger equation. This is the method we will follow, using the Fock states as a particular variational family.

2-c. Particular case of a time-independent Hamiltonian

If the Hamiltonian H is time-independent, one can look for time-independent kets $\left|\overline{\Psi}\right\rangle$ to make the functional S stationary. The function to be integrated in the definition of the functional S also becomes time-independent, and we can write S as:

$$S = (t_1 - t_0) \left\langle \overline{\Psi} \right| H \left| \overline{\Psi} \right\rangle \tag{17}$$

Since the two times t_0 and t_1 are fixed, the stationarity of S is equivalent to that of the diagonal matrix element of the Hamiltonian $\left\langle \overline{\Psi} \right| H \left| \overline{\Psi} \right\rangle$. We find again the stationarity condition of the time-independent variational method (Complement E_{XI}), which appears as a particular case of the more general method of the time-dependent variations. Consequently, it is not surprising that the Hartree-Fock methods, time-dependent or not, lead to the same Hartree-Fock potential, as we now show.

3. Computing the optimizer

The family of the state vectors we consider is the set of Fock kets $\widetilde{|\Psi(t)\rangle}$ defined in (1). We first compute the function to be integrated in the functional (10) when $|\Psi(t)\rangle$ takes the value $\widetilde{|\Psi(t)\rangle}$.

3-a. Average energy

For the term in $H(t)$, the calculation is identical to the one we already did in § 1-b of Complement E_{XV}. We first add to the series of orthonormal states $|\theta_i(t)\rangle$ with $i = 1, 2, ..., N$ other orthonormal states $|\theta_i(t)\rangle$ with $i = N+1, N+2, ...$, to obtain a complete orthonormal basis in the space of individual states. Using this basis, we can express the one-particle and two-particle operators according to relations (B-12) and (C-16) of Chapter XV. This presents no difficulty since the average values of creation and annihilation operator products are easily obtained in a Fock state (they only differ

from zero if the product of operators leaves the populations of the individual states unchanged). Relations (52), (53) and (57) of Complement E_{XV} are still valid when the $|\theta_i\rangle$ become time-dependent. We thus get for the average kinetic energy:

$$\left\langle \widehat{H}_0 \right\rangle = \sum_{i=1}^{N} \langle \theta_i(t)| \frac{\mathbf{P}^2}{2m} |\theta_i(t)\rangle \tag{18}$$

for the external potential energy:

$$\left\langle \widehat{V}_{\rm ext}(t) \right\rangle = \sum_{i=1}^{N} \langle \theta_i(t)| V_1(\mathbf{R},t) |\theta_i(t)\rangle \tag{19}$$

and for the interaction energy:

$$\left\langle \widehat{W}_{\rm int} \right\rangle = \frac{1}{2} \sum_{i,j}^{N} \langle 1:\theta_i(t); 2:\theta_j(t)| W_2(1,2)\,[1 - P_{\rm ex}(1,2)]\,|1:\theta_i(t); 2:\theta_j(t)\rangle \tag{20}$$

3-b. Hartree-Fock potential

We recognize in (20) the diagonal element ($i = k$) of the Hartree-Fock potential operator $W_{HF}(1,t)$ whose matrix elements have been defined in a general way by relation (58) of Complement E_{XV}:

$$\begin{aligned} &\langle \theta_i(t)|W_{HF}(1,t)\,|\theta_k(t)\rangle \\ &\qquad = \sum_{j}^{N} \langle 1:\theta_i(t); 2:\theta_j(t)| W_2(1,2)\,[1 - P_{\rm ex}(1,2)]\,|1:\theta_k(t); 2:\theta_j(t)\rangle \end{aligned} \tag{21}$$

We also noted in that complement E_{XV} that $W_{HF}(1,t)$ is a Hermitian operator.

It is often handy to express the Hartree-Fock potential using a partial trace:

$$\boxed{W_{HF}(1,t) = {\rm Tr}_2\Big\{ P_N(2,t)\; W_2(1,2)\,[1 - P_{\rm ex}(1,2)] \Big\}} \tag{22}$$

where P_N is the projector onto the subspace spanned by the N kets $|\theta_i(t)\rangle$:

$$P_N(2,t) = \sum_{i=1}^{N} |2:\theta_i(t)\rangle\,\langle 2:\theta_i(t)| \tag{23}$$

As we have seen before, this projector is actually nothing bu the one-particle reduced density operator $\widehat{\rho}_1$ normalized by imposing its trace to be equal to the total particle number N:

$$P_N(2,t) = \widehat{\rho}_1(2,t) \tag{24}$$

The average value of the interaction energy can then be written as:

$$\boxed{\left\langle \widehat{W}_{\rm int} \right\rangle = \frac{1}{2} {\rm Tr}_1\Big\{ \widehat{\rho}_1(1,t)\; W_{HF}(1,t) \Big\}} \tag{25}$$

3-c. Time derivative

As for the time derivative term, the function it contains can be written as:

$$\sum_{i=1}^{N} \langle 0|\, a_{\theta_N}(t)...\; a_{\theta_i}(t)...\; a_{\theta_1}(t)\; a^{\dagger}_{\theta_1}(t)...\left[\frac{\mathrm{d}}{\mathrm{d}t}a^{\dagger}_{\theta_i}(t)\right]...\; a^{\dagger}_{\theta_N}(t)\,|0\rangle \tag{26}$$

In this summation, all terms involving the individual states j other than the state i (which is undergoing the derivation) lead to an expression of the type:

$$\langle 0|\, a_{\theta_j}(t)a^{\dagger}_{\theta_j}(t)\,|0\rangle \tag{27}$$

which equals 1 since this expression is the square of the norm of the state $a^{\dagger}_{\theta_j}(t)\,|0\rangle$, which is simply the Fock state $|n_j = 1\rangle$. As for the state i, it leads to a factor written in the form of a scalar product in the one-particle state space:

$$\langle 0|\, a_{\theta_i}(t)\frac{\mathrm{d}}{\mathrm{d}t}a^{\dagger}_{\theta_i}(t)\,|0\rangle = \langle \theta_i(t)|\,\frac{\mathrm{d}}{\mathrm{d}t}\,\theta_i(t)\rangle \tag{28}$$

3-d. Functional value

Regrouping all these results, we can write the value of the functional S in the form:

$$\begin{aligned} S\left[\widetilde{|\Psi(t)\rangle}\right] = \sum_{i=1}^{N}\int_{t_0}^{t_1}\mathrm{d}t\;\Bigg\{ & i\hbar\;\langle\theta_i(t)|\,\frac{\mathrm{d}}{\mathrm{d}t}\,\theta_i(t)\rangle - \langle\theta_i(t)|\left[\frac{\mathbf{P}^2}{2m}+V_1(t)\right]|\theta_i(t)\rangle \\ & - \frac{1}{2}\sum_{j=1}^{N}\langle 1:\theta_i(t);2:\theta_j(t)|\,W_2(1,2)\,[1-P_{\mathrm{ex}}(1,2)]\,|1:\theta_i(t);2:\theta_j(t)\rangle\Bigg\} \end{aligned} \tag{29}$$

4. Equations of motion

We now vary the ket $|\theta_k(t)\rangle$ according to:

$$|\theta_k(t)\rangle \Rightarrow |\theta_k(t)\rangle + |\delta\theta_k(t)\rangle \qquad \text{with} \qquad k \leq N \tag{30}$$

As in complement $\mathrm{E_{XV}}$, we will only consider variations $|\delta\theta_k(t)\rangle$ that lead to an actual variation of the ket $\widetilde{|\Psi(t)\rangle}$; those where $|\delta\theta_k(t)\rangle$ is proportional to one of the occupied states $|\theta_l(t)\rangle$ with $l \leq N$ yield no change for $\widetilde{|\Psi(t)\rangle}$ (or at the most to a phase change) and are thus irrelevant for the value of S. As we did in relations (32) or (69) of Complement $\mathrm{E_{XV}}$, we assume that:

$$|\delta\theta_k(t)\rangle = \delta f(t)e^{i\chi}\,|\theta_l(t)\rangle \qquad \text{with} \qquad l > N \tag{31}$$

where $\delta f(t)$ is an infinitesimal time-dependent function.

The computation is then almost identical to that of § 2-b in Complement $\mathrm{E_{XV}}$. When $|\theta_k(t)\rangle$ varies according to (31), all the other occupied states remaining constant, the only changes in the first line of (29) come from the terms $i = k$. In the second line, the changes come from either the $i = k$ terms, or the $j = k$ terms. As the $W_2(1,2)$

operator is symmetric with respect to the two particles, these variations are the same and their sum cancels the 1/2 factor. All these variations involve terms containing either the ket $e^{i\chi}\,|\delta\theta_k(t)\rangle$, or the bra $\langle\delta\theta_k(t)|\,e^{-i\chi}$. Now their sum must be zero for any value of χ, and this is only possible if each of the terms is zero. Inserting the variation (31) of $|\theta_k(t)\rangle$, and canceling the term in $e^{-i\chi}$ leads to the following equality:

$$\int_{t_0}^{t_1} \mathrm{d}t\ \delta f(t) \left\{ i\hbar \langle\theta_l(t)| \frac{\mathrm{d}}{\mathrm{d}t}\ \theta_k(t)\rangle - \langle\theta_l(t)| \left[\frac{\mathbf{P}^2}{2m} + V_1(t)\right] |\theta_k(t)\rangle \right.$$
$$\left. - \sum_{j=1}^{N} \langle 1:\theta_l(t); 2:\theta_j(t)|\, W_2(1,2)\,[1 - P_{\mathrm{ex}}(1,2)]\, |1:\theta_k(t); 2:\theta_j(t)\rangle \right\} = 0 \tag{32}$$

As we recognize in the function to be integrated the Hartree-Fock potential operator $W_{HF}(1,t)$ defined in (21), we can write:

$$\int_{t_0}^{t_1} \mathrm{d}t\ \delta f(t) \left\{ \langle\theta_l(t)| \left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - \frac{\mathbf{P}^2}{2m} - V_1(t) - W_{HF}(t)\right] |\theta_k(t)\rangle \right\} = 0 \tag{33}$$

with $l > N$.

4-a. Time-dependent Hartree-Fock equations

As the choice of the function $\delta f(t)$ is arbitrary, for expression (33) to be zero for any $\delta f(t)$ requires the function inside the curly brackets to be zero at all times t. Stationarity therefore requires the ket:

$$\left[i\hbar\frac{\mathrm{d}}{\mathrm{d}t} - \frac{\mathbf{P}^2}{2m} - V_1(t) - W_{HF}(t)\right] |\theta_k(t)\rangle \tag{34}$$

to have no components on any of the non-occupied states $|\theta_l(t)\rangle$ with $(l > N)$. In other words, stationarity will be obtained if, for all values of k between 1 and N, we have:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\,|\theta_k(t)\rangle = \left[\frac{\mathbf{P}^2}{2m} + V_1(t) + W_{HF}(t)\right] |\theta_k(t)\rangle + |\xi_k(t)\rangle \tag{35}$$

where $|\xi_k(t)\rangle$ is any linear combination of the occupied states $|\theta_l(t)\rangle$ $(l \leq N)$. As we pointed out at the beginning of § 4, adding to one of the $|\theta_k(t)\rangle$ a component on the already occupied individual states has no effect on the N-particle state (aside from an eventual change of phase), and therefore does not change the value of S; consequently, the stationarity of this functional does not depend on the value of the ket $|\xi_k(t)\rangle$, which can be any ket, for example the zero ket.

Finally, if the $|\theta_n(t)\rangle$ are equal to the solutions $|\varphi_n(t)\rangle$ of the N equations:

$$\boxed{i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\,|\varphi_n(t)\rangle = \left[\frac{\mathbf{P}^2}{2m} + V_1(t) + W_{HF}(t)\right] |\varphi_n(t)\rangle} \tag{36}$$

the functional S is indeed stationary for all times. Furthermore, as we saw in Complement E_{XV} that $W_{HF}(t)$ is Hermitian, so is the operator on the right-hand side of (36).

Consequently, the N kets $|\varphi_n(t)\rangle$ follow an evolution similar to the usual Schrödinger evolution, described by a unitary evolution operator (Complement $\mathrm{F_{III}}$). Such an operator does not change either the norm nor the scalar products of the kets: if the kets $|\varphi_n(t)\rangle$ initially formed an orthonormal set, this remains true at any later time. The whole calculation just presented is thus consistent; in particular, the norm of the N-particle state vector $|\widetilde{\Psi(t)}\rangle$ is constant over time.

Relations (36) are the time-dependent Hartree-Fock equations. Introducing the one-particle mean field operator allowed us not only to compute the stationary energy levels, but also to treat time-dependent problems.

4-b. Particles in a single spin state

Let us return to the particular case of fermions all having the same spin state, as in § 1 of Complement $\mathrm{E_{XV}}$. We can then write the Hartree-Fock equations in terms of the wave functions as:

$$i\hbar\frac{\partial}{\partial t}\,\varphi_n(\mathbf{r},t) = \left[-\frac{\hbar^2}{2m}\Delta + V_1(\mathbf{r}) + V_{\mathrm{dir}}(\mathbf{r};t)\right]\varphi_n(\mathbf{r},t) - \int \mathrm{d}^3r'\; V_{\mathrm{ex}}(\mathbf{r},\mathbf{r}';t)\varphi_n(\mathbf{r}',t) \tag{37}$$

using definitions of (46) of that complement for the direct and exchange potentials, which are now time-dependent. There is obviously a close relation between the Hartree-Fock equations, whether they are time-dependent or not.

4-c. Discussion

As encountered in the search for a ground state with the time-independent Hartree-Fock equations, there is a strong similarity between equations (36) and an ordinary Schrödinger equation for a single particle. Here again, an exact solution of these equations is generally not possible, and we must use successive approximations. Assume for example that the external time-dependent potential $V_1(t)$ is zero until time t_0 and that for $t < t_0$, the physical system is in a stationary state. With the time-independent Hartree-Fock method we can compute an approximate value for this state and hence a series of initial values for the individual states $|\varphi_n(t_0)\rangle$. This determines the initial Hartree-Fock potential. Between time t_0 and a slightly later time $t_0 + \Delta t$, the evolution equation (36) describes the effect of the external potential $V_1(t)$ on the individual kets, and allows obtaining the $|\varphi_n(t_0 + \Delta t)\rangle$. We can then compute a new value for the Hartree-Fock potential, and use it to extend the computation of the evolution of the $|\varphi_n(t)\rangle$ until a later time $t_0 + 2\Delta t$. Proceeding step by step, we can obtain this evolution until the final time t_1. For the approach to be precise, Δt must be small enough for the Hartree-Fock potential to change only slightly from one time step to another.

Another possibility is to proceed as in the search for the stationary states. We start from a first family of orthonormal kets, now time-dependent, and which are not too far from the expected solution over the entire time interval; we then try to improve it by successive iterations. Inserting in (21) the first series of orthonormal trial functions, we get a first approximation of the Hartree-Fock potential and its associated dynamics. We then solve the corresponding equation of motion, with the same initial conditions at

$t = t_0$, which yields a new series of orthonormal functions. Using again (21), we get a value for the Hartree-Fock potential, a priori different from the previous one. We start the same procedure anew until an acceptable convergence is obtained.

Applications of this method are quite numerous, in particular in atomic, molecular, and nuclear physics. They allow, for example, the study of the electronic cloud oscillations in an atom, a molecule or a solid, placed in an external time-dependent electric field (dynamic polarisability), or the oscillations of nucleons in their nucleus. We mentioned in the conclusion of Complement E_{XV} that the time-independent Hartree-Fock method is sometimes replaced by the functional density method; this is also the case when dealing with time-dependent problems.

In concluding this complement we underline the close analogy between the Hartree-Fock theory and a time-independent or a time-dependent mean field theory. In both cases the same Hartree-Fock potential operators come into play. Even though they are the result of an approximation, these operators have a very large range of applicability.

Complement G_{XV}

Fermions or Bosons: Mean field thermal equilibrium

1 Variational principle 1712
- 1-a Notation, statement of the problem 1712
- 1-b A useful inequality 1713
- 1-c Minimization of the thermodynamic potential 1715

2 Approximation for the equilibrium density operator 1716
- 2-a Trial density operators 1716
- 2-b Partition function, distributions 1717
- 2-c Variational grand potential 1721
- 2-d Optimization 1722

3 Temperature dependent mean field equations 1725
- 3-a Form of the equations 1726
- 3-b Properties and limits of the equations 1727
- 3-c Differences with the zero-temperature Hartree-Fock equations (fermions) 1728
- 3-d Zero-temperature limit (fermions) 1729
- 3-e Wave function equations 1729

Understanding the thermal equilibrium of a system of interacting identical particles is important for many physical problems: conductor or semiconductor electronic properties, liquid Helium or ultra-cold gas properties, etc. It is also essential for studying phase transitions, various and multiple examples of which occur in solid and liquid physics: spontaneous magnetism appearing below a certain temperature, changes in electrical conduction, and many others. However, even if the Hamiltonian of a system of identical particles is known, calculation of the equilibrium properties cannot, in general, be carried to completion: these calculations present real difficulties in the handling of state vectors and interaction operators, where non-trivial combinations of creation and annihilation operators occur. One must therefore use one or several approximations. The most common one is probably the mean field approximation, which, as we saw in Complement E_{XV}, is the base of the Hartree-Fock method. In that complement, we showed, in terms of state vectors, how this method could be used to obtain approximate values for the energy levels of a system of interacting particles. As we consider here the more complex problem of thermal equilibrium, which must be treated in terms of density operators, we show how the Hartree-Fock method can be extended to this more general case.

We are going to see that, thanks to this approach, one can obtain compact formulas for an approximate value of the density operator at thermal equilibrium, in the framework of the grand canonical ensemble. The equations to be solved are fairly similar[1] to those of Complement E_{XV}. The Hartree-Fock method also gives a value of the

[1] They are not simply the juxtaposition of that complement's equations: one could imagine writing

thermodynamic grand potential, which leads directly to the pressure of the system. The other thermodynamic quantities can then be obtained via partial derivatives with respect to the equilibrium parameters (volume, temperature, chemical potential, eventually external applied field, etc. – see Appendix VI). It is clearly a powerful method even though it still is an approximation as the particles interactions are treated via a mean field approach where certain correlations are not taken into account. Furthermore, for bosons, it can only be applied to physical systems far from Bose-Einstein condensation; the reasons for this limitation will be discussed in detail in § 4-a of Complement H_{XV}.

Once we have recalled the notation and a few generalities, we shall establish (§ 1) a variational principle that applies to any density operator. It will allow us to search in any family of operators for the one closest to the density operator at thermal equilibrium. We will then introduce (§ 2) a family of trial density operators whose form reflects the mean field approximation; the variational principle will help us determine the optimal operator. We shall obtain Hartree-Fock equations for a non-zero temperature, and study some of their properties in the last section (§ 3). Several applications of these equations will be presented in Complement H_{XV}.

The general idea and the structure of the computations will be the same as in Complement E_{XV}, and we keep the same notation: we establish a variational condition, choose a trial family, and then optimize the system description within this family. This is why, although the present complement is self-contained, it might be useful to first read Complement E_{XV}.

1. Variational principle

In order to use a certain number of general results of quantum statistical mechanics (see Appendix refappend-6 for a more detailed review), we first introduce the notation.

1-a. Notation, statement of the problem

We assume the Hamiltonian is of the form:

$$\widehat{H} = \widehat{H}_0 + \widehat{V}_{\text{ext}} + \widehat{W}_{\text{int}} \tag{1}$$

which is the sum of the particles' kinetic energy $\widehat{H}_0$, their coupling energy $\widehat{V}_{\text{ext}}$ with an external potential:

$$\widehat{V}_{\text{ext}} = \sum_q V_1(q) \tag{2}$$

and their mutual interaction $\widehat{W}_{\text{int}}$, which can be expressed as:

$$\widehat{W}_{\text{int}} = \frac{1}{2} \sum_{q \neq q'} W_2(\mathbf{R}_q, \mathbf{R}_{q'}) \tag{3}$$

We are going to use the "grand canonical" ensemble (Appendix VI, § 1-c), where the particle number is not fixed, but takes on an average value determined by the chemical

those equations independently for each energy level, and then performing a thermal average. We are going to see (for example in § 2-d-β) that the determination of each level's position already implies thermal averages, meaning that the levels are coupled.

potential μ. In this case, the density operator ρ is an operator acting in the entire Fock space $\mathcal{E}_F$ (where N can take on all the possible values), and not only in the state space $\mathcal{E}_N$ for N particles (which is is more restricted since it corresponds to a fixed value of N). We set, as usual:

$$\beta = \frac{1}{k_B T} \tag{4}$$

where k_B is the Boltzmann constant and T the absolute temperature. At the grand canonical equilibrium, the system density operator depends on two parameters, β and the chemical potential μ, and can be written as:

$$\rho_{\text{eq}} = \frac{1}{Z} \exp\left[-\beta\left(\widehat{H} - \mu\widehat{N}\right)\right] \tag{5}$$

with the relation that comes from normalizing to 1 the trace of ρ_{eq}:

$$Z = \text{Tr}\left\{\exp\left[-\beta\left(\widehat{H} - \mu\widehat{N}\right)\right]\right\} \tag{6}$$

The function Z is called the "grand canonical partition function" (see Appendix VI, § 1-c). The operator $\widehat{N}$ associated with the total particle number is defined in (B-17) of Chapter XV. The temperature T and the chemical potential μ are two intensive quantities, respectively conjugate to the energy and the particle number.

Because of the particle interactions, these formulas generally lead to calculations too complex to be carried to completion. We therefore look, in this complement, for approximate expressions of ρ_{eq} and Z that are easier to use and are based on the mean field approximation.

1-b. A useful inequality

Consider two density operators ρ and ρ', both having a trace equal to 1:

$$\text{Tr}\{\rho\} = \text{Tr}\{\rho'\} = 1 \tag{7}$$

As we now show, the following relation is always true:

$$\text{Tr}\{\rho \ln\rho\} \geq \text{Tr}\{\rho \ln\rho'\} \tag{8}$$

We first note that the function $x \ln x$, defined for $x \geq 0$, is always larger than the function $x - 1$, which is the equation of its tangent at $x = 1$ (Fig. 1). For positive values of x and y we therefore always have:

$$\frac{x}{y} \ln\frac{x}{y} \geq \frac{x}{y} - 1 \tag{9}$$

or, after multiplying by y:

$$x \ln x - x \ln y \geq x - y \tag{10}$$

the equality occurring only if $x = y$.

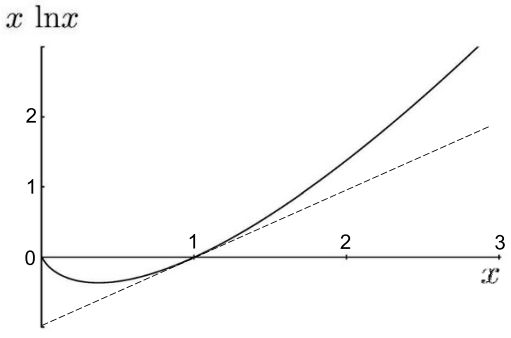


Figure 1: Plot of the function x lnx. At $x = 1$, this curve is tangent to the line $y = x - 1$ (dashed line) but always remains above it; the function value is thus always larger than $x - 1$.

Let us call p_n the eigenvalues of ρ corresponding to the normalized eigenvectors $|u_n\rangle$, and p'_m the eigenvalues of ρ' corresponding to the normalized eigenvectors $|v_m\rangle$. Used for the positive numbers p_n and p'_m, relation (10) yields:

$$p_n \ln p_n - p_n \ln p'_m \geq p_n - p'_m \tag{11}$$

We now multiply this relation by the square of the modulus of the scalar product:

$$|\langle u_n|\, v_m\rangle|^2 = \langle u_n|\, v_m\rangle\, \langle v_m|\, u_n\rangle \tag{12}$$

and sum over m and n. For the term in $p_n \ln p_n$ of (11), the summation over m yields in (12) the identity operator expanded on the basis $\{|v_m\rangle\}$; we then get $\langle u_n|\, u_n\rangle = 1$, and are left with the sum over n of $p_n \ln p_n$, that is the trace $\mathrm{Tr}\{\rho \ln \rho\}$. As for the term in $p_n \ln p'_m$, the summation over m introduces:

$$\sum_m \ln p'_m \;\; |v_m\rangle\, \langle v_m| = \ln \rho' \tag{13}$$

and we get:

$$\sum_n \langle u_n|\, p_n \ln \rho' \,|u_n\rangle = \mathrm{Tr}\,\{\rho \ln \rho'\} \tag{14}$$

As for the terms on the right-hand side of inequality (11), the term in p_n yields:

$$\sum_{m,n} \langle v_m|\, u_n\rangle\, p_n\, \langle u_n|\, v_m\rangle = \sum_m \langle v_m|\, \rho\, |v_m\rangle = \mathrm{Tr}\,\{\rho\} = 1 \tag{15}$$

and the one in p'_m also yields 1 for the same reasons, and both terms cancel out. We finally obtain the inequality:

$$\mathrm{Tr}\,\{\rho\ \ln\rho\} - \mathrm{Tr}\,\{\rho\ \ln\rho'\} \geq 0 \tag{16}$$

which proves (8).

Comment:

One may wonder under which conditions the above relation becomes an equality. This requires the inequality (11) to become an equality, which means $p_n = p'_m$ whenever the scalar product (12) is non-zero; consequently all the eigenvalues of the two operators ρ and ρ' must be equal. In addition, the eigenvectors of each operator corresponding to different eigenvalues must be orthogonal (their scalar product must be zero). In other words, the eigenvalues and the subspaces spanned by their eigenvectors are identical, which amounts to saying that $\rho = \rho'$.

1-c. Minimization of the thermodynamic potential

The entropy S associated with any density operator ρ having a trace equal to 1 is defined by relation (6) of Appendix VI:

$$S = -k_B\ \mathrm{Tr}\,\{\rho\ \ln\rho\} \tag{17}$$

The thermodynamic potential of the grand canonical ensemble is defined by the "grand potential" Φ, which can be expressed as a function of ρ by relation (Appendix VI, § 1-c-β):

$$\Phi = \left\langle \widehat{H} \right\rangle - TS - \mu \left\langle \widehat{N} \right\rangle = \mathrm{Tr}\left\{ \left[\widehat{H} + k_B T\ \ln\rho - \mu\widehat{N} \right] \rho \right\} \tag{18}$$

Inserting (5) into (18), we see that the value of Φ at equilibrium, Φ_{eq}, can be directly obtained from the partition function Z:

$$\begin{aligned} \Phi_{\mathrm{eq}} &= \mathrm{Tr}\left\{ \left[\widehat{H} + \beta^{-1} \left(-\beta\widehat{H} + \beta\mu\widehat{N} - \ln Z \right) - \mu\widehat{N} \right] \rho_{\mathrm{eq}} \right\} \\ &= -\beta^{-1} \ln Z\ \mathrm{Tr}\,\{\rho_{\mathrm{eq}}\} = -k_B T\ \ln Z \end{aligned} \tag{19}$$

We therefore have:

$$Z = e^{-\beta\Phi_{eq}} \qquad \text{or} \qquad \Phi_{eq} = -k_B T\ \ln Z \tag{20}$$

Consider now any density operator ρ and its associated function Φ obtained from (18). According to (5) and (20), we can write:

$$-\beta \left(\widehat{H} - \mu\widehat{N} \right) = \ln\rho_{eq} + \ln Z = \ln\rho_{eq} - \beta\Phi_{eq} \tag{21}$$

Inserting this result in (18) yields:

$$\Phi = \mathrm{Tr}\left\{ \beta^{-1} \left[-\ln\rho_{eq} + \beta\Phi_{eq} + \ln\rho \right] \rho \right\} = \beta^{-1} \mathrm{Tr}\left\{ \left[-\ln\rho_{eq} + \ln\rho \right] \rho \right\} + \Phi_{eq} \tag{22}$$

Now relation (16), used with $\rho' = \rho_{\mathrm{eq}}$, is written as:

$$\mathrm{Tr}\left\{ \left[\ln\rho - \ln\rho_{\mathrm{eq}} \right] \rho \right\} \geq 0 \tag{23}$$

Relation (22) thus implies that for any density operator ρ having a trace equal to 1, we have:

$$\Phi \geq \Phi_{\text{eq}} \tag{24}$$

the equality occurring if, and only if, $\rho = \rho_{\text{eq}}$.

Relation (24) can be used to fix a variational principle: choosing a family of density operators ρ having a trace equal to 1, we try to identify in this family the operator that yields the lowest value for Φ. This operator will then be the optimal operator within this family. Furthermore, this operator yields an upper value for the grand potential, with an error of second order with respect to the error made on ρ.

2. Approximation for the equilibrium density operator

We now use this variational principle with a family of density operators that leads to manageable calculations.

2-a. Trial density operators

The Hartree-Fock method is based on the assumption that a good approximation is to consider that each particle is independent of the others, but moving in the mean potential they create. We therefore compute an approximate value of the density operator by replacing the Hamiltonian $\widehat{H}$ by a sum of independent particles' Hamiltonians $\widetilde{h}(q)$:

$$\widetilde{H} = \sum_{q=1}^{N} \widetilde{h}(q) \tag{25}$$

We now introduce the basis of the creation and annihilation operators, associated with the eigenvectors of the one-particle operator $\widetilde{h}$:

$$a_k^\dagger |0\rangle = \widetilde{|\theta_k\rangle} \qquad \text{with} \quad \widetilde{h}\widetilde{|\theta_k\rangle} = \widetilde{e}_k \widetilde{|\theta_k\rangle} \tag{26}$$

The symmetric one-particle operator $\widetilde{H}$ can then be written, according to relation (B-14) of Chapter XV:

$$\widetilde{H} = \sum_k \widetilde{e}_k \; a_k^\dagger a_k \tag{27}$$

where the real constants $\widetilde{e}_k$ are the eigenvalues of the operator $\widetilde{h}$.

We choose as trial operators acting in the Fock space the set of operators $\widetilde{\rho}$ that can be written in the form corresponding to an equilibrium in the grand canonical ensemble – see relation (42) of Appendix VI. We then set:

$$\widetilde{\rho} = \frac{1}{\widetilde{Z}} \exp\left[-\beta\left(\widetilde{H} - \mu\widehat{N}\right)\right] \tag{28}$$

where $\widetilde{H}$ is any symmetric one-particle operator, the constant β the inverse of the temperature defined in (4), μ a real constant playing the role of a chemical potential, and $\widetilde{Z}$ the trace of $\widetilde{\rho}$:

$$\widetilde{Z} = \text{Tr}\left\{\exp\left[-\beta\left(\widetilde{H} - \mu\widehat{N}\right)\right]\right\} \tag{29}$$

Consequently, the relevant variables in our problem are the states $\widetilde{|\theta_k\rangle}$, which form an arbitrary orthonormal basis in the individual state space, and the energies $\widetilde{e}_k$. These variables determine the a_k as well as $\widetilde{H}$, and we have to find which of their values minimizes the function:

$$\widetilde{\Phi} = \mathrm{Tr}\left\{\left[\widehat{H} + k_B T\, \ln\widetilde{\rho} - \mu\widehat{N}\right]\widetilde{\rho}\right\} \tag{30}$$

Taking (27) and (28) into account, we can write:

$$\widetilde{\rho} = \frac{1}{\widetilde{Z}} \exp\left[-\sum_k \beta\left(\widetilde{e}_k - \mu\right) a_k^\dagger a_k\right] \tag{31}$$

The following computations are simplified since the Fock space can be considered to be the tensor product of independent spaces associated with the individual states $\widetilde{|\theta_k\rangle}$; consequently, the trial density operator (28) can be written as a tensor product of operators each acting on a single mode k:

$$\widetilde{\rho} = \frac{1}{\widetilde{Z}} \prod_k \exp\left[-\beta\left(\widetilde{e}_k\ - \mu\right) a_k^\dagger a_k\right] \tag{32}$$

2-b. Partition function, distributions

Equality (32) has the same form as relation (5) of Complement $\mathrm{B_{XV}}$, with a simple change: the replacement of the free particle energies $e_k = \hbar^2 k^2/2m$ by the energies $\widetilde{e}_k$, which are as yet unknown. As this change does not impact the mathematical structure of the density operator, we can directly use the results of Complement $\mathrm{B_{XV}}$.

α. Variational partition function

The function $\widetilde{Z}$ only depends on the variational energies $\widetilde{e}_k$, since the trace of (32) may be computed in the basis $\left\{\widetilde{|\theta_k\rangle}\right\}$, which yields:

$$\widetilde{Z} = \prod_k \sum_{n_k} \exp\left[-n_k \beta\left(\widetilde{e}_k - \mu\right)\right] \tag{33}$$

We simply get an expression similar to relation (7) of Complement $\mathrm{B_{XV}}$, obtained for an ideal gas. Since for fermions n_k can only take the values 0 and 1, we get:

$$\widetilde{Z}_{fermions} = \prod_k \left[1 + e^{-\beta\left(\widetilde{e}_k - \mu\right)}\right] \tag{34}$$

whereas for bosons n_k varies from 0 to infinity, so that:

$$\widetilde{Z}_{bosons} = \prod_k \frac{1}{1 - e^{-\beta\left(\widetilde{e}_k - \mu\right)}} \tag{35}$$

In both cases we can write:

$$\ln\widetilde{Z} = -\eta \sum_k \ln\left[1 - \eta e^{-\beta\left(\widetilde{e}_k - \mu\right)}\right] \tag{36}$$

with $\eta = +1$ for bosons, and $\eta = -1$ for fermions.

Computing the entropy can be done in a similar way. As the density operator $\widetilde{\rho}$ has the same form as the one describing the thermal equilibrium of an ideal gas, we can use for a system described by $\widetilde{\rho}$ the formulas obtained for the entropy of a system without interactions.

β. One particle, reduced density operator

Let us compute the average value of $a_i^\dagger a_j$ with the density operator $\widetilde{\rho}$:

$$\left\langle a_i^\dagger a_j \right\rangle_{\widetilde{\rho}} = \mathrm{Tr}\left\{ a_i^\dagger a_j \widetilde{\rho} \right\} \tag{37}$$

We saw in § 2-c of Complement B_{XV} that:

$$\mathrm{Tr}\left\{ a_i^\dagger a_j \widetilde{\rho} \right\} = \delta_{ij}\ f_\beta(\widetilde{e}_i - \mu) \tag{38}$$

where the distribution function f_β is noted f_β^{FD} for fermions, and f_β^{BE} for bosons:

$$f_\beta(e-\mu) = \begin{cases} f_\beta^{FD}(e-\mu) = \dfrac{1}{e^{\beta(e-\mu)}+1} & \text{for fermions} \\ f_\beta^{BE}(e-\mu) = \dfrac{1}{e^{\beta(e-\mu)}-1} & \text{for bosons} \end{cases} \tag{39}$$

When the system is described by the density operator $\widetilde{\rho}$, the average populations of the individual states $\widetilde{|\theta_k\rangle}$ are therefore determined by the usual Fermi-Dirac or Bose-Einstein distributions. From now on, and to simplify the notation, we shall write simply $|\theta_k\rangle$ for the kets $\widetilde{|\theta_k\rangle}$.

We can introduce a "one-particle reduced density operator" $\widetilde{\rho}_1(1)$ by[2] :

$$\widetilde{\rho}_1(1) = \sum_k f_\beta(\widetilde{e}_k - \mu)\ \ |1:\theta_k\rangle\langle 1:\theta_k| \tag{40}$$

where the 1 enclosed in parentheses and the subscript 1 on the left-hand side emphasize we are dealing with an operator acting in the one-particle state space (as opposed to $\widetilde{\rho}$ that acts in the Fock space); needless to say, this subscript has nothing to do with the initial numbering of the particles, but simply refers to any single particle among all the system particles. The diagonal elements of $\widetilde{\rho}_1(1)$ are the individual state populations. With this operator, we can compute the average value over $\widetilde{\rho}$ of any one-particle operator $\widehat{M}$:

$$\left\langle \widehat{M} \right\rangle_{\widetilde{\rho}} = \mathrm{Tr}\left\{ \widehat{M}\ \widetilde{\rho} \right\} \tag{41}$$

as we now show. Using the expression (B-12) of Chapter XV for any one-particle opera-

[2] Contrary to what is usually the case for a density operator, the trace of this reduced operator is not equal to 1, but to the average particle number – see relation (44). This different normalization is often more useful when studying systems composed of a large number of particles.

tor[3], as well as (38), we can write:

$$\left\langle \widehat{M} \right\rangle_{\widetilde{\rho}} = \sum_{k,l} \langle \theta_k | \, \widehat{m} \, | \theta_l \rangle \, \mathrm{Tr} \left\{ a_k^\dagger a_l \widetilde{\rho} \right\} = \sum_k \langle \theta_k | \, \widehat{m} \, | \theta_k \rangle \;\; f_\beta(\widetilde{e}_k - \mu)$$
$$= \sum_k \langle \theta_k | \, \widehat{m} \;\, \widetilde{\rho}_1(1) \, | \theta_k \rangle \tag{42}$$

that is:

$$\left\langle \widehat{M} \right\rangle_{\widetilde{\rho}} = \mathrm{Tr}_1 \left\{ \widehat{m} \;\, \widetilde{\rho}_1(1) \right\} \tag{43}$$

As we shall see, the density operator $\widetilde{\rho}_1(1)$ is quite useful since it allows obtaining in a simple way all the average values that come into play in the Hartree-Fock computations. Our variational calculations will simply amount to varying $\widetilde{\rho}_1(1)$. This operator presents, in a certain sense, all the properties of the variational density operator $\widetilde{\rho}$ chosen in (28) in the Fock space. It plays the same role[4] as the projector P_N (which also represents the essence of the variational N-particle ket) played in Complement $\mathrm{E_{XV}}$. In a general way, one can say that the basic principle of the Hartree-Fock method is to reduce the binary correlation functions of the system to products of single-particle correlation functions (more details on this point will be given in § 2-b of Complement $\mathrm{C_{XVI}}$).

The average value of the operator $\widehat{N}$ for the total particle number is written:

$$\left\langle \widehat{N} \right\rangle_{\widetilde{\rho}} = \mathrm{Tr} \left\{ \widehat{N} \;\, \widetilde{\rho} \right\} = \mathrm{Tr}_1 \left\{ \widetilde{\rho}_1(1) \right\} = \sum_{i=1}^{\infty} f_\beta(\widetilde{e}_i - \mu) \tag{44}$$

Both functions f_β^{FD} and f_β^{BE} increase as a function of μ and, for any given temperature, the total particle number is controlled by the chemical potential. For a large physical system whose energy levels are very close, the orbital part of the discrete sum in (44) can be replaced by an integral. Figure 1 of Complement $\mathrm{B_{XV}}$ shows the variations of the Fermi-Dirac and Bose-Einstein distributions. We also mentioned that for a boson system, the chemical potential could not exceed the lowest value e_0 of the energies e_l; when it approaches that value, the population of the corresponding level diverges, which is the Bose-Einstein condensation phenomenon we will come back to in the next complement. For fermions, on the other hand, the chemical potential has no upper boundary, as, whatever its value, the population of states having an energy lower than μ cannot exceed 1.

γ. *Two particles, distribution functions*

We now consider an arbitrary two-particle operator $\widehat{G}$ and compute its average value with the density operator $\widetilde{\rho}$. The general expression of a symmetric two-particle

[3] We have changed the notation $\widehat{F}$ and $\widehat{f}$ of Chapter XV into $\widehat{M}$ and $\widehat{m}$ to avoid any confusion with the distribution functions f_β.

[4] For fermions, and when the temperature approaches zero, the distribution function f_β^{FD} included in the definition of $\rho_I(1)$ becomes a step function and $\rho_I(1)$ does indeed coincide with $P_N(1)$.

operator is given by relation (C-16) of Chapter XV, and we can write:

$$\begin{aligned}\left\langle \widehat{G} \right\rangle_{\widetilde{\rho}} &= \mathrm{Tr}\left\{ \widehat{G}\ \widetilde{\rho} \right\} \\ &= \frac{1}{2} \sum_{i,j,k,l} \langle 1:\theta_i; 2:\theta_j |\, \widehat{g}(1,2) \,|1:\theta_k; 2:\theta_l\rangle\, \mathrm{Tr}\left\{ a_i^\dagger a_j^\dagger a_l a_k\ \widetilde{\rho} \right\}\end{aligned} \tag{45}$$

We follow the same steps as in § 2-a of Complement E_{XV}: we use the mean field approximation to replace the computation of the average value of a two-particle operator by that of average values for one-particle operators. We can, for example, use relation (43) of Complement B_{XV}, which shows that:

$$\mathrm{Tr}\left\{ a_i^\dagger a_j^\dagger a_l a_k\ \widetilde{\rho} \right\} = [\delta_{ik}\delta_{jl} + \eta\ \delta_{il}\delta_{jk}]\ \ f_\beta(\widetilde{e}_i - \mu) f_\beta(\widetilde{e}_j - \mu) \tag{46}$$

We then get:

$$\begin{aligned}&\left\langle \widehat{G} \right\rangle_{\widetilde{\rho}} = \tfrac{1}{2} \textstyle\sum_{i,j} f_\beta(\widetilde{e}_i - \mu)\ f_\beta(\widetilde{e}_j - \mu) \times \\ &\quad \Big[\langle 1:\theta_i; 2:\theta_j |\, \widehat{g}(1,2) \,|1:\theta_i; 2:\theta_j\rangle + \eta\ \langle 1:\theta_i; 2:\theta_j |\, \widehat{g}(1,2) \,|1:\theta_j; 2:\theta_i\rangle \Big]\end{aligned} \tag{47}$$

Which, according to (40), can also be written as:

$$\begin{aligned}\left\langle \widehat{G} \right\rangle_{\widetilde{\rho}} = \tfrac{1}{2} \textstyle\sum_{i,j} &\langle 1:\theta_i |\, \widetilde{\rho}_1(1) \,|1:\theta_i\rangle\ \ \langle 2:\theta_j |\, \widetilde{\rho}_1(2) \,|2:\theta_j\rangle \times \\ &\langle 1:\theta_i; 2:\theta_j |\, \widehat{g}(1,2)\, [1 + \eta P_{\mathrm{ex}}(1,2)] \,|1:\theta_i; 2:\theta_j\rangle\end{aligned} \tag{48}$$

where P_{ex} is the exchange operator between particles 1 and 2. Since:

$$\begin{aligned}&\langle 1:\theta_i |\, \widetilde{\rho}_1(1) \,|1:\theta_i\rangle\ \ \langle 2:\theta_j |\, \widetilde{\rho}_1(2) \,|2:\theta_j\rangle \\ &\qquad = \langle 1:\theta_i; 2:\theta_j |\, \widetilde{\rho}_1(1) \otimes \widetilde{\rho}_1(2) \,|1:\theta_i; 2:\theta_j\rangle\end{aligned} \tag{49}$$

and as the operators $\widetilde{\rho}_1(1)$ and $\widetilde{\rho}_1(2)$ are diagonal in the basis $|\theta_i\rangle$, we can write the right-hand side of (48) as:

$$\frac{1}{2} \sum_{i,j} \langle 1:\theta_i; 2:\theta_j |\, [\widetilde{\rho}_1(1) \otimes \widetilde{\rho}_1(2)]\, \widehat{g}(1,2)\, [1 + \eta P_{\mathrm{ex}}(1,2)] \,|1:\theta_i; 2:\theta_j\rangle \tag{50}$$

which is simply a (double) trace on two particles 1 and 2. This leads to:

$$\left\langle \widehat{G} \right\rangle_{\widetilde{\rho}} = \frac{1}{2} \mathrm{Tr}_{1,2} \left\{ [\widetilde{\rho}_1(1) \otimes \widetilde{\rho}_1(2)]\, \widehat{g}(1,2)\, [1 + \eta P_{\mathrm{ex}}(1,2)] \right\} \tag{51}$$

As announced above, the average value of the two-particle operator $\widehat{G}$ can be expressed, within the Hartree-Fock approximation, in terms of the one-particle reduced density operator $\widetilde{\rho}_1(1)$; this relation is not linear.

Comment:

The analogy with the computations of Complement $\mathrm{E_{XV}}$ becomes obvious if we regroup its equations (57) and (58) and write:

$$\left\langle \hat{W}_{\rm int} \right\rangle = \frac{1}{2} \mathrm{Tr}_{1,2} \left\{ [P_N(1) \otimes P_N(2)] \, W_2(1,2) \, [1 + \eta P_{\rm ex}(1,2)] \right\} \tag{52}$$

Replacing $W_2(1,2)$ by G, we get a relation very similar to (51), except for the fact that the projectors P_N must be replaced by the one-particle operators $\widetilde{\rho}_1$. In § 3-d, we shall come back to the correspondence between the zero and non-zero temperature results.

2-c. Variational grand potential

We now have to compute the grand potential $\widetilde{\Phi}$ written in (30). As the exponential form (28) for the trial operator makes it easy to compute $\ln\widetilde{\rho}$, we see that the terms in $\mu\widehat{N}$ cancel out, and we get:

$$\widetilde{\Phi} = \mathrm{Tr} \left\{ \left[\widehat{H} - \widetilde{H} \right] \widetilde{\rho} \right\} - k_B T \, \ln \widetilde{Z} \tag{53}$$

We now have to compute the average energy, with the density operator $\widetilde{\rho}$, of the difference between the Hamiltonians $\widehat{H}$ and $\widetilde{H}$ respectively defined by (1) and (25).

We first compute the trace:

$$\mathrm{Tr} \left\{ \widehat{H} \; \widetilde{\rho} \right\} = \left\langle \widehat{H} \right\rangle_{\widetilde{\rho}} \tag{54}$$

starting with the kinetic energy contribution $\widehat{H}_0$ in (1). We call K_0 the individual kinetic energy operator:

$$K_0 = \frac{\mathbf{P}^2}{2m} \tag{55}$$

(m is the particle mass). Equality (43) applied to $\widehat{H}_0$ yields the average kinetic energy when the system is described by $\widetilde{\rho}$:

$$\left\langle \widehat{H_0} \right\rangle_{\widetilde{\rho}} = \mathrm{Tr}_1 \left\{ K_0 \; \widetilde{\rho}_1(1) \right\} = \sum_i \langle \theta_i | \, K_0 \, | \theta_i \rangle \times f_\beta(\widetilde{e}_i - \mu) \tag{56}$$

This result is easily interpreted; each individual state contributes its average kinetic energy, multiplied by its population.

The computation of the average value $\left\langle \widehat{V_{\rm ext}} \right\rangle_{\widetilde{\rho}}$ follows the same steps:

$$\left\langle \widehat{V_{\rm ext}} \right\rangle_{\widetilde{\rho}} = \mathrm{Tr}_1 \left\{ V_1 \; \widetilde{\rho}_1(1) \right\} = \sum_i \langle \theta_i | \, V_1 \, | \theta_i \rangle \times f_\beta(\widetilde{e}_i - \mu) \tag{57}$$

(as in Complement $\mathrm{E_{XV}}$, operator V_1 is the one-particle external potential operator).

To complete the calculation of the average value of $\widehat{H}$, we now have to compute the trace $\mathrm{Tr}\{W_{\rm int}\widetilde{\rho}\}$, the average value of the interaction energy when the system is described by $\widetilde{\rho}$. Using relation (51) we can write this average value as a double trace:

$$\langle W_{\rm int} \rangle_{\widetilde{\rho}} = \frac{1}{2} \mathrm{Tr}_{1,2} \left\{ \widetilde{\rho}_1(1) \otimes \widetilde{\rho}_1(2) \;\; [W_2(1,2)] \, [1 + \eta P_{\rm ex}(1,2)] \right\} \tag{58}$$

We now turn to the average value of $\widetilde{H}$. The calculation is simplified since $\widetilde{H}$ is, like $\widehat{H_0}$, a one-particle operator; furthermore, the $|\theta_i\rangle$ have been chosen to be the eigenvectors of $\widetilde{h}$ with eigenvalues $\widetilde{e}_i$ – see relation (26). We just replace in (56), $\widehat{H_0}$ by $\widetilde{H}$, and obtain:

$$\left\langle \widetilde{H} \right\rangle_{\widetilde{\rho}} = \sum_i \langle\theta_i| \left(\widetilde{e}_i\right) |\theta_i\rangle \times f_\beta(\widetilde{e}_i - \mu) = \sum_i \widetilde{e}_i \; f_\beta(\widetilde{e}_i - \mu) \tag{59}$$

Regrouping all these results and using relation (36), we can write the variational grand potential as the sum of three terms:

$$\widetilde{\Phi} = \widetilde{\Phi}_1 + \widetilde{\Phi}_2 + \widetilde{\Phi}_3 \tag{60}$$

with:

$$\begin{aligned}
\widetilde{\Phi}_1 &= \mathrm{Tr}_1 \Big\{ \left[K_0 + V_1\right] \; \widetilde{\rho}_1(1) \Big\} \\
\widetilde{\Phi}_2 &= \frac{1}{2} \mathrm{Tr}_{1,2} \Big\{ \left[\widetilde{\rho}_1(1) \otimes \widetilde{\rho}_1(2)\right] W_2(1,2) \left[1 + \eta P_{\mathrm{ex}}(1,2)\right] \Big\} \\
\widetilde{\Phi}_3 &= \sum_i \left\{ -\widetilde{e}_i \; f_\beta(\widetilde{e}_i - \mu) + \eta k_B T \, \ln \left[1 - \eta e^{-\beta\left(\widetilde{e}_i - \mu\right)}\right] \right\}
\end{aligned} \tag{61}$$

2-d. Optimization

We now vary the eigenenergies $\widetilde{e}_k$ and eigenstates $|\theta_k\rangle$ of $\widetilde{H}$ to find the value of the density operator $\widetilde{\rho}$ that minimizes the average value $\widetilde{\Phi}$ of the potential. We start with the variations of the eigenstates, which induce no variation of $\widetilde{\Phi}_3$. The computation is actually very similar to that of Complement $\mathrm{E_{XV}}$, with the same steps: variation of the eigenvectors, followed by the demonstration that the stationarity condition is equivalent to a series of eigenvalue equations for a Hartree-Fock operator (a one-particle operator). Nevertheless, we will carry out this computation in detail, as there are some differences. In particular, and contrary to what happened in Complement $\mathrm{E_{XV}}$, the number of states $|\theta_i\rangle$ to be varied is no longer fixed by the particle number N; these states form a complete basis of the individual state space, and their number can go to infinity. This means that we can no longer give to one (or several) state(s) a variation orthogonal to all the other $|\theta_j\rangle$; this variation will necessarily be a linear combination of these states. In a second step, we shall vary the energies $\widetilde{e}_k$.

α. Variations of the eigenstates

As the eigenstates $|\theta_i\rangle$ vary, they must still obey the orthogonality relations:

$$\langle\theta_i \, |\theta_j\rangle = \delta_{ij} \tag{62}$$

The simplest idea would be to vary only one of them, $|\theta_l\rangle$ for example, and make the change:

$$|\theta_l\rangle \Rightarrow |\theta_l\rangle + |\mathrm{d}\theta_l\rangle \tag{63}$$

The orthogonality conditions would then require:

$$\langle\theta_i \, |\mathrm{d}\theta_l\rangle = 0 \qquad \text{for all } i \neq l \tag{64}$$

preventing $|\mathrm{d}\theta_l\rangle$ from having a component on any ket $|\theta_i\rangle$ other than $|\theta_l\rangle$: in other words, $|\mathrm{d}\theta_l\rangle$ and $|\theta_l\rangle$ would be colinear. As $|\theta_l\rangle$ must remain normalized, the only possible variation would thus be a phase change, which does not affect either the density operator $\widetilde{\rho}_1(1)$ or any average values computed with $\widetilde{\rho}$. This variation does not change anything and is therefore irrelevant.

It is actually more interesting to vary simultaneously two eigenvectors, which will be called $|\theta_l\rangle$ and $|\theta_m\rangle$, as it is now possible to give $|\theta_l\rangle$ a component on $|\theta_m\rangle$, and the reverse. This does not change the two-dimensional subspace spanned by these two states; hence their orthogonality with all the other basis vectors is automatically preserved. Let us give the two vectors the following infinitesimal variations (without changing their energies $\widetilde{e}_l$ and $\widetilde{e}_m$):

$$\left\{ \begin{array}{rcl} |\mathrm{d}\theta_l\rangle & = & \mathrm{d}\alpha\ e^{i\chi}\ |\theta_m\rangle \\ |\mathrm{d}\theta_m\rangle & = & -\mathrm{d}\alpha\ e^{-i\chi}\ |\theta_l\rangle \end{array} \right. \quad \text{with } l \neq m \tag{65}$$

where $\mathrm{d}\alpha$ is an infinitesimal real number and χ an arbitrary but fixed real number. For any value of χ, we can check that the variation of $\langle\theta_l\ |\theta_l\rangle$ is indeed zero (it contains the scalar products $\langle\theta_l\ |\theta_m\rangle$ or $\langle\theta_m\ |\theta_l\rangle$ which are zero), as is the symmetrical variation of $\langle\theta_m\ |\theta_m\rangle$, and that we have:

$$\mathrm{d}\,\langle\theta_l\ |\theta_m\rangle = \mathrm{d}\alpha\ e^{-i\chi}\,\langle\theta_m\ |\theta_m\rangle - \mathrm{d}\alpha\ e^{-i\chi}\,\langle\theta_l\ |\theta_l\rangle = 0 \tag{66}$$

The variations (65) are therefore acceptable, for any real value of χ.

We now compute how they change the operator $\widetilde{\rho}_1(1)$ defined in (40). In the sum over k, only the $k = l$ and $k = m$ terms will change. The $k = l$ term yields a variation:

$$\mathrm{d}\alpha\ f_\beta(\widetilde{e}_l - \mu)\left[e^{i\chi}\,|\theta_m\rangle\,\langle\theta_l| + e^{-i\chi}\,|\theta_l\rangle\,\langle\theta_m|\right] \tag{67}$$

whereas the $k = m$ term yields a similar variation but where $f_\beta(\widetilde{e}_l - \mu)$ is replaced by $-f_\beta(\widetilde{e}_m - \mu)$. This leads to:

$$\mathrm{d}\widetilde{\rho}_1(1) = \mathrm{d}\alpha\ \left[f_\beta(\widetilde{e}_l - \mu) - f_\beta(\widetilde{e}_m - \mu)\right]\left[e^{i\chi}\,|\theta_m\rangle\,\langle\theta_l| + e^{-i\chi}\,|\theta_l\rangle\,\langle\theta_m|\right] \tag{68}$$

We now include these variations in the three terms of (61); as the distributions f are unchanged, only the terms $\widetilde{\Phi}_1$ and $\widetilde{\Phi}_2$ will vary. The infinitesimal variation of $\widetilde{\Phi}_1$ is written as:

$$\mathrm{d}\widetilde{\Phi}_1 = \mathrm{Tr}_1\left\{[K_0 + V_1]\ \ \mathrm{d}\widetilde{\rho}_1(1)\right\} \tag{69}$$

As for $\mathrm{d}\widetilde{\Phi}_2$, it contains two contributions, one from $\mathrm{d}\widetilde{\rho}_1(1)$ and one from $\mathrm{d}\widetilde{\rho}_1(2)$. These two contributions are equal since the operator $W_2(1,2)$ is symmetric (particles 1 and 2 play an equivalent role). The factor $1/2$ in $\widetilde{\Phi}_2$ disappears and we get:

$$\mathrm{d}\widetilde{\Phi}_2 = \mathrm{Tr}_{1,2}\Big\{\mathrm{d}\widetilde{\rho}_1(1)\otimes\widetilde{\rho}_1(2)\ \ [W_2(1,2)]\,[1 + \eta P_{\mathrm{ex}}(1,2)]\Big\} \tag{70}$$

We can regroup these two contributions, using the fact that for any operator $O(1,2)$, it can be shown that:

$$\mathrm{Tr}_{1,2}\left\{\mathrm{d}\widetilde{\rho}_1(1)\ O(1,2)\right\} = \mathrm{Tr}_1\Big\{\mathrm{d}\widetilde{\rho}_1(1)\ \mathrm{Tr}_2\left\{O(1,2)\right\}\Big\} \tag{71}$$

This equality is simply demonstrated[5] by using the definition of the partial trace $\mathrm{Tr}_2\{O(1,2)\}$ of operator $O(1,2)$ with respect to particle 2. We then get:

$$\begin{aligned} \mathrm{d}\widetilde{\Phi} &= \mathrm{d}\widetilde{\Phi}_1 + d\widetilde{\Phi}_2 \\ &= \mathrm{Tr}_1\Big\{ \mathrm{d}\widetilde{\rho}_1(1) \Big[K_0 + V_1 + \mathrm{Tr}_2\big\{ \widetilde{\rho}_1(2)\ W_2(1,2)\ \ [1+\eta P_{\mathrm{ex}}(1,2)] \big\} \Big] \Big\} \end{aligned} \tag{72}$$

Inserting now the expression (68) for $d\widetilde{\rho}_1(1)$, we get two terms, one proportional to $e^{i\chi}$, another one to $e^{-i\chi}$, whose value is:

$$\begin{aligned} &\mathrm{d}\alpha\ e^{-i\chi}\ \ [f_\beta(\widetilde{e}_l-\mu) - f_\beta(\widetilde{e}_m-\mu)] \\ &\quad \times \mathrm{Tr}_1\Big\{ |\theta_l\rangle\langle\theta_m| \Big[K_0 + V_1 + \mathrm{Tr}_2\big\{ \widetilde{\rho}_1(2)\ W_2(1,2)\ \ [1+\eta P_{\mathrm{ex}}(1,2)] \big\} \Big] \Big\} \end{aligned} \tag{73}$$

Now, for any operator $O(1)$, we can write:

$$\begin{aligned} \mathrm{Tr}_1\{|\theta_l\rangle\langle\theta_m| O(1)\} &= \sum_i \langle\theta_i\ |\theta_l\rangle\langle\theta_m| O(1)|\theta_i\rangle = \sum_i \langle\theta_m| O(1)|\theta_i\rangle\langle\theta_i\ |\theta_l\rangle \\ &= \langle\theta_m| O(1)|\theta_l\rangle \end{aligned} \tag{74}$$

so that the variation (73) can be expressed as:

$$\begin{aligned} &\mathrm{d}\alpha\ e^{-i\chi}\ \ [f_\beta(\widetilde{e}_l-\mu) - f_\beta(\widetilde{e}_m-\mu)] \\ &\qquad \langle\theta_m| \Big[K_0 + V_1 + \mathrm{Tr}_2\big\{ \widetilde{\rho}_1(2)\ W_2(1,2)\ \ [1+\eta P_{\mathrm{ex}}(1,2)] \big\} \Big] |\theta_l\rangle \end{aligned} \tag{75}$$

The term in $e^{i\chi}$ has a similar form, but it does not have to be computed for the following reason. The variation $\mathrm{d}\widetilde{\Phi}$ is the sum of a term in $e^{i\chi}$ and another in $e^{-i\chi}$:

$$\mathrm{d}\widetilde{\Phi} = \mathrm{d}\alpha \left[c_1 e^{-i\chi} + c_2 e^{i\chi} \right] \tag{76}$$

and the stationarity condition requires $\mathrm{d}\widetilde{\Phi}$ to be zero for any choice of χ. Choosing $\chi = 0$, yields $c_1 + c_2 = 0$; choosing $\chi = \pi/2$, and multiplying by $-i$, we get $c_1 - c_2 = 0$. Adding and subtracting those two relations shows that both coefficients c_1 and c_2 must be zero. Consequently, it suffices to impose the terms in $e^{\pm i\chi}$, and hence expression (75), to be zero. When $\widetilde{e}_l \neq \widetilde{e}_m$, the distribution functions f_β are not equal, and we get:

$$\langle\theta_m| \Big[K_0 + V_1 + \mathrm{Tr}_2\big\{ \widetilde{\rho}_1(2)\ W_2(1,2)\ \ [1+\eta P_{\mathrm{ex}}(1,2)] \big\} \Big] |\theta_l\rangle = 0 \tag{77}$$

(if $\widetilde{e}_l = \widetilde{e}_m$, however, we have not yet obtained any particular condition to be satisfied[6]).

[5] The definition of partial traces is given in § 5-b of Complement $\mathrm{E}_{\mathrm{III}}$. The left hand side of (71) can be written as $\sum_{i,j} \langle 1:\theta_i; 2:\theta_j| d\rho_I(1) O(1,2) |1:\theta_i; 2:\theta_j\rangle$. We then insert, after $d\rho_I(1)$, a closure relation on the kets $|1:\theta_k; 2:\theta_{k'}\rangle$, with $k' = j$ since $d\rho_I(1)$ does not act on particle 2. This yields: $\sum_{i,k} \langle 1:\theta_i| d\rho_I(1) |1:\theta_k\rangle \sum_j \langle 1:\theta_k; 2:\theta_j| O(1,2) |1:\theta_i; 2:\theta_j\rangle$, where the sum over j is the definition of the matrix element between $\langle 1:\theta_k|$ and $|1:\theta_i\rangle$ of the partial trace over particle 2 of the operator $O(1,2)$. We then get the right-hand side of (71).

[6] This was expected, since this choice does not lead to any variation of the trial density operator.

β. Variation of the energies

Let us now see what happens if the energy $\widetilde{e}_l$ varies by $\mathrm{d}\widetilde{e}_l$. The function $f_\beta(\widetilde{e}_l-\mu)$ then varies by $\mathrm{d}f^l_\beta$ which, according to relation (40), induces a variation of $\widetilde{\rho}_1$:

$$\mathrm{d}\widetilde{\rho}_1 = |\theta_l\rangle\,\langle\theta_l|\,\mathrm{d}f^l_\beta \tag{78}$$

and thus leads to variations of expressions (61) of $\widetilde{\Phi}_1$ and $\widetilde{\Phi}_2$. Their sum is:

$$\mathrm{d}\widetilde{\Phi}_1 + \mathrm{d}\widetilde{\Phi}_2 = \mathrm{Tr}_1\Big\{\mathrm{d}\widetilde{\rho}_1(1)\Big[K_0 + V_1 + \mathrm{Tr}_2\big\{\widetilde{\rho}_1(2)\ \ [W_2(1,2)]\,[1+\eta P_{\mathrm{ex}}(1,2)]\ \big\}\Big]\Big\} \tag{79}$$

where the factor $1/2$ in $\widetilde{\Phi}_2$ has been canceled since the variations induced by $\widetilde{\rho}_1(1)$ and $\widetilde{\rho}_1(2)$ double each other. Inserting (78) in this relation and using again (74), we get:

$$\mathrm{d}\widetilde{\Phi}_1 + \mathrm{d}\widetilde{\Phi}_2 = \mathrm{d}f^l_\beta\,\langle\theta_l|\,\Big[K_0 + V_1 + \mathrm{Tr}_2\big\{\widetilde{\rho}_1(2)\ \ [W_2(1,2)]\,[1+\eta P_{\mathrm{ex}}(1,2)]\ \big\}\Big]\,|\theta_l\rangle \tag{80}$$

As for $\widetilde{\Phi}_3$, its variation is the sum of a term in $\mathrm{d}\widetilde{e}_l$ coming from the explicit presence of the energies $\widetilde{e}_i$ in its definition (61), and a term in $\mathrm{d}f^l_\beta$. If we let only the energy $\widetilde{e}_l$ vary (not taking into account the variations of the distribution function), we get a zero result, since:

$$\begin{aligned}\Bigg[-f_\beta(\widetilde{e}_l-\mu) + \eta k_B T\ \ \Big[(-\eta)\,(-\beta)\ \ e^{-\beta\left(\widetilde{e}_l-\mu\right)}\Big]\,\frac{1}{1-\eta e^{-\beta\left(\widetilde{e}_l-\mu\right)}}\Bigg]\,\mathrm{d}\widetilde{e}_l \\ = \big[-f_\beta(\widetilde{e}_l-\mu) + f_\beta(\widetilde{e}_l-\mu)\big]\mathrm{d}\widetilde{e}_l = 0\end{aligned} \tag{81}$$

Consequently, we just have to vary by $\mathrm{d}f^l_\beta$ the distribution function, and we get:

$$\mathrm{d}\widetilde{\Phi}_3 = -\widetilde{e}_l\ \mathrm{d}f^l_\beta \tag{82}$$

Finally, after simplification by $\mathrm{d}f^l_\beta$ (which, by hypothesis, is different from zero), imposing the variation $\mathrm{d}\widetilde{\Phi}$ to be zero leads to the condition:

$$\begin{aligned}&\mathrm{d}\widetilde{\Phi}_1 + \mathrm{d}\widetilde{\Phi}_2 + \mathrm{d}\widetilde{\Phi}_3 \\ &\quad = \langle\theta_l|\,[K_0 + V_1 - \widetilde{e}_l + \mathrm{Tr}_2\,\{\widetilde{\rho}_1(2)\ \ [W_2(1,2)]\,[1+\eta P_{\mathrm{ex}}(1,2)]\}]\,|\theta_l\rangle = 0\end{aligned} \tag{83}$$

This expression does look like the stationarity condition at constant energy (77), but now the subscripts l and m are the same, and a term in $-\widetilde{e}_l$ is present in the operator.

3. Temperature dependent mean field equations

Introducing a Hartree-Fock operator acting in the single particle state space allows writing the stationarity relations just obtained in a more concise and manageable form, as we now show.

3-a. Form of the equations

Let us define a temperature dependent Hartree-Fock operator as the partial trace that appears in the previous equations:

$$\boxed{W_{HF}(\beta) = \mathrm{Tr}_2\left\{\widetilde{\rho}_1(2) W_2(1,2)\left[1+\eta P_{\mathrm{ex}}(1,2)\right]\right\}} \tag{84}$$

It is thus an operator acting on the single particle 1. It can be defined just as well by its matrix elements between the individual states:

$$\begin{aligned}&\langle\theta_k|W_{HF}(\beta)\,|\theta_l\rangle \\ &\quad = \sum_j f_\beta(\widetilde{e}_j-\mu)\,\langle 1:\theta_k; 2:\theta_j|\,W_2(1,2)\left[1+\eta P_{\mathrm{ex}}(1,2)\right]|1:\theta_l; 2:\theta_j\rangle\end{aligned} \tag{85}$$

Equation (77) is valid for any two chosen values l and m, as long as $\widetilde{e}_l \neq \widetilde{e}_m$. When l is fixed and m varies, it simply means that the ket:

$$\left[K_0+V_1+W_{HF}(\beta)\right]|\theta_l\rangle \tag{86}$$

is orthogonal to all the eigenvectors $|\theta_m\rangle$ having an eigenvalue $\widetilde{e}_m$ different from $\widetilde{e}_l$; it has a zero component on each of these vectors. As for equation (83), it yields the component of this ket on $|\theta_l\rangle$, which is equal to $\widetilde{e}_l\,|\theta_l\rangle$. The set of $|\theta_m\rangle$ (including those having the same eigenvalue as $|\theta_l\rangle$) form a basis of the individual state space, defined by (26) as the basis of eigenvectors of the individual operator $\widetilde{h}$. Two cases must be distinguished:

(i) If $\widetilde{e}_l$ is a non-degenerate eigenvalue of $\widetilde{h}$, the set of equations (77) and (83) determine all the components of the ket $\left[K_0+V_1+W_{HF}(\beta)\right]|\theta_l\rangle$. This shows that $|\theta_l\rangle$ is an eigenvector of the operator $K_0+V_1+W_{HF}$ with the eigenvalue $\widetilde{e}_l$.

(ii) If this eigenvalue of $\widetilde{h}$ is degenerate, relation (77) only proves that the eigen-subspace of $\widetilde{h}$, with eigenvalue $\widetilde{e}_l$, is stable under the action of the operator $K_0+V_1+W_{HF}$; it does not yield any information on the components of the ket (86) inside that subspace. It is possible though to diagonalize $K_0+V_1+W_{HF}$ inside each of the eigen-subspace of $\widetilde{h}$, which leads to a new eigenvectors basis $|\varphi_n\rangle$, now common to $\widetilde{h}$ and $K_0+V_1+W_{HF}$.

We now reason in this new basis where all the $\left[K_0+V_1+W_{HF}(\beta)\right]|\varphi_n\rangle$ are proportional to $|\varphi_n\rangle$. Taking (83) into account, we get:

$$\boxed{\left[K_0+V_1+W_{HF}(\beta)\right]|\varphi_n\rangle = \widetilde{e}_n\,|\varphi_n\rangle} \tag{87}$$

As we just saw, the basis change from the $|\theta_l\rangle$ to the $|\varphi_n\rangle$ only occurs within the eigen-subspaces of $\widetilde{h}$ corresponding to given eigenvalues $\widetilde{e}$; one can then replace the $|\theta_l\rangle$ by the $|\varphi_n\rangle$ in the definition (40) of $\widetilde{\rho}_1(1)$ and write:

$$\boxed{\widetilde{\rho}_1(1) = \sum_n f_\beta(\widetilde{e}_n-\mu)\;\;|1:\varphi_n\rangle\,\langle 1:\varphi_n|} \tag{88}$$

Inserting this relation in the definition (84) of $W_{HF}(\beta)$ leads to a set of equations only involving the eigenvectors $|\varphi_n\rangle$.

For all the values of n we get a set of equations (87), which, associated with (84) and (88) defining the potential $W_{HF}(\beta)$ as a function of the $|\varphi_n\rangle$, are called the temperature dependent Hartree-Fock equations.

3-b. Properties and limits of the equations

We now discuss how to apply the mean field equations we have obtained, and their limit of validity, which are more stringent for bosons than for fermions.

α. *Using the equations*

Hartree-Fock equations concern a self-consistent and nonlinear system: the eigenvectors $|\varphi_n\rangle$ and eigenvalues of the density operator $\widetilde{\rho}_1(1)$ are solutions of an eigenvalue equation (87) which itself depends on $\widetilde{\rho}_1(1)$. This situation is reminiscent of the one encountered with the zero-temperature Hartree-Fock equations, and, a priori, no exact solutions can be found.

As for the zero-temperature case, we proceed by iteration: starting from a physically reasonable density operator $\widetilde{\rho}_1(1)$, we use it in (84) to compute a first value of the Hartree-Fock potential operator. We then diagonalize this operator to get its eigenkets and eigenvalues $\bar{e}_n$. Next, we build the operator $\overline{\rho}_1$ that has the same eigenkets, but whose eigenvalues are the $f_\beta(\bar{e}_n - \mu)$. Inserting this new operator $\overline{\rho}_1$ in (84), we get a second iteration of the Hartree-Fock operator. We again diagonalize this operator to compute new eigenvalues and eigenvectors, on which we build the next approximation $\overline{\overline{\rho}}_1$ of $\widetilde{\rho}_1$, and so on. After a few iterations, we may expect convergence towards a self-consistent solution.

β. *Validity limit*

For a fermion system, there is no fundamental general limit for using the Hartree-Fock approximation. The pertinence of the final result obviously depends on the nature of the interactions, and whether a mean field treatment of these interactions is a good approximation. One can easily understand that the larger the interaction range, the more each particle will be submitted to the action of many others. This will lead to an averaging effect improving the mean field approximation. If, on the other hand, each particle only interacts with a single partner, strong binary correlations may appear, which cannot be correctly treated by a mean field acting on independent particles.

For bosons, the same general remarks apply, but the populations are no longer limited to 1. When, for example, Bose-Einstein condensation occurs, one population becomes much larger than the others, and presents a singularity that is not accounted for in the calculations presented above. The Hartree-Fock approximation has therefore more severe limitations than for the fermions, and we now discuss this problem.

For a boson system in which many individual states have comparable populations, taking into account the interactions by the Hartree-Fock mean field yields as good an approximation as for a fermion system. If the system however is close to condensation, or already condensed, the mean field equations we have written are no longer valid. This is because the trial density operator in relation (31) contains a distribution function associated with each individual quantum state and varies as for an ideal gas, i.e. as an exponentially decreasing function of the occupation numbers. Now we saw in § 3-b-β of Complement $\mathrm{B_{XV}}$ that, in an ideal gas, the fluctuations of the particle numbers in each of the individual states are as large as the average values of those particle numbers. If the individual state has a large population, these fluctuations can become very important, which is physically impossible in the presence of repulsive interactions. Any population fluctuation increases the average value of the square of the occupation number (equal

to the sum of the squared average value and the squared fluctuation), and hence of the interaction energy (proportional to the average value of the square). A large fluctuation in the populations would lead to an important increase of the interaction repulsive energy, in contradiction with the minimization of the thermodynamic potential. In other words, the finite compressibility of the physical system, introduced by the interactions, prevents any large fluctuation in the density. Consequently, the fluctuations in the number of condensed particles predicted by the trial Hartree-Fock density operator are not physically acceptable, in the presence of condensation.

It is worth analyzing more precisely the origin of this Hartree-Fock approximation limit, in terms of correlations between the particles. Relation (51) concerns any two-particle operator $\widehat{G}$. It shows that, using the trial density operator (31), the two-particle reduced density operator can be written as:

$$\widetilde{\rho}_2(1,2) = \widetilde{\rho}_1(1) \otimes \widetilde{\rho}_1(1)\,[1 + P_{\rm ex}(1,2)] \tag{89}$$

Its diagonal matrix elements are then written:

$$\begin{aligned} &\langle 1:\theta_i; 2:\theta_j|\,\widetilde{\rho}_2(1,2)\,|1:\theta_i; 2:\theta_j\rangle \\ &= \langle 1:\theta_i|\,\widetilde{\rho}_1\,|1:\theta_i\rangle\,\langle 2:\theta_j|\,\widetilde{\rho}_1\,|2:\theta_j\rangle + \langle 1:\theta_i|\,\widetilde{\rho}_1\,|1:\theta_j\rangle\,\langle 2:\theta_j|\,\widetilde{\rho}_1\,|2:\theta_i\rangle \end{aligned} \tag{90}$$

and are the sum of a direct term, and an exchange term. When $i \neq j$, the presence of an exchange term is not surprising, and corresponds to the general discussion of § C-5 in Chapter XV. It is similar to the expression of the spatial correlation function written in (C-34) of that chapter, which is also the sum of two contributions, a direct one (C-32) and an exchange one (C-33). Since this last contribution is positive when $r_1 \simeq r_2$, the physical consequence of the exchange is a spatial bunching of the bosons. What is surprising though is that the exchange term still exists in (90) when $i = j$, even though the notion of exchange is meaningless: when dealing with a single individual state, the four expressions (C-21) of Chapter XV reduce to a single one, the direct term. We can also check that the exchange term (C-34) of Chapter XV includes the explicit condition $i \neq j$, which means it receives no contribution from $i = j$. We shall furthermore confirm in § 3 of Complement A_{XVI} that bosons all placed in the same individual quantum state are not spatially correlated, and therefore present neither bunching nor exchange effects. The mathematical expression of the trial two-particle Hartree-Fock density operator thus contains too many exchange terms. This does not really matter as long as the boson system remains far from Bose-Einstein condensation: the error involved is small since the $i = j$ terms play a negligible role compared to the $i \neq j$ terms in the summations over i and j appearing in the interaction energy. However, as soon as an individual state becomes highly populated, significant errors can occur and the Hartree-Fock method must be abandoned. There exist, however, more elaborate theoretical treatments better adapted to this case.

3-c. Differences with the zero-temperature Hartree-Fock equations (fermions)

The main difference between the approach we just used and that of Complements C_{XV} and E_{XV} is that these complements were only looking for a single eigenstate of the Hamiltonian $\hat{H}$, generally its ground state. If we are now interested in several of these states, we have to redo the computation separately for each of them. To study the

properties of thermal equilibrium, one could imagine doing the calculations a great many times, and then weigh the results with occupation probabilities. This method obviously leads to heavy computations, which become impossible for a macroscopic system having an extremely large number of levels. In the present complement, the Hartree-Fock equations yield immediately thermal averages, as well as eigenvectors of a one-particle density operator with their energies.

Another important difference is that the Hartree-Fock operator now depends on the temperature, because of the presence in (85) of a temperature dependent distribution function – or, which amounts to the same thing, of the presence in (84) of an operator dependent on β, and which replaces the projector $P_N(2)$ onto all the populated individual states. The equations obtained remind us of those governing independent particles, each finding its thermodynamic equilibrium while moving in the self-consistent mean field created by all the others, also including the exchange contribution (which can be ignored in the simplified "Hartree" version).

We must keep in mind, however, that the Hartree-Fock potential associated with each individual state now depends on the populations of an infinity of other individual states, and these populations are function of their energy as well as of the temperature. In other words, because of the nonlinear character of the Hartree-Fock equations, the computation is not merely a juxtaposition of separate mean field calculations for stationary individual states.

3-d. Zero-temperature limit (fermions)

Let us check that the Hartree-Fock method for non-zero temperature yields the same results as the zero temperature method explained in Complement $\mathrm{E_{XV}}$ for fermions.

In § 2-d of Complement $\mathrm{B_{XV}}$, we introduced for an ideal gas the concept of a degenerate quantum gas. It can be generalized to a gas with interactions: in a fermion system, when $\beta\mu \gg 1$, the system is said to be strongly degenerate. As the temperature goes to zero, a fermion system becomes more and more degenerate. Can we be certain that the results of this complement are in agreement with those of Complement $\mathrm{E_{XV}}$, valid at zero temperature?

We saw that the temperature comes into play in the definition (85) of the mean Hartree-Fock potential, W_{HF}. In the limit of a very strong degeneracy, the Fermi-Dirac distribution function appearing in the definition (40) of $\widetilde{\rho}_1(1)$ becomes practically a step function, equal to 1 for energies e_j less than the chemical potential μ, and zero otherwise (Figure 1 of Complement $\mathrm{B_{XV}}$. In other words, the only populated states (and by a single fermion) are the states having energies less than μ, i.e. less than the Fermi level. Under such conditions, the $\widetilde{\rho}_1(2)$ of (84) becomes practically equal to the projector $P_N(2)$ which, in Complement $\mathrm{E_{XV}}$, appears in the definition (52) of the zero-temperature Hartree-Fock potential; in other words, the partial trace appearing in this relation (85) is then strictly limited to the individual states having the lowest energies. We thus obtain the same Hartree Fock equations as for zero temperature, leading to the determination of a set of individual eigenstates on which we can build a unique N-particle state.

3-e. Wave function equations

Let us write the Hartree-Fock equations (87) in terms of wave functions: these equations are strictly equivalent to (87), written in terms of operators and kets, but their

form is sometimes easier to use, in particular for numerical calculations.

Assuming the particles have a spin, we shall note the wave functions $\varphi^{\nu}(\mathbf{r})$, with:

$$\varphi^{\nu}(\mathbf{r}) = \langle \mathbf{r}, \nu \, | \varphi \rangle \tag{91}$$

where the spin quantum number ν can take $(2S+1)$ values; according to the nature of the particles, the possible spins S are $S = 0$, $S = 1/2$, $S = 1$ etc. As in Complement E_{XV} (§ 2-d), we introduce a complete basis $\{|\varphi_k^{\nu_k}\rangle\}$ for the individual state space, built from kets that are all eigenvectors of the spin component along the quantization axis, with eigenvalue ν_k. For each value of n, the spin index ν takes on a given value ν_n and is not, therefore, an independent index. As for the potentials, we assume here again that V_1 is diagonal in ν, but that its diagonal elements $V_1^{\nu}(\mathbf{r})$ may depend on ν. The interaction potential, however, is described by a function $W_2(\mathbf{r}, \mathbf{r}')$ that only depends on $\mathbf{r} - \mathbf{r}'$, but does not act on the spins.

To obtain the matrix elements of $\widehat{W}_{HF}(\beta)$ in the representation $\{|\mathbf{r}, \nu\rangle\}$, we use (85) after replacing the $|\theta\rangle$ by the $|\varphi\rangle$(we showed in § 3 that this was possible). We now multiply both sides by $\langle \mathbf{r}, \nu \, | \varphi_k^{\nu_k} \rangle$ and $\langle \varphi_l^{\nu_l} \, | \mathbf{r}', \nu \rangle$, and sum over the subscripts k and l; we recognize in both sides the closure relations:

$$\sum_k \langle \mathbf{r}, \nu \, | \varphi_k^{\nu_k} \rangle \langle \varphi_k^{\nu_k} | = \langle \mathbf{r}, \nu | \quad \text{and} \quad \sum_l | \varphi_l^{\nu_l} \rangle \langle \varphi_l^{\nu_l} \, | \mathbf{r}', \nu' \rangle = | \mathbf{r}', \nu' \rangle \tag{92}$$

This leads to:

$$\begin{aligned} &\langle \mathbf{r}, \nu | \, \widehat{W}_{HF}(\beta) \, | \mathbf{r}', \nu' \rangle = \\ &\qquad \sum_j f_\beta(\widetilde{e}_j - \mu) \; \left\langle 1 : \mathbf{r}, \nu; 2 : \varphi_j^{\nu_j} \right| W_2(1,2) \left[1 + \eta P_{ex}(1,2) \right] \left| 1 : \mathbf{r}', \nu'; 2 : \varphi_j^{\nu_j} \right\rangle \end{aligned} \tag{93}$$

As in § C-5 of Chapter XV, we get the sum of a direct term (the term 1 in the central bracket) and an exchange term (the term in ηP_{ex}). This expression contains the same matrix element as relation (87) of Complement E_{XV}, the only difference being the presence of a coefficient $f_\beta(\widetilde{e}_j - \mu)$ in each term of the sum (plus the fact that the summation index goes to infinity).

(i) For the direct term, as we did in that complement, we insert a closure relation on the particle 2 position:

$$\left| 2 : \varphi_j^{\nu_j} \right\rangle = \int d^3 r_2 \; \varphi_j^{\nu_j}(\mathbf{r}_2) \, | 2 : \mathbf{r}_2, \nu_j \rangle \tag{94}$$

Since the interaction operator is diagonal in the position representation, the part of the matrix element of (93) that does not contain the exchange operator becomes:

$$\int d^3 r_2 \; \left| \varphi_j^{\nu_j}(\mathbf{r}_2) \right|^2 \; \langle 1 : \mathbf{r}, \nu; 2 : \mathbf{r}_2 | \, W_2(1,2) \, | 1 : \mathbf{r}', \nu'; 2 : \mathbf{r}_2 \rangle \tag{95}$$

The direct term of (93) is then written:

$$\delta_{\nu\nu'} \; \delta(\mathbf{r} - \mathbf{r}') \int \mathrm{d}^3 r_2 \; W_2(\mathbf{r}, \mathbf{r}_2) \sum_j f_\beta(\widetilde{e}_j - \mu) \left| \varphi_j^{\nu_j}(\mathbf{r}_2) \right|^2 \tag{96}$$

which is equivalent to relation (91) of Complement E_{XV}.

(ii) The exchange term is obtained by permutation of the two particles in the ket appearing on the right-hand side of (93); the diagonal character of $W_2(1,2)$ in the position representation leads to the expression:

$$\left\langle 1:\mathbf{r},\nu \left| 1:\varphi_j^{\nu_j} \right\rangle \left\langle 2:\varphi_j^{\nu_j} \right| 2:\mathbf{r}',\nu' \right\rangle W_2(\mathbf{r},\mathbf{r}') \tag{97}$$

For the first scalar product to be non-zero, the subscript j must be such that $\nu_j = \nu$; in the same way, for the second product to be non-zero, we must have $\nu_j = \nu'$. For both conditions to be satisfied, we must impose $\nu = \nu'$, and the exchange term (93) is equal to:

$$\eta\ \delta_{\nu\nu'} \sum_{\nu_j=\nu} f_\beta(\widetilde{e}_j - \mu)\ \left[\varphi_j^{\nu_j}(\mathbf{r})\right] \left[\varphi_j^{\nu_j}(\mathbf{r}')\right]^* W_2(\mathbf{r},\mathbf{r}') \tag{98}$$

where the summation is on all the values of j such that $\nu_j = \nu$: this term only exists if the two interacting particles are totally indistinguishable, which requires that they be in the same spin state (see the discussion in Complement $\mathrm{E_{XV}}$).

We now define the direct and exchange potentials by:

$$\begin{aligned} V_{\mathrm{dir}}(\mathbf{r}) &= \int d^3r''\ W_2(\mathbf{r},\mathbf{r}'') \sum_j f_\beta(\widetilde{e}_j - \mu) \left|\varphi_j^{\nu_j}(\mathbf{r}'')\right|^2 \\ V_{\mathrm{ex}}^{\nu}(\mathbf{r},\mathbf{r}') &= W_2(\mathbf{r},\mathbf{r}') \sum_{\nu_j=\nu} f_\beta(\widetilde{e}_j - \mu)\ \left[\varphi_j^{\nu_j}(\mathbf{r})\right] \left[\varphi_j^{\nu_j}(\mathbf{r}')\right]^* \end{aligned} \tag{99}$$

The equalities (87) then lead to the Hartree-Fock equations in the position representation:

$$\boxed{\left[-\frac{\hbar^2}{2m}\Delta + V_1^{\nu_n}(\mathbf{r}) + V_{\mathrm{dir}}(\mathbf{r})\right] \varphi_n^{\nu_n}(\mathbf{r}) + \eta \int \mathrm{d}^3r'\ V_{\mathrm{ex}}^{\nu_n}(\mathbf{r},\mathbf{r}')\varphi_n^{\nu_n}(\mathbf{r}') = \widetilde{e}_n\ \varphi_n^{\nu_n}(\mathbf{r})} \tag{100}$$

The general discussion of § 3-b can be applied here without any changes. These equations are both nonlinear and self-consistent, as the direct and exchange potentials are themselves functions of the solutions $\varphi_j^{\nu}(\mathbf{r})$ of the eigenvalue equations (100). This situation is reminiscent of the zero-temperature case, and we can, once again, look for solutions using iterative methods. The number of equations to be solved, however, is infinite and no longer equal to the finite number N, as already pointed out in § 3-c. The set of solutions must span the entire individual state space. Along the same line, in the definitions (99) of the direct and exchange potentials, the summations over j are not limited to N states, but go to infinity. However, even though the number of these wave functions is in principle infinite, it is limited in practice (for numerical calculations) to a high but finite number. As for the initial conditions to start the iteration process, one can choose for example the states and energies of a free fermion gas, but any other conjecture is equally possible.

Conclusion

There are many applications of the previous calculations, and more generally of the mean field theory. We give a few examples in the next complement, which are far from showing the richness of the possible application range. The main physical idea is to

reduce, whenever possible, the calculation of the various physical quantities to a problem similar to that of an ideal gas, where the particles have independent dynamics. We have indeed shown that the individual level populations, as well as the total particle number, are given by the same distribution functions f_β as for an ideal gas – see relations (38) and (44). The same goes for the system entropy S, as already mentioned at the end of § 2-b-α. If we replace the free particle energies by the modified energies $\widetilde{e}_{\mathbf{k},\nu}$, the analogy with independent particles is quite strong.

If we now want to compute other thermodynamic quantities, as for example the average energy, we can no longer use the ideal gas formulas; we must go back to the equations of § 2-c. The grand potential may be calculated by inserting in (61) the $|\theta_i\rangle$ and the $\widetilde{e}_i$ obtained from the Hartree-Fock equations. Another method uses the fact that $\left\langle \widehat{N} \right\rangle$ is given by ideal gas formulas that contain the distribution f_β, and hence do not require any further calculations. As:

$$\left\langle \widehat{N} \right\rangle = \beta^{-1} \frac{\partial}{\partial \mu} \ln Z \tag{101}$$

we can integrate $\left\langle \widehat{N} \right\rangle$ over μ (between $-\infty$ and the current value μ, for a fixed value of β) to obtain $\ln Z$, and hence the grand potential. From this grand potential, all the other thermodynamic quantities can be calculated, taking the proper derivatives (for example a derivative with respect to β to get the average energy). We shall see an example of this method in § 4-a of the next complement.

We must however keep in mind that all these calculations derive from the mean field approximation, in which we replaced the exact equilibrium density operator by an operator of the form (32). In many cases this approximation is good, even excellent, as is the case, in particular, for a long-range interaction potential: each particle will interact with several others, therefore enhancing the averaging effect of the interaction potential. It remains, however, an approximation: if, for example, the particles interact via a "hard core" potential (infinite potential when the mutual distance becomes less than a certain microscopic distance), the particles, in the real world, can never be found at a distance from each other smaller than the hard core diameter; now this impossibility is not taken into account in (32). Consequently, there is no guarantee of the quality of a mean field approximation in all situations, and there are cases for which it is not sufficient.

Complement H_{XV}

Applications of the mean field method for non-zero temperature (fermions and bosons)

1 Hartree-Fock for non-zero temperature, a brief review . . . 1733
2 Homogeneous system . . . 1734
2-a Calculation of the energies . . . 1735
2-b Quasi-particules . . . 1736
3 Spontaneous magnetism of repulsive fermions . . . 1737
3-a A simple model . . . 1737
3-b Resolution of the equations by graphical iteration . . . 1739
3-c Physical discussion . . . 1743
4 Bosons: equation of state, attractive instability . . . 1745
4-a Repulsive bosons . . . 1746
4-b Attractive bosons . . . 1747

In the previous complement, we presented the Hartree-Fock (mean field) method for non-zero temperatures, which has numerous applications – a few of them will be discussed in this complement. We start in § 1 with a brief review of the results obtained with this method in the previous complement, and which will be used in the present complement. The general properties of a homogeneous system are then studied in § 2, as this particular case is often encountered, hence giving it a special importance. The last two sections are concerned with the study of phase transitions in homogeneous systems. Section § 3 studies fermions; we show how the mean field theory predicts the existence of a transition where a fermion system becomes spontaneously magnetic because of the repulsion between particles (even though this repulsion is supposed to be completely independent of the spins). Finally, the last section deals with bosons and the study of their equation of state. This will allow us to show, in particular, the appearance of an instability when the bosons are attractive and close to Bose-Einstein condensation.

1. Hartree-Fock for non-zero temperature, a brief review

We start with a brief review of the results obtained previously (§ 2 of Complement B_{XV} and § 3 of Complement G_{XV}, which will be useful for what follows.

For an ideal gas, the distribution function f_β^{FD} for fermions, or f_β^{BE} for bosons, is given by:

$$f_\beta(e-\mu) = \begin{cases} f_\beta^{FD}(e-\mu) = \dfrac{1}{e^{\beta(e-\mu)}+1} & \text{for fermions} \\ f_\beta^{BE}(e-\mu) = \dfrac{1}{e^{\beta(e-\mu)}-1} & \text{for bosons} \end{cases} \tag{1}$$

where $\beta = 1/k_B T$ and μ is the chemical potential. The average total particle number $\left\langle \widehat{N} \right\rangle$ is then obtained by a sum over all the individual accessible states, labeled by the

subscript i:

$$\left\langle \widehat{N} \right\rangle = \sum_{i=1}^{\infty} f_\beta(e_i - \mu) \tag{2}$$

The temperature dependent Hartree-Fock equations (mean field equations) in the position representation are given by relation (100) of Complement G_{XV}:

$$\left[-\frac{\hbar^2}{2m}\Delta + V_1^{\nu_n}(\mathbf{r}) + V_{\rm dir}(\mathbf{r}) \right] \varphi_n^{\nu_n}(\mathbf{r}) + \eta \int \mathrm{d}^3 r' \; V_{\rm ex}^{\nu_n}(\mathbf{r},\mathbf{r}') \varphi_n^{\nu_n}(\mathbf{r}') = \widetilde{e}_n \; \varphi_n^{\nu_n}(\mathbf{r}) \tag{3}$$

where $\eta = +1$ for bosons, $\eta = -1$ for fermions, and where $V_{\rm dir}(\mathbf{r})$ and $V_{\rm ex}^{\nu}(\mathbf{r})$ are given by relation (99) of that same complement (we assume the interaction potential does not act on the spin quantum numbers ν):

$$\begin{aligned} V_{\rm dir}(\mathbf{r}) &= \int d^3 r'' \; W_2(\mathbf{r},\mathbf{r}'') \sum_j f_\beta(\widetilde{e}_j - \mu) \left| \varphi_j^{\nu_j}(\mathbf{r}'') \right|^2 \\ V_{\rm ex}^{\nu}(\mathbf{r},\mathbf{r}') &= W_2(\mathbf{r},\mathbf{r}') \sum_{\nu_j = \nu} f_\beta(\widetilde{e}_j - \mu) \; \left[\varphi_j^{\nu_j}(\mathbf{r}) \right] \left[\varphi_j^{\nu_j}(\mathbf{r}') \right]^* \end{aligned} \tag{4}$$

2. Homogeneous system

We assume from now on that the physical system is subjected to boundary conditions created by a one-body potential, which confines the particles inside a cubic box of edge length L; this potential is zero ($\widehat{V}_1 = 0$) inside the box, and takes on an infinite value outside. To take this confinement into account, we shall use the periodic boundary conditions (Complement C_{XIV}, § 1-c), for which the normalized eigenfunctions of the kinetic energy are written as:

$$\frac{1}{L^{3/2}} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{5}$$

where the possible wave vectors $\mathbf{k}$ are those whose three components are integer multiples of $2\pi/L$. Because of the spin, the eigenvectors of the kinetic energy are labeled by the values of both $\mathbf{k}$ and ν, and are written $|\mathbf{k},\nu\rangle$, with:

$$\langle \mathbf{r}, \nu \, | \mathbf{k}, \nu \rangle = \varphi_{\mathbf{k}}^{\nu}(\mathbf{r}) = \frac{1}{L^{3/2}} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{6}$$

The index n (or j) that labeled the basis vectors in the previous complements is now replaced by two indices, $\mathbf{k}$ and ν (which are independent, as opposed to the indices n and ν_n used in § 3-e of Complement G_{XV}). We shall finally assume that the particle interaction is invariant under translation: $W_2(\mathbf{r}_1,\mathbf{r}_2)$ only depends on $\mathbf{r}_1 - \mathbf{r}_2$.

We are going to see that, in such a case, solutions of the Hartree-Fock equations can be found without having to search for the eigenfunctions of the (Hartree-Fock) operator written on the left-hand side of (3); these solutions are simply the plane waves written in (5). Only the operator's eigenvalues $\widetilde{e}_{\mathbf{k},\nu}$ remain to be calculated, and can be interpreted as the energies of independent objects called "quasi-particles" (§ 2-b).

Comment:

We shall verify that these plane waves are solutions of the Hartree-Fock equations, while neither being necessarily the only ones, nor even those leading to the lowest energy of the total system. A phenomenon called "symmetry breaking" (translation symmetry in this case) could occur and introduce solutions whose moduli vary in space and correspond to lower energies. The Wigner crystal of electrons is such an example, where the particle density spontaneously shows a periodic spatial modulation. Another example of spontaneous symmetry breaking will be discussed in § 3-c of this complement. There are many other cases (in nuclear physics in particular) where the Hartree-Fock method can be used to study symmetry breaking phenomena.

2-a. Calculation of the energies

As the plane waves are obviously eigenfunctions of the kinetic energy, and since the potential is zero inside the box, we just have to demonstrate that they are also eigenfunctions of the direct and exchange potentials. Inserting (5) in (4), we get:

$$\begin{aligned} V_{\text{dir}}(\mathbf{r}) &= \frac{1}{L^3} \sum_{\mathbf{k}',\nu'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu'} - \mu) \int \mathrm{d}^3 r'' \;\; W_2(\mathbf{r} - \mathbf{r}'') \\ &= \frac{\overline{V}_0}{L^3} \sum_{\mathbf{k}',\nu'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu'} - \mu) \end{aligned} \tag{7}$$

where $\overline{V}_0$ is defined as (with a change of variable $\mathbf{r} - \mathbf{r}'' = \mathbf{s}$):

$$\overline{V}_0 = \int \mathrm{d}^3 r'' \;\; W_2(\mathbf{r} - \mathbf{r}'') = \int \mathrm{d}^3 s \; W_2(\mathbf{s}) \tag{8}$$

The direct potential is therefore a constant, independent of the position $\mathbf{r}$; multiplying an exponential $e^{i\mathbf{k}\cdot\mathbf{r}}$, it yields a function proportional to it. This means that $e^{i\mathbf{k}\cdot\mathbf{r}}$ is an eigenfunction of the direct potential. As for the exchange potential, using the second relation of (4), we get:

$$V_{\text{ex}}^{\nu}(\mathbf{r}, \mathbf{r}') = \frac{1}{L^3} \sum_{\mathbf{k}'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu} - \mu) \; e^{i\mathbf{k}'\cdot(\mathbf{r}-\mathbf{r}')} \; W_2(\mathbf{r} - \mathbf{r}') \tag{9}$$

The exchange potential is thus also translation-invariant (it only depends on the difference $\mathbf{r} - \mathbf{r}'$). Consequently, the last term on the left-hand side of (3) can be written as:

$$\begin{aligned} &\frac{\eta}{L^3} \sum_{\mathbf{k}'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu} - \mu) \int \mathrm{d}^3 r' e^{i\mathbf{k}'\cdot(\mathbf{r}-\mathbf{r}')} \; W_2(\mathbf{r} - \mathbf{r}') \frac{e^{i\mathbf{k}\cdot\mathbf{r}'}}{L^{3/2}} \\ &\qquad = \frac{e^{i\mathbf{k}\cdot\mathbf{r}}}{L^{3/2}} \sum_{\mathbf{k}'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu} - \mu) \; \frac{\eta}{L^3} \overline{V}_{(\mathbf{k}'-\mathbf{k})} \end{aligned} \tag{10}$$

where (with the change of variable $\mathbf{r} - \mathbf{r}' = \mathbf{s}$):

$$\overline{V}_{(\mathbf{k}'-\mathbf{k})} = \int \mathrm{d}^3 s \; e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{s}} \; W_2(\mathbf{s}) \tag{11}$$

Consequently, the exchange term simply multiplies the plane wave $e^{i\mathbf{k}\cdot\mathbf{r}}/L^{3/2}$ by:

$$\frac{\eta}{L^3}\sum_{\mathbf{k}'} f_\beta(\widetilde{e}_{\mathbf{k},\nu}-\mu)\ \overline{V}_{(\mathbf{k}'-\mathbf{k})} \tag{12}$$

To sum up, we showed that, for a uniform system, the plane waves are indeed solutions of the Hartree-Fock equations (3). It is no longer necessary to solve the eigenvector equations, but we simply have to replace in (3) the $\varphi^\nu_{\mathbf{k}}(\mathbf{r})$ by plane waves. This leads to:

$$\boxed{\widetilde{e}_{\mathbf{k},\nu} = e_k + \frac{\overline{V}_0}{L^3}\sum_{\mathbf{k}',\nu'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu'}-\mu) + \frac{\eta}{L^3}\sum_{\mathbf{k}'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu}-\mu)\overline{V}_{(\mathbf{k}'-\mathbf{k})}} \tag{13}$$

where e_k is the kinetic energy of a free particle:

$$e_k = \frac{\hbar^2\mathbf{k}^2}{2m} \tag{14}$$

We have obtained self-consistent conditions for the eigenvalues, which are a set of coupled nonlinear equations because of the $f_\beta(\widetilde{e}_{\mathbf{k}',\nu}-\mu)$ dependence on the energies $\widetilde{e}_{\mathbf{k}',\nu}$.

Comment:

The exchange term contains the Fourier transform at the spatial frequency $\mathbf{k}'-\mathbf{k}$ of the particle interaction potential; the direct term, however, contains the Fourier component at zero spatial frequency. This property is easy to understand from a physical point of view. Consider two particles, having respectively an initial momentum $\hbar\mathbf{k}$ and $\hbar\mathbf{k}'$. We saw in Chapter VIII (§ B-4-a) that the effect, to first order (Born approximation), of an interaction potential is proportional to the Fourier transform of that potential, calculated at the value of the variation of the relative momentum between the two particles (Chapter VII, § B-2-a); this variation is none other than the momentum transfer between the particle as they interact. Consequently, it is normal that the system energy is the sum of two terms: a direct term where no particle changes its momentum (no momentum transfer, hence a Fourier variable equal to zero); and another one where the two particles exchange their momenta, so that the relative momentum changes sign and the Fourier variable is proportional to the difference $\mathbf{k}'-\mathbf{k}$.

2-b. Quasi-particules

Equations (13) yield the individual energies $\widetilde{e}_{\mathbf{k},\nu}$, which are the sum of the free particle energy $\hbar^2k^2/2m$ and a contribution from the interactions. One can look at them as energies of individual objects[1], often referred to as "quasi-particles". The populations of the corresponding levels, as well as the total number of quasi-particles, are given by the same distribution functions f_β as for an ideal gas – see relations (39) and (44) of Complement G_{XV}. The same is true for the system entropy S, as we already mentioned at the end of § 2-b-α of that complement. Provided we replace the free particle energy by the modified energies $\widetilde{e}_{\mathbf{k},\nu}$, the analogy with independent particles is quite strong.

[1]The concept of quasi-particle is not necessarily limited to systems whose interacting particles are free inside a box; it remains valid for non-zero V_1 potentials (a harmonic potential for example). The first term in (13) must then be replaced by the particle energy in the potential V_1, and the direct and exchange terms will have a different expression.

3. Spontaneous magnetism of repulsive fermions

Consider a system of spin 1/2 fermions, contained in a box. To make the computations easier, we shall make a few simplifying hypotheses. They lead to a simple model, giving a good illustration of the nonlinear character of the Hartree-Fock theory. They involve the resolution of a set of nonlinear equations, containing only two variables – equations we will write in (22).

3-a. A simple model

We assume the mutual interaction potential to be repulsive and to have a very short range r_0. In relation (13) the only vectors $\mathbf{k}$ and $\mathbf{k}'$ that matter are those for which the distributions $f_\beta(\widetilde{e}_{\mathbf{k},\nu}-\mu)$ and $f_\beta(\widetilde{e}_{\mathbf{k}',\nu}-\mu)$ are not negligible. If, for all these vectors, the products kr_0 and $k'r_0$ are very small compared to 1, the product $(\mathbf{k}'-\mathbf{k})\cdot\mathbf{s}$ may be replaced by zero in (11), and we get:

$$\overline{V}_{(\mathbf{k}'-\mathbf{k})} \simeq \overline{V}_0 \tag{15}$$

α. Energy of the quasi-particles; spin state populations

As $\eta = -1$ for fermions, equation (13) can be simplified:

$$\widetilde{e}_{\mathbf{k},\nu} = e_k + \frac{\overline{V}_0}{L^3}\left[\sum_{\mathbf{k}',\nu'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu'}-\mu) - \sum_{\mathbf{k}'} f_\beta(\widetilde{e}_{\mathbf{k}',\nu}-\mu)\right] \tag{16}$$

or else:

$$\widetilde{e}_{\mathbf{k},\nu} = e_k + \frac{\overline{V}_0}{L^3}\sum_{\mathbf{k}',\nu'\neq\nu} f_\beta(\widetilde{e}_{\mathbf{k}',\nu'}-\mu) \tag{17}$$

Consequently, the energy of a quasi-particle with a given ν is only modified by the interaction with quasi-particles having a different spin component ν' (opposite spin if the particles have a spin $S = 1/2$). This result was to be expected since if the spins of the two quasi-particles are parallel, they cannot be distinguished; the Pauli principle then forbids them to approach at a distance closer than r_0, and they cannot interact. On the other hand, if their spins are opposite, they can be identified by the direction of their spin (we have assumed the interaction does not act on the spins): they behave as distinguishable particles, the exclusion principle does not apply, and they now interact.

We note $\left\langle \widehat{N}_+ \right\rangle$ and $\left\langle \widehat{N}_- \right\rangle$ the total particle numbers respectively in the spin state $+$ or $-$:

$$\left\langle \widehat{N}_\pm \right\rangle = \sum_{\mathbf{k}'} f_\beta(\widetilde{e}_{\mathbf{k}',\pm}-\mu) \tag{18}$$

Equation (17) shows that the energies of the $+$ and $-$ spin states are modified according to:

$$\widetilde{e}_{\mathbf{k},+} = e_k + g\left\langle \widehat{N}_- \right\rangle \quad ; \quad \widetilde{e}_{\mathbf{k},-} = e_k + g\left\langle \widehat{N}_+ \right\rangle \tag{19}$$

where the coupling constant g (having the dimension of an energy) is defined by:

$$g = \frac{\overline{V}_0}{L^3} \tag{20}$$

Since the particle numbers only depend on the difference between the energies $\widetilde{e}$ and the chemical potential μ, we can account for the terms in g appearing in (19) by keeping the energies e_k for free particles, but lowering the chemical potentials $\mu_\pm$ by the quantity $g\left\langle \widehat{N}_\pm \right\rangle$. Calling $N_{ig}^{FD}(\beta,\mu)$, as in relation (47) of Complement B_{XV}, the total number of fermions in an ideal gas:

$$N_{ig}^{FD}(\beta,\mu) = \left(\frac{L}{2\pi}\right)^3 \int \mathrm{d}^3k \frac{1}{e^{\beta(e_k-\mu)}+1} \tag{21}$$

we get, for an interacting gas:

$$\begin{cases} \left\langle \widehat{N}_+ \right\rangle = N_{ig}^{FD}\left(\beta, \mu - g\left\langle \widehat{N}_- \right\rangle\right) \\ \left\langle \widehat{N}_- \right\rangle = N_{ig}^{FD}\left(\beta, \mu - g\left\langle \widehat{N}_+ \right\rangle\right) \end{cases} \tag{22}$$

These equations determine the populations of the two spin states as a function of the parameters β (or the temperature), μ and finally the volume $\mathcal{V}$. These are, however, two coupled equations since the population $\left\langle \widehat{N}_+ \right\rangle$ depends on $\left\langle \widehat{N}_- \right\rangle$ and conversely. Finding their solution is not obvious, and we shall use a change of variables and resort to a graphic construction.

β. Change of variables

It is useful to write the previous relations in terms of dimensionless variables. We shall thus introduce the "thermal wavelength" λ_T by:

$$\lambda_T = \hbar\sqrt{\frac{2\pi}{mk_BT}} = \hbar\sqrt{\frac{2\pi\beta}{m}} \tag{23}$$

We can now make the same change of integration variable as in § 4-a of Complement B_{XV}:

$$\kappa = k\,\frac{\lambda_T}{2\sqrt{\pi}} \tag{24}$$

Relation (21) then becomes:

$$N_{ig}^{FD}(\beta,\mu) = \left(\frac{L}{\lambda_T}\right)^3 I_{3/2}(\beta\mu) \tag{25}$$

where:

$$I_{3/2}(\beta\mu) = \pi^{-3/2}\int \mathrm{d}^3\kappa \frac{1}{e^{\kappa^2-\beta\mu}+1} \tag{26}$$

These relations are just the same as those written in (51) and (52) in Complement B_{XV}. The value of $I_{3/2}$ only depends on a dimensionless variable, the product $\beta\mu$. As opposed

to N_{ig}^{FD} which is an "extensive" quantity (proportional[2] to the volume $\mathcal{V} = L^3$ of the system), $I_{3/2}(\beta\mu)$ is an "intensive" quantity (independent of the volume).

We can also replace the two unknowns, which are the populations $\left\langle \widehat{N}_{\pm} \right\rangle$ of the two spin states, by two dimensionless and intensive variables $x_{\pm}$:

$$x_{\pm} = \left(\frac{\lambda_T}{L} \right)^3 \left\langle \widehat{N}_{\pm} \right\rangle \tag{27}$$

To characterize the interactions appearing in relations (22) via the constant g, we introduce the dimensionless parameter $\overline{g}$:

$$\overline{g} = \beta g \left(\frac{L}{\lambda_T} \right)^3 = \frac{m\overline{V}_0}{2\pi\hbar^2\lambda_T} \tag{28}$$

Replacing in equations (22) the $\left\langle \widehat{N}_{\pm} \right\rangle$ by the $x_{\pm}$, and g by $\overline{g}$, we get the simpler form:

$$\begin{cases} x_+ = F^{(1)}(x_-) \\ x_- = F^{(1)}(x_+) \end{cases} \tag{29}$$

where the function $F^{(1)}$ is defined by:

$$F^{(1)}(x) = I_{3/2}\left(\beta\mu - \overline{g}x\right) \tag{30}$$

(this function depends not only on x, but also on the parameters $\beta\mu$ and $\overline{g}$). As (22), the system (29) contains two coupled equations: x_- allows computing directly x_+, and vice versa.

3-b. Resolution of the equations by graphical iteration

We now show how to solve equations (29) by a graphical method. The two variables x_+ and x_- can be uncoupled by noting that:

$$x_+ = F^{(1)}\left[F^{(1)}(x_+)\right] \tag{31}$$

with the same equation for x_-. We now introduce a second iterated function $F^{(2)}$ of the function $F^{(1)}$ (function of the same function) as:

$$F^{(2)}(x) = F\left[F(x)\right] \tag{32}$$

This leads to:

$$x_+ = F^{(2)}(x_+) \tag{33}$$

Applying the function $F^{(2)}$ to the variable x yields the same value x, which is said to be the abscissa of a "fixed point" of this function. Graphically, the fixed points of any function F are at the intersections of the curve representing the function F with the first bisector.

[2] In a more general way, Appendix VI recalls that a quantity is said to be extensive if, in the limit of large volumes (for T and μ constant), its ratio to the volume $\mathcal{V}$ tends toward a constant; this does not prevent the quantity from containing terms in L^2 for example. On the other hand, it is said to be intensive if, in this same limit (and without dividing by the volume), it tends toward a constant.

α. Iterations of a function

Consider, from a general point of view, the equation:

$$x = G(x) \tag{34}$$

whose solutions correspond to the fixed points of the function G. These solutions can be found by iteration: starting from an approximate value x_1 of the solution, we compute $G(x_1)$, then use $x_2 = G(x_1)$ as a new value of the variable to compute $x_3 = G(x_2)$, etc. It can be shown that this iteration process converges toward the solution of equation (34), hence toward the fixed point on the first bisector, if the slope of the function G at that point is included between -1 and $+1$, that is if:

$$-1 < G'(x) < +1 \tag{35}$$

where G' is the derivative of the function G. The fixed point of the application G is then said to be stable. On the other hand, if that slope is outside the interval $[-1, +1]$, the fixed point on the first bisector becomes unstable; the iteration method for G no longer converges.

We can also introduce the "second iterated function" $G^{(2)}(x) = G\,[G(x)]$. Any fixed point of G is necessarily a fixed point of $G^{(2)}$. The inverse is not true, as it is possible to get a "two-order cycle" where two different values of x are swapped under the effect of G:

$$\begin{aligned} x_2 &= G\,(x_1) \\ x_1 &= G\,(x_2) \end{aligned} \tag{36}$$

In such a case, x_1 and x_2 are both fixed points of $G^{(2)}$, but not of G (we shall see below an example of such a situation, illustrated by Figure 3). These fixed points can be stable for $G^{(2)}$, in which case they constitute a "stable cycle of order two" for the initial function G. After a certain number of iterations of G, the solution converges toward a series taking alternatively two distinct values, x_1 and x_2.

The process may repeat itself: it is possible for the fixed point of $G^{(2)}$ to next become unstable, and yield fixed points for an iterated function of a higher order, and hence to a stable cycle of that order.

β. Form of the function $F^{(1)}$

Relation (25) shows that the variations of $I_{3/2}(x)$ as a function of x are very similar to the variations of $N_{ig}^{FD}(\beta, \mu)$ as a function of μ, already studied in Complement B_{XV} (Figure 2). The equality (30) shows that the plot of $F^{(1)}(x)$ can be deduced from that of $I_{3/2}(x)$:

- reversing the variable x (symmetry with respect to the vertical axis)
- multiplying this variable by $\overline{g}$ (scale change of the abscissa)
- and finally shifting to the right the abscissa origin by the value $\beta\mu$.

This leads to the solid line curve in Figure 1, that plots a constantly decreasing function (for fixed values of β, μ and $\overline{g}$).

As the parameter $\overline{g}$ changes (for β and μ constant), we get a set of curves representing $F^{(1)}(x)$ for each value of $\overline{g}$. For $x = 0$, all these curves go through the same point of ordinate $I_{3/2}(\beta\mu)$. If $\overline{g} = 0$, the curve is a simple horizontal line going through this point. When $\overline{g}$ becomes slightly positive, the curve starts decreasing but still extends along the abscissa axis; it has a small negative slope at the origin. As $\overline{g}$ increases more and more, the curve contracts more and more toward the ordinate axis and its slope at the origin is more and more negative. In the limit where $\overline{g}$ goes to infinity, the curve becomes a straight vertical line.

γ. Form of the function $F^{(2)}$

Figure 1 also shows the geometric construction used to go from $F^{(1)}$ to $F^{(2)}$. For a given value of x, we start from point M_1 of ordinate $F^{(1)}(x)$, and draw a horizontal line until it crosses the first bisector in M_2, which transfers the ordinate of M_1 onto the abscissa. From the intersection point M_2 we draw a vertical line that intersects the function $F^{(1)}$ at point M_3 of ordinate $F^{(2)}(x)$, which we simply transfer to the initial abscissa x to get the final point (surrounded by a triangle).

This construction shows that $F^{(2)}(x)$ is an increasing function of x confined between two horizontal asymptotes: the abscissa axis, and a horizontal line of ordinate $I_{3/2}(\beta\mu)$. The larger the value of $\overline{g}$, the faster the increase of $F^{(2)}(x)$.

δ. Influence of the coupling parameter on the fixed points

We now discuss the influence of the parameter $\overline{g} \sim \overline{V}_0$ on the stability of single or multiple fixed points.

(i) The trivial case where $\overline{V}_0 = \overline{g} = 0$ (no interaction between the fermions) is particularly simple: the two curves are now identical horizontal lines, whose zero slope makes their intersection with the first bisector obviously stable. We then get the ideal gas results, with equal x_+ and x_- densities.

(ii) As long as $\overline{g}$ (hence $\overline{V}_0$) is weak enough, the slope of $F^{(1)}$ at the intersection point is small and the corresponding fixed point remains stable, as shown in Figure 1. This same point is obviously a fixed point for $F^{(2)}$ as well; as the derivative of a function $f[f(x)]$ with respect to x is $f'[f(x)]\,f'(x)$, that is $[f'(x)]^2$ at a point where $x = f(x)$, the slope of $F^{(2)}$ is less than 1, and that fixed point is also stable.

In such a case, both functions have only one common fixed point, which determines the only solution of the equations: the numbers x_+ and x_- are necessarily equal since they correspond to a fixed point of $F^{(1)}$. The only effect of the fermion repulsion is to lower in an equal way the densities associated with each of the spin states.

(iii) If now $\overline{g}$ (or $\overline{V}_0$) gets larger, we come to a situation, for a certain critical value g_c of g, where the slope of $F^{(1)}$ at the intersection with the first bisector becomes equal to -1, and that of $F^{(2)}$ now takes the value $+1$. The corresponding critical situation is plotted in Figure 2, where the curve representing the function $F^{(2)}$ is now tangent to the first bisector at their intersection (even osculating[3] to it, as their contact is of order two). For both functions, the fixed point is now right at the border of its stability domain.

[3]The first derivative of the function $f[f(x)]$ is equal to $f'(x) \times f'[f(x)]$, and its second derivative, to $f''(x)f'[f(x)]+[f'(x)]^2 f''[f(x)]$. At a fixed point, that second derivative becomes $f'(x)f''(x)[1 + f'(x)]$, which cancels out when $f'(x) = -1$.

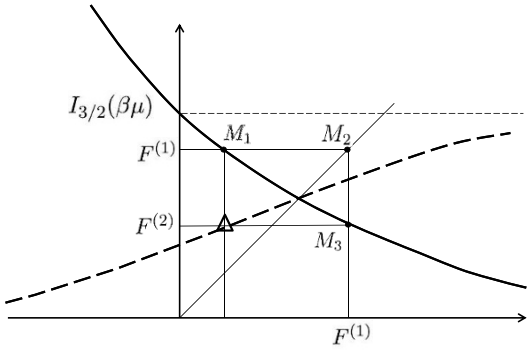


Figure 1: Plots of the functions $F^{(1)}$ (solid line) and $F^{(2)}$ (dashed line) as a function of x. Starting from any initial value of the variable x, we place a point M_1 on the first iterated $F^{(1)}$ curve, whose ordinate we transfer on the x axis by using the first bisector (point M_2); a new vertical intersection with the solid line (point M_3) yields the second iterated value $F^{(2)}$, that must be simply transferred to the initial value of the variable x (final point surrounded by a triangle). The whole dashed line curve can thus be constructed point by point. This method shows that when $x \to -\infty$, that curve is asymptotic to the abscissa axis; when $x \to +\infty$, the curve now has another horizontal asymptote, of ordinate $F^{(1)}(0) = I_{3/2}(\beta\mu)$, represented by a line with smaller dashes. The general form of the second iterated function is plotted in the figure: a uniformly increasing function between those two asymptotic values. In the case represented here, the coupling constant $\overline{g}$ is supposed to be weak enough for the two curves to intersect the first bisector with slopes of moduli less than 1; we then get a unique stable solution where x_- and x_+ are equal, which corresponds to a non-polarized spin system.

(iv) Beyond that situation, as shown in Figure 3, the curve representing $F^{(2)}$ intersects the first bisector in three points; the middle one is unstable as it corresponds to a slope larger than 1, but the two points on each sides are stable since they are associated with slopes between -1 and $+1$. As far as the central point is concerned, the slightest perturbation moves the iteration away from that point. On the other hand, the other two points are fixed stable points for $F^{(2)}$; they correspond to a physically acceptable solution of equations (29). As those two points are not fixed for $F^{(1)}$, two different values, x_+ and x_-, are swapped under the action of the function $F^{(1)}$ (two-cycle fixed points, represented by the arrows in the figure). We get a solution of the equations where the spin state populations are different: the gas develops a spontaneous polarization when the repulsion goes beyond a certain critical value and a phase transition occurs.

Comment:

For convenience, we discussed the emergence of the spontaneous polarization as a function of the parameter $\overline{g} \sim \overline{V}_0$, for fixed values of β and μ: the plot of the curves is then simply

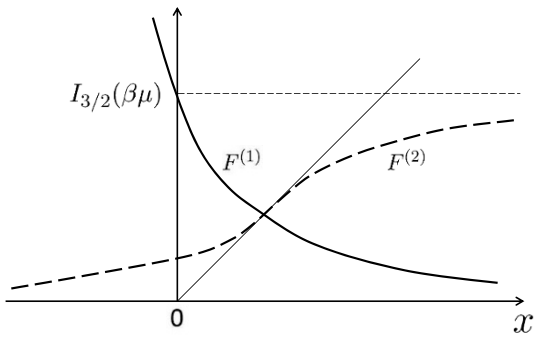


Figure 2: Plots of the functions $F^{(1)}$ (solid line) and $F^{(2)}$ (dashed line) in the critical case $g = g_c$, where the function $F^{(1)}$ intersects the first bisector with a slope equal to -1; the slope of $F^{(2)}$ is then equal to $+1$, and this function is not only tangent but also osculating to the first bisector (in other words it intersects this bisector at three points grouped together).

obtained by a scale change along the x axis. In general, however, the phase transition is observed by changing either the physical system density, or its temperature, while keeping the interactions constant. Our line of reasoning can be applied to this case, keeping the interactions constant while changing either the chemical potential μ that controls the particle density, or β, which is the inverse of the temperature. When either of these two parameters gets larger, the ordinate at the origin of $I_{3/2}(\beta\mu)$ gets higher, which increases the absolute values of the slopes of $F^{(1)}$ and $F^{(2)}$; the same phenomenon as above (instability and phase transition) will thus occur when the temperature is lowered or the density increased. If $I_{3/2}(\beta\mu) \gg 1$, relation (25) shows that the particle number contained in a volume $(\lambda_T)^3$ is large compared to one, which means that the average distance between the particles is smaller than the thermal wavelength; the fermion gas is then degenerate.

3-c. Physical discussion

As the spins carry a magnetic moment, a spontaneous polarization implies a transition towards a ferromagnetic phase. The origin of this phenomenon comes from an equilibrium between two opposite tendencies. On one hand, the "motor" of the transition comes from the fact that, to minimize the repulsion energy, the system tends to put all its particles in the same spin state (polarized system), which prevents them from interacting. This is because the Pauli principle forbids them to be at the same point in space, and as we assumed their interaction potential to be of zero-range, they can no longer interact. On the other hand, the system polarization (for a fixed total density) increases its kinetic energy: the same number of particles must be placed in a single Fermi sphere, instead of two, which results in a sphere with a larger radius, i.e. a higher Fermi level. It also changes the system entropy. The compromise between gain and loss (for

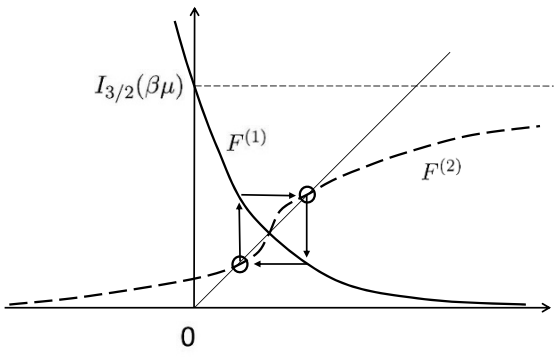


Figure 3: Beyond the critical point, the function $F^{(1)}$ intersects the first bisector with a slope less than -1, and there are now three distinct intersection points of the function $F^{(2)}$ and the bisector. The middle point corresponding to an $F^{(2)}$ slope larger than 1 is unstable, but the other two points (surrounded by circles) are stable. These two points are swapped under the action of $F^{(1)}$ (two-cycle fixed points, represented by the arrows). They yield different values for the spin densities, which leads to the appearance of a spontaneous spin polarization.

the grand potential) varies as a function of the parameters; when those parameters take a value where gain and loss balance each other perfectly, a spontaneous ferromagnetic transition occurs.

A more detailed study is possible; examining the shapes of the curves we plotted, we deduce that the conditions that favor the transition are: strong repulsion, high density, low temperature. It is worth noting that no Hamiltonian acting on the spins comes into play in this phase transition. Even though the interactions are totally independent of the spin, the Fermi-Dirac statistics has an effect on the spins, and can induce a transition polarizing those spins.

At the critical point (Figure 2), the two new stable points appear at the same place, and move away from each other in a continuous way. The phase transition is therefore continuous, which puts it into the category of second order phase transitions. The study of critical transitions is a very large domain of physics that we cannot discuss here in a general way. We can, however, take the analysis a little further, without too much difficulty: we note, from the equations written above, that the distance between the two stable points increases, beyond the critical point where $\mu = \mu_c$ and $\beta = \beta_c$, as the square root of the difference $\mu - \mu_c$ (or $\beta - \beta_c$). In other words, the system spontaneous magnetization varies as the square root of the distance to the critical point, which is typical of the so-called "Hopf bifurcation". In addition, at the critical point, the magnetic susceptibility of the spin system diverges.

Comments:

(i) A very general concept plays a role here: spontaneous symmetry breaking. The first symmetry breaking concerns the two opposed directions along the quantization axis

Oz. Equations (29) are invariant upon a permutation of x_+ and x_-; for any solution of these equations, there exists another one where these two variables are interchanged, and where the spin magnetization points in the opposite direction. This was to be expected since nothing physically distinguishes those two directions. The symmetry is said to be broken if the stable solutions of the set of equations are asymmetric, corresponding to different values of x_+ and x_-; there are then necessarily two (or more) distinct solutions, symmetric to one another.

Furthermore, the quantization axis Oz we used is arbitrary; had we chosen a different direction, we would have found that the spontaneous magnetization could point in any spatial direction. This again was to be expected since our problem is rotation invariant. The ferromagnetic transition phenomenon we have just studied corresponds to a spontaneous breaking of the rotational symmetry of the usual space, often called, in terms of group symmetry, "$SO(3)$ symmetry breaking". There are many other second order transitions that break various symmetry groups, as for example the symmetry $U(1)$ for the superfluid transition, etc.

(ii) A mean field theory like the one we used – i.e. an approximate theory – may identify the existence of a critical transition (second order transition) as explained above, but does not allow an exhaustive study of all its aspects, in particular in the vicinity of the critical point. Several critical phenomena (large wavelength critical fluctuations for example) cannot be accounted for with such an approximation, and require more elaborate theoretical methods.

4. Bosons: equation of state, attractive instability

For bosons, the equations (3) are very similar to those we used for fermions, except for a change of sign of η, and hence of the exchange potential. For a barely degenerate system, this modifies the interaction effects, but does not drastically change their consequences. On the other hand, for a system of degenerate bosons, the situation is radically different since expression (1) presents a singularity when $(e - \mu)$ is zero – whereas none occurs for fermions. As pointed out in Complement B_{XV}, this is the origin of the "Bose-Einstein condensation" phenomenon: as the chemical potential μ increases, the singularity becomes significant when μ gets close (through lower values) to the lowest individual energy among the e_i, that is close to the ground level energy e_0. The population of this level then increases more and more and can become "extensive" (proportional to the system volume $\mathcal{V}$ in the limit of large volumes).

Actually, using the Hartree-Fock equations for condensed boson systems leads to some difficulties, which will be briefly discussed below – see Comment (ii) of § 4-a. We shall limit ourselves to the study of non-condensed systems, not excluding the possibility that they approach condensation. We assume the bosons are without spin, and, as we did for fermions, that the range of the interaction potential $W_2(\mathbf{s})$ appearing in relation (11) is short enough so that:

$$\overline{V}_{(\mathbf{k}'-\mathbf{k})} = V_0 \tag{37}$$

In that case, the direct and exchange contributions in (13) are equal. For a homogeneous system, this equation then becomes:

$$\widetilde{e}_{\mathbf{k}} = e_k + \frac{2\overline{V}_0}{L^3} \sum_{\mathbf{k}'} f_\beta(\widetilde{e}_{\mathbf{k}'} - \mu) = e_k + \frac{2\overline{V}_0}{L^3} \langle N \rangle \tag{38}$$

where $\langle N \rangle$ is the average total number of particles:

$$\langle N \rangle = \sum_{\mathbf{k}} f_\beta(\widetilde{e}_{\mathbf{k}} - \mu) = \sum_{\mathbf{k}} f_\beta(e_{\mathbf{k}} - \mu - \Delta\mu) \tag{39}$$

with:

$$\Delta\mu = -\frac{2\overline{V}_0}{L^3} \langle N \rangle \tag{40}$$

We therefore find that the average total number of particles is the same as for a boson gas without interactions, provided the chemical potential μ is replaced by an effective chemical potential $\widetilde{\mu} = \mu + \Delta\mu$. The same holds true for the average population of each individual state $\mathbf{k}$.

As in Complement B_{XV}, we note $N_{ig}^{BE}(\beta, \mu)$ the function yielding the particle number for an ideal gas of bosons:

$$N_{ig}^{BE}(\beta, \mu) = \sum_{\mathbf{k}} f_\beta(e_{\mathbf{k}} - \mu) = \frac{\mathcal{V}}{(2\pi)^3} \int \mathrm{d}^3k \frac{1}{e^{\beta(e_k - \mu)} - 1} \tag{41}$$

(the second equality is valid for large volumes). Equation (39) then becomes:

$$\langle N \rangle = N_{ig}^{BE}(\beta, \widetilde{\mu}) \equiv N_{ig}^{BE}(\beta, \mu + \Delta\mu) \tag{42}$$

4-a. Repulsive bosons

For repulsive interactions, Figure 4 shows how to graphically obtain, by a geometric construction, the system density predicted by equation (42). For a given chemical potential, the particle number decreases because of the repulsion, which takes the system further away from condensation; consequently, its description by the Hartree-Fock equations is a good approximation.

To first order in $\overline{V}_0$, relation (42) may be approximated by:

$$\begin{aligned} \langle N \rangle &\simeq N_{ig}^{BE}(\beta, \mu) + \Delta\mu \frac{\partial}{\partial\mu} N_{ig}^{BE}(\beta, \mu) \\ &= N_{ig}^{BE}(\beta, \mu) - 2\overline{V}_0 \frac{N_{ig}^{BE}(\beta, \mu)}{L^3} \frac{\partial}{\partial\mu} N_{ig}^{BE}(\beta, \mu) \end{aligned} \tag{43}$$

Noting $\Phi(\beta, \mu)$ the grand potential, relation (62) of Appendix VI shows that:

$$\langle N \rangle = \frac{1}{\beta} \frac{\partial}{\partial\mu} \ln Z = -\frac{\partial}{\partial\mu} \Phi(\beta, \mu) \tag{44}$$

Integrating over μ relation (43) from $-\infty$ to the value μ, we get the grand potential:

$$\Phi(\beta, \mu) = \Phi_{ig}(\beta, \mu) + \frac{\overline{V}_0}{L^3} \left[N_{ig}^{BE}(\beta, \mu) \right]^2 \tag{45}$$

where $\Phi_{ig}(\beta, \mu)$ is the grand potential for the ideal gas, at the same temperature and chemical potential. In addition, relation (62) of Appendix VI shows that the grand potential is equal to minus the product of the volume and the pressure P:

$$\Phi(\beta, \mu) = -P\mathcal{V} \tag{46}$$

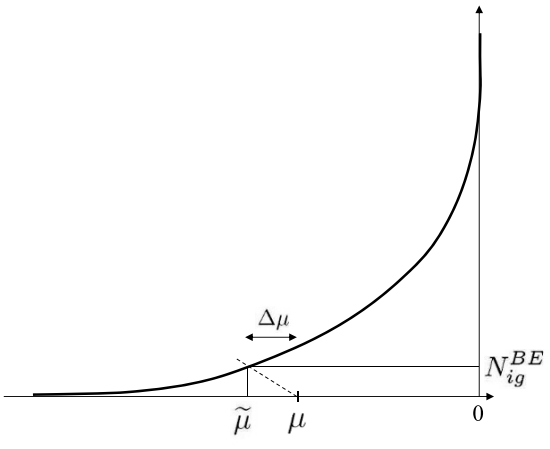


Figure 4: Geometric solutions of equations (40) and (42) for a gas of repulsive bosons. On the graph plotting the total number of particles as a function of the chemical potential, we draw a line starting from the point on the abscissa axis with chemical potential μ, and having the slope $-L^3/2\overline{V}_0$. It intersects the curve at a point whose abscissa is $\widetilde{\mu}$ and ordinate N_{ig}^{BE}. As $\overline{V}_0$ is positive for a repulsive gas, we see that the interactions lower the density (at constant temperature and chemical potential). The text explains how this geometric construction yields the equation of state for the interacting gas.

This means that if, at constant β, we vary the parameter μ in (43) and (45), we obtain in the plane $(\langle N\rangle, P)$ a curve representing the pressure as a function of the particle number in the volume $\mathcal{V}$, which is an isothermal line of the equation of state. Repeating this plot for several values of β, we get a set of curves covering the whole equation of state, taking into account the changes introduced by the interactions.

Comments:

(i) To keep the computations as simple as possible, we limited ourselves to the first order in $\overline{V}_0$. It is however possible to make the graphical construction of Fig. 4 more precise by including the higher order terms.

(ii) We discussed in § 3-b-β of Complement $\mathrm{G_{XV}}$ the limits of the Hartree-Fock approximation for bosons, which can no longer be used when the physical system gets too close to Bose-Einstein condensation. The graphical construction shown in Figure 4 loses its physical meaning if the intersection point on the curve is too close to the contact point of the curve with the vertical axis.

4-b. Attractive bosons

Attractive interactions ($\overline{V}_0 < 0$) result in an increase of the effective chemical potential, and consequently raises the value of $\langle N\rangle$. This in turn increases the effective chemical potential, and this positive feedback may even induce an avalanche effect leading to an instability if μ is too close to zero.

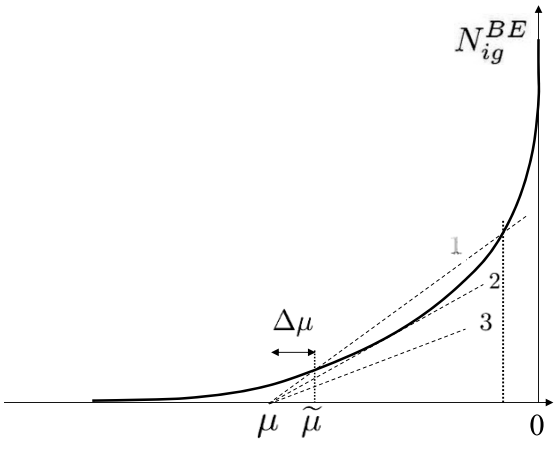


Figure 5: Graphical construction similar to that of Figure 4, but for an attractive boson gas (where $\overline{V}_0$ is negative). When the attractive potential $\overline{V}_0$ is not too large, the line noted 1 in the figure yields two possible solutions, only one of which is close to the solution in the absence of interactions, and hence suitable for our approximation. As $\overline{V}_0$ increases, for a certain critical value we only get one solution (tangent line 2), then none (line 3). In this last case, no solution signals an instability of the gas, which collapses onto itself because of the attractive interactions. Starting from a nearly condensed ideal gas, the closer it is to condensation, the weaker the attractive interactions necessary to trigger the instability.

The geometrical construction that yields $\Delta\mu$ and $\langle N \rangle$ from the intersection of a straight line with a curve is shown on Fig. 5. If $\overline{V}_0$ is weak enough, and for a fixed value of μ, we get two intersection points, corresponding to possible solutions. We only keep the first one, yielding the lowest value of $\Delta\mu$. The other point yields a high value of $\Delta\mu$, which changes radically and increases considerably the system density; in that case, chances are the approximate mean field treatment of the interactions is no longer valid. Beyond the value of $\overline{V}_0$ for which the straight line becomes tangent to the curve, the couple of equations (39) and (40) do not have a solution: there no longer exists any stable solution.

Figure 5 also shows that as the chemical potential gets closer to zero, the effect of the attractive interactions between bosons becomes more and more important; weak interactions are enough to render the system unstable. The reason we did not find any solution to the equations is that we assumed, in the computations, that the system was perfectly homogeneous; now this homogeneity cannot be maintained beyond a certain attraction intensity. We must therefore enlarge the theoretical framework, and include the possibility for the system to become spontaneously inhomogeneous. A more precise study would show that the system may develop local instabilities, hence breaking spontaneously the translation invariance symmetry. In the limit of large systems (thermodynamic limit), condensed bosons tend to collapse onto themselves under the effect

of an attractive interaction, however weak it may be[4].

As a general conclusion, the Hartree-Fock method applied to fermions yields results valid in a very large parameter range. As an example, it allowed computing effects of the interactions on the particle number and the pressure of the system. In addition, this method was able to predict the existence of phase transitions. This is also true for non-degenerate bosons, and the mean field method actually has a very large number of applications that we cannot detail here. We must, however, keep in mind that when Bose-Einstein condensation occurs, certain predictions pertaining to the condensate may not be realistic from a physical point of view, as they depend too closely on the mean field approximation which does not properly account for the correlations between particles.

[4]If the interaction potential is attractive at large distance, but strongly repulsive at short range (hard core for example), the system spontaneously forms a high density liquid or solid.

Chapter XVI

Field operator

A Definition of the field operator 1752
A-1 Definition 1752
A-2 Commutation and anticommutation relations 1754
B Symmetric operators 1755
B-1 General expression 1755
B-2 Simple examples 1756
B-3 Field spatial correlation functions 1758
B-4 Hamiltonian operator 1761
C Time evolution of the field operator (Heisenberg picture) . 1763
C-1 Contribution of the kinetic energy 1763
C-2 Contribution of the potential energy 1764
C-3 Contribution of the interaction energy 1764
C-4 Global evolution 1765
D Relation to field quantization 1765

Introduction

This chapter is a continuation of the previous chapter and uses the same mathematical tools. The main difference is that, until now, we have mainly used discrete bases in the individual state space, $\{|u_i\rangle\}$ or $\{|v_s\rangle\}$. In this chapter, we shall use a continuous basis, which is the basis, for spinless particles, of the position eigenvectors (see Chapter II, § E). As they now depend on the position $\mathbf{r}$, the creation and annihilation operators become field operators depending on a continuous subscript $\mathbf{r}$. They are the operator analog of the classical fields (which are numbers and not operators), and are often called "field operators". They are useful for concisely describing numerous properties of identical particle systems. They have commutation relations for bosons, and anticommutation relations for fermions. This chapter is a preparation for Chapters XIX and XX, where we will introduce the quantization of the electromagnetic field.

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

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

After defining these operators in § A-1, we discuss some of their properties. Their commutation and anticommutation relations are examined in § A-2. As in Chapter XV, we then study (§ B) symmetric operators and their expression as a function of the field operators; special attention will be given to the operators associated with the field correlation functions. In § C, we shall use the Heisenberg picture to study the time dependence of these operators. As a conclusion, we shall briefly come back (in § D) to the field quantization procedure and its link to the concept of identical particles.

A. Definition of the field operator

The field operator is defined as the annihilation operator a_i, but associated with a basis of individual states labeled by the continuous position index $\mathbf{r}$ instead of a discrete index i. Our starting point will be the basis change relation (A-52) of Chapter XV:

$$a_{v_s} = \sum_i \langle v_s |\ u_i \rangle\ \ a_{u_i} \tag{A-1}$$

where the subscripts i and s label the kets of the two orthonormal bases $\{|u_i\rangle\}$ and $\{|v_s\rangle\}$ in the individual state space. In what follows, and as already done in Chapter XV, we will simplify the notation and often replace the subscript u_i by i, and the subscript v_s by s.

A-1. Definition

We will first define the field operator for spinless particles, and later generalize it to the case with spin.

A-1-a. Spinless particles

We replace, in relation (A-1), the basis $\{|v_s\rangle\}$ by the basis of vectors $\{|\mathbf{r}\rangle\}$, where $\mathbf{r}$ symbolizes three continuous indices (the three vector components). The operator a_s now becomes an operator depending on the continuous index $\mathbf{r}$, that we shall call "*field operator*" for the system of identical particles we consider. We could write it simply as $a_{\mathbf{r}}$, but we shall prefer the commonly used classical notation $\Psi(\mathbf{r})$. Like any annihilation operator, $\Psi(\mathbf{r})$ acts in the Fock space where it lowers by one unit the particle number. In (A-1), the coefficient appearing in the sum is now the wave function $u_i(\mathbf{r})$ associated with the ket $|u_i\rangle$:

$$u_i(\mathbf{r}) = \langle \mathbf{r} |\ u_i \rangle \tag{A-2}$$

and this relation becomes:

$$\boxed{\Psi(\mathbf{r}) = \sum_i u_i(\mathbf{r})\ a_i} \tag{A-3}$$

Formally, definition (A-3) looks like the expansion of a wave function on the basis functions $u_i(\mathbf{r})$; here, however, the "components" a_i are operators and no longer simple complex numbers. In the same way as a_s annihilates a particle in the state $|v_s\rangle$, the operator $\Psi(\mathbf{r})$ now annihilates a particle at point $\mathbf{r}$.

It does not depend on the basis $\{|u_i\rangle\}$, i.e. on the wave functions chosen to define it in (A-3), as we now show. We insert in this equality the closure relation in any arbitrarily chosen basis $\{|v_s\rangle\}$, and use again (A-1):

$$\Psi(\mathbf{r}) = \sum_i \sum_s \langle \mathbf{r}|\ v_s\rangle\ \langle v_s|\ u_i\rangle\ \ a_{u_i} = \sum_s \langle \mathbf{r}|\ v_s\rangle\ \ a_{v_s} \tag{A-4}$$

(we temporarily came back to the explicit notation for the annihilation operators). We can thus write:

$$\Psi(\mathbf{r}) = \sum_s v_s(\mathbf{r})\ a_{v_s} \tag{A-5}$$

which means $\Psi(\mathbf{r})$ satisfies, in the new basis, a relation similar to (A-3).

We now take the Hermitian conjugate of (A-3):

$$\Psi^\dagger(\mathbf{r}) = \sum_i u_i^*(\mathbf{r})\ a_i^\dagger \tag{A-6}$$

The operator $\Psi^\dagger(\mathbf{r})$ creates a particle at point $\mathbf{r}$, as can be shown for example by computing the ket resulting from its action on the vacuum:

$$\Psi^\dagger(\mathbf{r})\,|0\rangle = \sum_i u_i^*(\mathbf{r})\ a_i^\dagger\,|0\rangle = \sum_i u_i^*(\mathbf{r})\ \ |u_i\rangle \tag{A-7}$$

that is:

$$\Psi^\dagger(\mathbf{r})\,|0\rangle = \sum_i\ \ |u_i\rangle\,\langle u_i|\ \mathbf{r}\rangle = |\mathbf{r}\rangle \tag{A-8}$$

which represents, as announced, a particle localized at point $\mathbf{r}$.

One can easily invert formulas (A-3) and (A-6) by writing, for example:

$$\begin{aligned}\int \mathrm{d}^3r\ u_i^*(\mathbf{r})\ \Psi(\mathbf{r}) = &\int \mathrm{d}^3r\ u_i^*(\mathbf{r})\ \sum_j u_j(\mathbf{r})\ a_j \\ &= \sum_j \delta_{ij}\ a_j\ =\ a_i\end{aligned} \tag{A-9}$$

or else, by Hermitian conjugation:

$$\int \mathrm{d}^3r\ u_i(\mathbf{r})\ \Psi^\dagger(\mathbf{r}) = a_i^\dagger \tag{A-10}$$

A-1-b. Particles with spin

When the particles have a spin S, the basis vector $\{|\mathbf{r}\rangle\}$ used above must be replaced by the basis vector $\{|\mathbf{r},\nu\rangle\}$, where ν is the spin index, which can take $2S+1$ discrete values ($\nu = -S,\ -S+1,\ ...,+S$). To all the summations over d^3r, we must now add a summation on the $2S+1$ values of the spin index ν. As an example, a basis vector $|u_i\rangle$ in the individual state space must now be written:

$$|u_i\rangle = \sum_{\nu=-S}^{+S} \int \mathrm{d}^3r\ u_i(\mathbf{r},\nu)\,|\mathbf{r},\nu\rangle \tag{A-11}$$

with:

$$u_i(\mathbf{r},\nu) = \langle \mathbf{r},\nu \,|u_i\rangle \tag{A-12}$$

The variables $\mathbf{r}$ and ν play a similar role. The first one is, however, continuous, whereas the second is discrete. Writing them in the same parenthesis might hide this difference and one often prefers noting the discrete index as a superscript of the function u_i, writing for example:

$$u_i^{\nu}(\mathbf{r}) = \langle \mathbf{r},\nu \,|u_i\rangle \tag{A-13}$$

Let us again use relation (A-1). On the left-hand side, the index v_s now symbolizes both the position $\mathbf{r}$ and the spin quantum number ν, which leads us to define a field operator $\Psi_\nu(\mathbf{r})$ having $2S+1$ spin components. Inserting (A-13) in the right-hand side of (A-1), we get:

$$\boxed{\Psi_\nu(\mathbf{r}) = \sum_i u_i^{\nu}(\mathbf{r})\; a_i} \tag{A-14}$$

The Hermitian conjugate operator $\Psi_\nu^\dagger(\mathbf{r})$ now creates a particle at point $\mathbf{r}$ with a spin ν:

$$\Psi_\nu^\dagger(\mathbf{r})\,|0\rangle = \sum_i \langle u_i|\;\mathbf{r},\nu\rangle \;\; |u_i\rangle = |\mathbf{r},\nu\rangle \tag{A-15}$$

As we did above, we can invert those relations. In relation (A.51) of Chapter XV (basis change), we replace v_s by i, and u_i by $\mathbf{r},\nu$ (which means that the summation over i is replaced by an integral over d^3r and a summation over ν), and use equality (A-11); we therefore get:

$$a_i^\dagger = \sum_{\nu=-S}^{+S} \int \mathrm{d}^3r \; u_i^{\nu}(\mathbf{r}) \; \Psi_\nu^\dagger(\mathbf{r}) \tag{A-16}$$

which is the analog, in the presence of spin, of relation (A-10).

A-2. Commutation and anticommutation relations

Commutation relations for field operators are analogous to those obtained in § A-5 of Chapter XV, but the discrete index i is now replaced by a continuous index.

A-2-a. Spinless particles

The commutator (or anticommutator) of two field operators:

$$[\Psi(\mathbf{r}) \;, \Psi(\mathbf{r}')]_{-\eta} = \sum_{i,j} u_i(\mathbf{r}) u_j(\mathbf{r}') \,[a_i, a_j]_{-\eta} = 0 \tag{A-17}$$

is indeed zero, as expected from expression (A-48) of Chapter XV. In the same way, by Hermitian conjugation:

$$\left[\Psi^\dagger(\mathbf{r}) \;, \Psi^\dagger(\mathbf{r}')\right]_{-\eta} = \sum_{i,j} u_i^*(\mathbf{r}) u_j^*(\mathbf{r}') \left[a_i^\dagger, a_j^\dagger\right]_{-\eta} = 0 \tag{A-18}$$

However, when we (anti)commute the field operator and the adjoint operator, we get:

$$\left[\Psi(\mathbf{r})\ ,\Psi^\dagger(\mathbf{r}')\right]_{-\eta} = \sum_{i,j} u_i(\mathbf{r})u_j^*(\mathbf{r}')\left[a_i, a_j^\dagger\right]_{-\eta} \tag{A-19}$$

which yields, taking into account the commutation relations (A-49) of Chapter XV:

$$\left[\Psi(\mathbf{r})\ ,\Psi^\dagger(\mathbf{r}')\right]_{-\eta} = \sum_{i} u_i(\mathbf{r})u_i^*(\mathbf{r}') = \sum_{i} \langle\mathbf{r}|\ u_i\rangle\langle u_i|\ \mathbf{r}'\rangle = \langle\mathbf{r}|\ \mathbf{r}'\rangle \tag{A-20}$$

Finally, we get:

$$\boxed{\left[\Psi(\mathbf{r})\ ,\Psi^\dagger(\mathbf{r}')\right]_{-\eta} = \delta(\mathbf{r}-\mathbf{r}')} \tag{A-21}$$

which is the equivalent of the relations (A-49) of Chapter XV in the case of a continuous basis.

A-2-b. Particles with spin

Relation (A-17) now becomes:

$$[\Psi_\nu(\mathbf{r})\ ,\Psi_{\nu'}(\mathbf{r}')]_{-\eta} = \sum_{i,j} u_i^\nu(\mathbf{r})u_j^{\nu'}(\mathbf{r}')\,[a_i, a_j]_{-\eta} = 0 \tag{A-22}$$

Relation (A-18) remains valid even if the field operators have spin indices. Finally, relation (A-19) becomes:

$$\begin{aligned}\left[\Psi_\nu(\mathbf{r})\ ,\Psi_{\nu'}^\dagger(\mathbf{r}')\right]_{-\eta} &= \sum_{i,j} u_i(\mathbf{r},\nu)u_j^*(\mathbf{r}',\nu')\left[a_i, a_j^\dagger\right]_{-\eta} \\ &= \langle\mathbf{r},\nu|\ \mathbf{r}',\nu'\rangle\end{aligned} \tag{A-23}$$

that is:

$$\boxed{\left[\Psi_\nu(\mathbf{r})\ ,\Psi_{\nu'}^\dagger(\mathbf{r}')\right]_{-\eta} = \delta_{\nu\nu'}\delta(\mathbf{r}-\mathbf{r}')} \tag{A-24}$$

B. Symmetric operators

In the previous chapter, we wrote one- or two-particle symmetric operators in terms of creation and annihilation operators in the discrete states $|u_i\rangle$. We are now going to express those operators in terms of the field operator (and its Hermitian conjugate).

B-1. General expression

We start with spinless particles. We can either directly transpose expressions (B-12) and (C-16) of Chapter XV to a continuous basis $\{|\mathbf{r}\rangle\}$ (replacing the sums by integrals), or insert in those expressions the form (A-9) for the operators a_i. In both cases, we get:

$$\widehat{F} = \int \mathrm{d}^3r \int \mathrm{d}^3r'\,\langle\mathbf{r}|\,\widehat{f}\,|\mathbf{r}'\rangle\ \ \Psi^\dagger(\mathbf{r})\Psi(\mathbf{r}') \tag{B-1}$$

and:

$$\widehat{G} = \frac{1}{2}\int \mathrm{d}^3 r \int \mathrm{d}^3 r' \int \mathrm{d}^3 r'' \int \mathrm{d}^3 r''' \\ \langle 1:\mathbf{r}; 2:\mathbf{r}'|\,\widehat{g}\,|1:\mathbf{r}''; 2:\mathbf{r}'''\rangle \;\; \Psi^\dagger(\mathbf{r})\Psi^\dagger(\mathbf{r}')\Psi(\mathbf{r}''')\Psi(\mathbf{r}'') \tag{B-2}$$

where, as in relation (C-16) of Chapter XV, the order of the annihilation operators is the inverse of the order that appears in the ket of the matrix element.

Expression (B-1) reminds us of the average value $\left\langle \widehat{f}(1) \right\rangle$ of the operator $\widehat{f}(1)$ for a single particle (without spin), described by the wave function $\psi(\mathbf{r}_1)$:

$$\left\langle \widehat{f}(1) \right\rangle = \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_1' \, \langle \mathbf{r}_1|\, \widehat{f}\,|\mathbf{r}_1'\rangle \;\; \psi^*(\mathbf{r}_1)\psi(\mathbf{r}_1') \tag{B-3}$$

Both expressions are not equivalent since (B-1) concerns any number of identical particles, rather than a single one; furthermore, the Ψ are now operators, and their respective order matters – as opposed to the order of the ψ in (B-3). As for formula (B-2), it can be compared to the average value of an operator $\widehat{g}(1,2)$ acting on two particles, 1 and 2, both described by the same wave function ψ. Here again, the order of the field operators is important, as opposed to the order in a product of wave functions.

For particles with spins, we simply complement each integral over $\mathbf{r}$ with a sum over the spin index ν, we include this index in the matrix elements and add a spin index to the field operator. As an example, relation (B-1) is generalized to:

$$\widehat{F} = \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \sum_{\nu=-S}^{S} \sum_{\nu'=-S}^{S} \langle \mathbf{r}, \nu|\, f\,|\mathbf{r}', \nu'\rangle \;\; \Psi_\nu^\dagger(\mathbf{r})\Psi_{\nu'}(\mathbf{r}') \tag{B-4}$$

As for relation (B-2), we get four summations over the spin indices ν, and the operator matrix elements are taken between bras and kets where an index ν is added to the variable $\mathbf{r}$.

B-2. Simple examples

We start with a few examples concerning operators for one spinless particle. For a single particle, the operator associated with the local density at point $\mathbf{r}_0$ is:

$$|\mathbf{r}_0\rangle \, \langle \mathbf{r}_0| \tag{B-5}$$

and its matrix elements are:

$$\langle \mathbf{r}|\, \mathbf{r}_0\rangle \, \langle \mathbf{r}_0 \,|\mathbf{r}'\rangle = \delta(\mathbf{r} - \mathbf{r}_0) \;\; \delta(\mathbf{r}' - \mathbf{r}_0) \tag{B-6}$$

The corresponding N-particle operator is written as:

$$\widehat{D}^{(N)}(\mathbf{r}_0) = \sum_{q=1}^{N} |q:\mathbf{r}_0\rangle \, \langle q:\mathbf{r}_0| \tag{B-7}$$

Replacing $\widehat{f}$ by $|\mathbf{r}_0\rangle \, \langle \mathbf{r}_0|$ in (B-1) yields the operator acting in the Fock space:

$$\widehat{D}(\mathbf{r}_0) = \Psi^\dagger(\mathbf{r}_0)\Psi(\mathbf{r}_0) \tag{B-8}$$

This operator annihilates a particle at point $\mathbf{r}_0$, and immediately recreates it at the same point. The average value:

$$\widehat{D}(\mathbf{r}_0) = \langle\Phi|\,\widehat{D}(\mathbf{r}_0)\,|\Phi\rangle \tag{B-9}$$

in a normalized state $|\Phi\rangle$ of the N-particle system yields the particle density associated with this state at point $\mathbf{r}_0$.

The operator $\widehat{N}$, total number of particles, has been written in (B-15) of Chapter XV. As the discrete summation index i is changed to a continuous index $\mathbf{r}$, the summation becomes an integral over all space:

$$\widehat{N} = \int \mathrm{d}^3r\ \Psi^\dagger(\mathbf{r})\Psi(\mathbf{r}) \tag{B-10}$$

As expected, it is the integral over d^3r of the operator $\widehat{D}(\mathbf{r})$.

The operator $V_1(\mathbf{r})$ describing the one-particle potential energy is also diagonal in the position representation; in the Fock space, it becomes the operator $\widehat{V}_1$:

$$\widehat{V}_1 = \int \mathrm{d}^3r\ V_1(\mathbf{r})\Psi^\dagger(\mathbf{r})\Psi(\mathbf{r}) \tag{B-11}$$

As for the particle current, it can be deduced from the expression of the current $\widehat{\mathbf{j}}(\mathbf{r}_0)$ associated with a single particle of mass m; it is the product of the operators giving the local density $|\mathbf{r}_0\rangle\langle\mathbf{r}_0|$ and the velocity $\widehat{\mathbf{p}}/m$ (product that must obviously be symmetrized):

$$\widehat{\mathbf{j}}(\mathbf{r}_0) = \frac{1}{2}\left[|\mathbf{r}_0\rangle\langle\mathbf{r}_0|\,\frac{\widehat{\mathbf{p}}}{m} + \frac{\widehat{\mathbf{p}}}{m}\,|\mathbf{r}_0\rangle\langle\mathbf{r}_0|\right] \tag{B-12}$$

If the particle is described by a wave function $\psi(\mathbf{r})$, a simple calculation[1] shows that the average value of this operator yields the usual expression for the probability current – see equation (D-17) of Chapter III:

$$\mathbf{J}(\mathbf{r}_0) = \frac{\hbar}{2mi}\left[\psi^*(\mathbf{r}_0)\boldsymbol{\nabla}\psi(\mathbf{r}_0) - \psi(\mathbf{r}_0)\boldsymbol{\nabla}\psi^*(\mathbf{r}_0)\right] \tag{B-14}$$

The current $\mathbf{J}(\mathbf{r}_0)$ of a system of identical particles is obtained by replacing in (B-1) the operator $\widehat{f}$ by $\widehat{\mathbf{j}}(\mathbf{r}_0)$:

$$\widehat{\mathbf{J}}(\mathbf{r}_0) = \frac{\hbar}{2mi}\left[\Psi^\dagger(\mathbf{r}_0)\boldsymbol{\nabla}\Psi(\mathbf{r}_0) - \Psi(\mathbf{r}_0)\boldsymbol{\nabla}\Psi^\dagger(\mathbf{r}_0)\right] \tag{B-15}$$

Another way of arriving at this equality is to use the substitution process mentioned in § B-1. To obtain the operator we are looking for in terms of $\Psi(\mathbf{r})$, we start from the

[1] Let us calculate the average value $\langle\Psi|\,\widehat{\mathbf{j}}(\mathbf{r}_0)\,|\Psi\rangle$. The first term on the right-hand side of (B-12) yields:

$$\frac{1}{2m}\,\langle\Psi\ |\mathbf{r}_0\rangle\langle\mathbf{r}_0|\,\widehat{\mathbf{p}}\,|\boldsymbol{\Psi}\rangle = \frac{\hbar}{2mi}\Psi^*(\mathbf{r}_0)\boldsymbol{\nabla}\psi(\mathbf{r}_0) \tag{B-13}$$

since the action of the operator $\widehat{\mathbf{p}}$ in the position representation is given by $(\hbar/i)\boldsymbol{\nabla}$. The second term is its complex conjugate, and we therefore get (B-14).

expression of the average value for a single particle, described by the wave function $\psi(\mathbf{r})$, which is then replaced by the field operator $\Psi(\mathbf{r})$.

For particles with spin, the local density at point $\mathbf{r}_0$, with spin ν, is written in the same way:

$$\widehat{D}^{(N)}(\mathbf{r}_0,\nu) = \sum_{q=1}^{N} |q:\mathbf{r}_0,\nu\rangle \langle q:\mathbf{r}_0,\nu| \tag{B-16}$$

and yields, in the Fock space, the operator:

$$\widehat{D}(\mathbf{r}_0,\nu) = \Psi_\nu^\dagger(\mathbf{r}_0)\Psi_\nu(\mathbf{r}_0) \tag{B-17}$$

The total density is obtained by a summation over ν:

$$\widehat{D}(\mathbf{r}_0) = \sum_{\nu=-S}^{S} \Psi_\nu^\dagger(\mathbf{r}_0)\Psi_\nu(\mathbf{r}_0) \tag{B-18}$$

For particles with spin, the operator $\widehat{N}$ associated with the total particle number, or the operator probability current $\widehat{\mathbf{J}}_{\mathbf{r}_0}$, can be obtained in a similar way.

B-3. Field spatial correlation functions

Operators associated with spatial correlation functions can also be defined with field operators; their average values are very useful for characterizing the field properties at different points of space. When we reason in terms of fields, we generally characterize each correlation function by the number of points concerned, which is different from the number of particles involved: the two-point functions characterize the properties of a single particle, the four-point ones concern two particles, etc. The reason is simple: a one-particle density operator ρ_I is characterized by non-diagonal elements $\langle \mathbf{r}_0'| \rho_I |\mathbf{r}_0\rangle$ depending on two positions, a two-particle density operator ρ_{II} involves elements depending on four positions, etc.

B-3-a. Two-point correlation functions

Defining a non-diagonal operator depending on two parameters $\mathbf{r}_0$ and $\mathbf{r}_0'$, we can generalize relation (B-5):

$$|\mathbf{r}_0\rangle \langle \mathbf{r}_0'| \tag{B-19}$$

A calculation very similar to the one leading to equation (B-7) – we simply add a "prime" to the second $\mathbf{r}_0$ – gives the N-particle symmetric operator:

$$\widehat{D}^{(N)}(\mathbf{r}_0,\mathbf{r}_0') = \sum_{q=1}^{N} |q:\mathbf{r}_0\rangle \langle q:\mathbf{r}_0'| \tag{B-20}$$

which yields, in the Fock space, the operator:

$$\widehat{D}(\mathbf{r}_0,\mathbf{r}_0') = \Psi^\dagger(\mathbf{r}_0)\Psi(\mathbf{r}_0') \tag{B-21}$$

We thus obtain an operator that annihilates a particle at point $\mathbf{r}'_0$ and recreates it at a different point $\mathbf{r}_0$.

When the N-particle system is described by a quantum state $|\Phi\rangle$, we call *two-point field correlation function*, $G_1(\mathbf{r}_0, \mathbf{r}'_0)$, the average value:

$$G_1(\mathbf{r}_0, \mathbf{r}'_0) = \left\langle \Psi^\dagger(\mathbf{r}_0)\Psi(\mathbf{r}'_0) \right\rangle = \langle\Phi|\, \Psi^\dagger(\mathbf{r}_0)\Psi(\mathbf{r}'_0)\, |\Phi\rangle \tag{B-22}$$

which also yields the matrix element, in the $\{|\mathbf{r}\rangle\}$ representation, of the one-particle operator[2] ρ_I:

$$\left\langle \Psi^\dagger(\mathbf{r}_0)\Psi(\mathbf{r}'_0) \right\rangle = \langle \mathbf{r}'_0|\, \rho_I\, |\mathbf{r}_0\rangle \tag{B-23}$$

Demonstration:

The matrix elements of the density operator $\rho_I^{(q)}$ of particle q are:

$$\left\langle \mathbf{r}'_0 \right| \rho_I^{(q)} \left| \mathbf{r}_0 \right\rangle = \mathrm{Tr}\left\{ \rho_I^{(q)} \left| q : \mathbf{r}_0 \right\rangle \left\langle q : \mathbf{r}'_0 \right| \right\} = \left\langle \left| q : \mathbf{r}_0 \right\rangle \left\langle q : \mathbf{r}'_0 \right| \right\rangle \tag{B-24}$$

For an N-particle system, we define the one-particle density operator ρ_I by a sum over all the particles:

$$\rho_I = \sum_{q=1}^{N} \rho_I^{(q)} \tag{B-25}$$

(be careful: the trace of this operator is N, not 1). Its matrix elements are the sum of the average values written in (B-24), i.e. the average value of the one-particle symmetric operator obtained by summing over q the $|q : \mathbf{r}_0\rangle \langle q : \mathbf{r}'_0|$. This result is simply the operator $\widehat{D}^{(N)}(\mathbf{r}_0, \mathbf{r}'_0)$ of (B-20), which, as seen above, yields in the Fock space expression (B-21). We then simply take the average value of each side of this expression to get equality (B-23).

The average value (B-22) at two different points plays an important role in the study of Bose-Einstein condensation. For a system at thermal equilibrium, this average value generally tends to zero rapidly as the distance between $\mathbf{r}$ and $\mathbf{r}'$ increases; it only remains non-zero in a domain of microscopic size. However, for a Bose-Einstein condensed gas, this average value behaves quite differently as it tends toward a non-zero value at large distance. This difference is actually the "Penrose and Onsager condensation criterion "; they have defined the existence of such a condensation as the appearance of a non-zero value of the matrix element of ρ_I at large distance; this definition is quite general as it applies not only to the ideal gas but also to systems of interacting particles.

Particles with spin:

If the particles have a non-zero spin S, we use, as a basis, the kets $|\mathbf{r}, \nu\rangle$ where ν takes the $(2S+1)$ values $-S$, $-S+1$, .., $+S$, and we add a ν index to the field operators. We then define $(2S+1)^2$ two-point field correlation functions as the average values:

$$\left\langle \Psi^\dagger_\nu(\mathbf{r}_0)\Psi_{\nu'}(\mathbf{r}'_0) \right\rangle = G_1(\mathbf{r}_0, \nu; \mathbf{r}'_0, \nu') \tag{B-26}$$

[2]Note the inversion of the variable order between the function G_1 (or the variables of $\Psi^\dagger$ and Ψ) and the matrix element of ρ_I.

The same computation that led to (B-23) for spinless particles, can be repeated with no changes other than the simple replacement of the kets (or bras) $|\mathbf{r}\rangle$ by $|\mathbf{r},\nu\rangle$; it shows that these average values yield the matrix elements of the one-particle density operator:

$$\left\langle \mathbf{r}_0',\nu' \right| \rho_I \left| \mathbf{r}_0,\nu \right\rangle = \left\langle \Psi_\nu^\dagger(\mathbf{r}_0)\Psi_{\nu'}(\mathbf{r}_0') \right\rangle \tag{B-27}$$

(here again we have chosen to normalize to N the trace of the one-particle density operator).

B-3-b. Higher order correlation functions

One can also start with the two-particle operator, which now depends on four positions:

$$\widehat{g}(1:\mathbf{r}_0,\mathbf{r}_0';2:\mathbf{r}_0'',\mathbf{r}_0''') = |1:\mathbf{r}_0\rangle\langle 1:\mathbf{r}_0'| \otimes |2:\mathbf{r}_0''\rangle\langle 2:\mathbf{r}_0'''| \tag{B-28}$$

In this case, the expression of $\widehat{g}$ is not symmetric with respect to the exchange of particles 1 and 2, as opposed to what happens for an interaction energy. The operator $\widehat{G}$ is then defined without the 1/2 factor of relation (C-1) of Chapter XV:

$$\widehat{G}^{(N)}(\mathbf{r}_0,\mathbf{r}_0',\mathbf{r}_0'',\mathbf{r}_0''') = \sum_{q,q'=1;\ q\neq q'}^{N} \widehat{g}(q:\mathbf{r}_0,\mathbf{r}_0';q':\mathbf{r}_0'',\mathbf{r}_0''') \tag{B-29}$$

and yields in the Fock space the operator $\widehat{G}(\mathbf{r}_0,\mathbf{r}_0',\mathbf{r}_0'',\mathbf{r}_0''')$. Relation (B-2), without this factor 1/2, then leads to:

$$\widehat{G}(\mathbf{r}_0,\mathbf{r}_0',\mathbf{r}_0'',\mathbf{r}_0''') = \Psi^\dagger(\mathbf{r}_0)\Psi^\dagger(\mathbf{r}_0'')\Psi(\mathbf{r}_0''')\Psi(\mathbf{r}_0') \tag{B-30}$$

In this case, the operator annihilates two particles at two points and recreates them at two others.

A computation very similar to the one leading to (B-22) and (B-23) enables us to show, using (B-2), that the matrix elements of the two-particle density operator ρ_{II} can be written[3]:

$$\langle 1:\mathbf{r}_0';2:\mathbf{r}_0'''|\ \rho_{II}\ |1:\mathbf{r}_0;2:\mathbf{r}_0''\rangle = \left\langle \Psi^\dagger(\mathbf{r}_0)\Psi^\dagger(\mathbf{r}_0'')\Psi(\mathbf{r}_0''')\Psi(\mathbf{r}_0') \right\rangle \tag{B-31}$$

This density operator, whose trace is equal to $N(N-1)$, plays an essential role in the study of correlations between particles.

A particularly important example of a higher order correlation function corresponds to the case where $\mathbf{r}_0 = \mathbf{r}_0'$ and $\mathbf{r}_0'' = \mathbf{r}_0'''$. We then get:

$$\begin{aligned}\widehat{G}^{(N)}(\mathbf{r}_0,\mathbf{r}_0,\mathbf{r}_0'',\mathbf{r}_0'') &= \sum_{q,q'=1;\ q\neq q'}^{N} |q:\mathbf{r}_0\rangle\langle q:\mathbf{r}_0| \otimes |q':\mathbf{r}_0''\rangle\langle q':\mathbf{r}_0''| \\ &= \sum_{q,q'=1;\ q\neq q'}^{N} |q:\mathbf{r}_0;\ q':\mathbf{r}_0''\rangle\langle q:\mathbf{r}_0;\ q':\mathbf{r}_0''|\end{aligned} \tag{B-32}$$

[3] One can also use relation (C-19) of Chapter XV to get the same result.

which yields in the Fock space the operator:

$$\widehat{G}(\mathbf{r}_0, \mathbf{r}_0, \mathbf{r}_0'', \mathbf{r}_0'') = \Psi^\dagger(\mathbf{r}_0)\Psi^\dagger(\mathbf{r}_0'')\Psi(\mathbf{r}_0'')\Psi(\mathbf{r}_0) \tag{B-33}$$

The expression on the right-hand side of (B-32) characterizes the probability of finding any particle at $\mathbf{r}_0$ and any other one at $\mathbf{r}_0''$. In the same way as the average value (B-9) gives the one-particle density, the average value:

$$\left\langle \widehat{G}(\mathbf{r}_0, \mathbf{r}_0, \mathbf{r}_0'', \mathbf{r}_0'') \right\rangle = \langle \Phi | \, \widehat{G}(\mathbf{r}_0, \mathbf{r}_0, \mathbf{r}_0'', \mathbf{r}_0'') \, | \Phi \rangle = G_2(\mathbf{r}_0, \mathbf{r}_0'') \tag{B-34}$$

gives the two-particle "double density", which contains information on all the binary correlations between the particle positions.

We are now in a position to again obtain expression (C-28) of Chapter XV, and more precisely justify the interpretation we gave of the average value of the interaction energy written in (C-27) of that chapter. We replace in (B-33) the field operators (or their adjoints) by their expansion (A-3) on the operators a_i (or the $a_i^\dagger$); we then get (C-28) of Chapter XV, $\mathbf{r}_0$ being replaced by $\mathbf{r}_1$ and $\mathbf{r}_0'$ by $\mathbf{r}_2$.

If $\mathbf{r}_0 \neq \mathbf{r}_0''$, we can check[4] that this operator is equal to the product of the simple densities defined in (B-8):

$$\widehat{G}(\mathbf{r}_0, \mathbf{r}_0, \mathbf{r}_0'', \mathbf{r}_0'') = \widehat{D}(\mathbf{r}_0) \; \widehat{D}(\mathbf{r}_0'') \tag{B-35}$$

Obviously this relation between operators does not mean that the double density $G_2(\mathbf{r}_0, \mathbf{r}_0'')$ is merely the product $\left\langle \widehat{D}(\mathbf{r}_0) \right\rangle \left\langle \widehat{D}(\mathbf{r}_0'') \right\rangle$ of the simple densities: the average value of a product of operators is not, in general, equal to the product of the average values. When we studied the function G_2 (§ C-5-b in Chapter XV), we did find the presence of an exchange term that introduces "statistical correlations" between particles, even in the absence of interactions.

Particles with spin:

For particles with non-zero spin, we just have to add an index ν to each of the kets or bras, as well as to the field operators; this brings up to $(2S+1)^4$ the number of 4-point correlation functions. The matrix elements of the two-body density operator are then given by the average values:

$$\left\langle 1 : \mathbf{r}', \nu'; 2 : \mathbf{r}''', \nu''' \right| \rho_{II} \left| 1 : \mathbf{r}, \nu; 2 : \mathbf{r}'', \nu'' \right\rangle = \left\langle \Psi_\nu^\dagger(\mathbf{r}) \Psi_{\nu''}^\dagger(\mathbf{r}'') \Psi_{\nu'''}(\mathbf{r}''') \Psi_{\nu'}(\mathbf{r}') \right\rangle \tag{B-36}$$

B-4. Hamiltonian operator

We now establish the expression, in terms of the field operator, of the Hamiltonian operator for a system of identical (spinless) particles. Two formulas will be useful for this computation. The first one transposes to three dimensions the formula (34) of Appendix II:

$$\delta(\mathbf{r} - \mathbf{r}') = \frac{1}{(2\pi)^3} \int \mathrm{d}^3 k \; e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \tag{B-37}$$

[4] If the particles are bosons, we just permute the commuting operators to bring $\Psi(\mathbf{r}_0)$ to the second position and obtain the result. If we are dealing with fermions, two successive anticommutations are necessary to get that result, and the corresponding two minus signs cancel each other.

The second one is obtained from (B-37) by a double derivation with respect to $\mathbf{r}$:

$$\Delta\delta(\mathbf{r}-\mathbf{r}') = -\frac{1}{(2\pi)^3}\int \mathrm{d}^3k\ k^2\ e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')} \tag{B-38}$$

The matrix elements of a single particle's kinetic energy are written as:

$$\langle\mathbf{r}|\,\frac{P^2}{2m}\,|\mathbf{r}'\rangle = \int \mathrm{d}^3k\ \langle\mathbf{r}|\,\mathbf{k}\rangle\,\frac{\hbar^2k^2}{2m}\,\langle\mathbf{k}\,|\mathbf{r}'\rangle = -\frac{\hbar^2}{2m}\Delta\delta(\mathbf{r}-\mathbf{r}') \tag{B-39}$$

In the Fock space, it corresponds to the following operators (an integration by parts[5] was used to go from the first to the second relation):

$$\widehat{H}_0 = -\frac{\hbar^2}{2m}\int \mathrm{d}^3r\ \Psi^\dagger(\mathbf{r})\Delta\Psi(\mathbf{r}) = \frac{\hbar^2}{2m}\int \mathrm{d}^3r\ \boldsymbol{\nabla}\Psi^\dagger(\mathbf{r})\cdot\boldsymbol{\nabla}\Psi(\mathbf{r}) \tag{B-40}$$

As in § B-1, we obtain here an expression similar to the average value of an operator (here the kinetic energy) for one particle; but the gradient of the wave function must be replaced by that of a field operator, and the order of the operators can matter.

The system Hamiltonian includes in general an interaction term, which makes it a two-particle operator and requires using formula (B-2). For a two-particle system, we know that the interaction yields an operator that is diagonal in the $\{|\mathbf{r},\mathbf{r}'\rangle\}$ representation; furthermore, it only depends on the relative position $\mathbf{r}-\mathbf{r}'$ (and not on $\mathbf{r}$ and $\mathbf{r}'$ separately). Consequently, the matrix element in (B-2) takes the form:

$$\langle\mathbf{r},\mathbf{r}'|\,g\,|\mathbf{r}'',\mathbf{r}'''\rangle = \delta(\mathbf{r}-\mathbf{r}'')\ \delta(\mathbf{r}'-\mathbf{r}''')\times W_{2.}(\mathbf{r}-\mathbf{r}') \tag{B-41}$$

where $W_2(\mathbf{r}-\mathbf{r}')$ is the interaction potential energy between two particles located at a relative position $\mathbf{r}-\mathbf{r}'$ (this interaction is often isotropic, in which case W_2 only depends on the relative distance $|\mathbf{r}-\mathbf{r}'|$). Finally, starting from (B-2), we get the following expression for the Hamiltonian operator:

$$\begin{aligned}\widehat{H} = \int \mathrm{d}^3r\ &\left[\frac{\hbar^2}{2m}\boldsymbol{\nabla}\Psi^\dagger(\mathbf{r})\cdot\boldsymbol{\nabla}\Psi(\mathbf{r})\ + V_1(\mathbf{r})\Psi^\dagger(\mathbf{r})\Psi(\mathbf{r})\right]\\ &+\frac{1}{2}\int \mathrm{d}^3r'\int \mathrm{d}^3r''\ W_2(\mathbf{r}'-\mathbf{r}'')\ \Psi^\dagger(\mathbf{r}')\Psi^\dagger(\mathbf{r}'')\Psi(\mathbf{r}'')\Psi(\mathbf{r}')\end{aligned} \tag{B-42}$$

The first term corresponds to the particles' kinetic energy, the second to the external potential $V_1(\mathbf{r})$ acting separately on each particle, and the third one to the mutual interaction between particles; note that this last term involves four field operators, whereas the first two involve only two. The same comment as above still applies: this expression is reminiscent of the average energy of a system of two particles, both described by the same wave function; but now we are dealing with operators that do not commute.

This Hamiltonian can also be expressed directly via the one-particle simple density $\widehat{D}(\mathbf{r}_0,\mathbf{r}_0')$ and the two-particle double density $\widehat{G}(\mathbf{r}_0,\mathbf{r}_0,\mathbf{r}_0'',\mathbf{r}_0'')$ operators, as we now show. Inserting relations (B-21) and (B-30) in expression (B-42), we obtain:

$$\begin{aligned}\widehat{H} = \int \mathrm{d}^3r\ &\left[\frac{\hbar^2}{2m}\boldsymbol{\nabla}_{\mathbf{r}_0}\cdot\boldsymbol{\nabla}_{\mathbf{r}_0'}\ \widehat{D}(\mathbf{r}_0,\mathbf{r}_0')\Big|_{\mathbf{r}_0=\mathbf{r}_0'=\mathbf{r}} + V_1(\mathbf{r})\widehat{D}(\mathbf{r},\mathbf{r})\right]\\ &+\frac{1}{2}\int \mathrm{d}^3r'\int d^3r''\ W_2(\mathbf{r}'-\mathbf{r}'')\ G_2(\mathbf{r}',\mathbf{r}',\mathbf{r}'',\mathbf{r}'')\end{aligned} \tag{B-43}$$

[5] The value of the already integrated terms must be taken at infinity, and we assume that all the states of the physical system are limited to a finite volume; as a result, those terms do not play any role and can be ignored.

where the notation $\boldsymbol{\nabla}_{\mathbf{r}_0}$ and $\boldsymbol{\nabla}_{\mathbf{r}'_0}$ represents the gradient taken with respect to the variables $\mathbf{r}_0$ (for the first one), and $\mathbf{r}'_0$ (for the second); once these gradients have been computed in the kinetic energy term of (B-43), both variables take on the same value $\mathbf{r}$. The fact that the Hamiltonian operator can be directly expressed in terms of the simple and double density operators can be useful. For example, to determine the ground state energy of an N-particle system, we do not have to compute the state wave function, which involves all the correlations of order 1 to N between the particles; it is sufficient to know the average values of these two densities. There exist, in certain cases, approximation methods that yield directly good estimates of these simple and double densities, hence allowing an access to the N-body energy. Complements $\mathrm{E_{XV}}$ and $\mathrm{G_{XV}}$ discuss the Hartree-Fock method, which is based on an approximation where the two-particle density operator is simply expressed as a function of the one-particle density operator, i.e. the double density as a function of the simple density (Complement $\mathrm{G_{XV}}$, § 2-b-γ); this allows convenient mean field calculations.

C. Time evolution of the field operator (Heisenberg picture)

The operators we have considered until now correspond to the "Schrödinger picture", where the time evolution of the system is determined by the time evolution of its state vector. It may, however, be more convenient to adopt the Heisenberg picture (Complement $\mathrm{G_{III}}$), where this time evolution is transferred to the operators associated with the system's physical quantities. For spinless particles, let us call $\Psi_H(\mathbf{r};t)$ the operator corresponding, in the Heisenberg picture, to $\Psi(\mathbf{r})$:

$$\Psi_H(\mathbf{r};t) = e^{i\widehat{H}t/\hbar}\ \Psi(\mathbf{r})\ e^{-i\widehat{H}t/\hbar} \tag{C-1}$$

($\widehat{H}$ is the Hamiltonian operator), and whose time dependence follows the equation:

$$i\hbar\frac{\partial}{\partial t}\Psi_H(\mathbf{r};t) = \left[\Psi_H(\mathbf{r};t)\ ,\ \widehat{H}\right] \tag{C-2}$$

We are going to compute successively the commutator of $\Psi_H(\mathbf{r};t)$ with each of the three terms on the right-hand side of (B-42). The evolution equation for the field operator involves all the terms of the Hamiltonian (B-42): the kinetic, potential and interaction energies.

C-1. Contribution of the kinetic energy

In order to determine the commutator of the field operator with the kinetic energy, we first transpose the equations (A-17), (A-18) and (A-21) to the Heisenberg picture. Actually, they can be used without any changes: the unitary transform of a product by (C-1) is the product of the unitary transforms, that of the commutator (or of the anticommutator) is the commutator (or the anticommutator) of the transforms, and numbers like zero or the function $\delta(\mathbf{r}-\mathbf{r}')$ are invariant. Those three relations are therefore still valid in the Heisenberg picture, if we simply add an index H to the field operators. We now take their derivative with respect to the positions; only (A-21) yields a non-zero result:

$$\left[\Psi_H(\mathbf{r};t), \boldsymbol{\nabla}_{\mathbf{r}'}\Psi_H^\dagger(\mathbf{r}';t)\right]_{-\eta} = \boldsymbol{\nabla}_{\mathbf{r}'}\ \delta(\mathbf{r}-\mathbf{r}') = -\boldsymbol{\nabla}_{\mathbf{r}}\ \delta(\mathbf{r}-\mathbf{r}') \tag{C-3}$$

Taking (B-40) into account, we can write the commutator to be evaluated as:

$$\Big[\Psi_H(\mathbf{r};t), \widehat{H}_{0H}(t)\Big] = \frac{\hbar^2}{2m}\int \mathrm{d}^3r' \ \Big[\Psi_H(\mathbf{r};t), \boldsymbol{\nabla}\Psi_H^\dagger(\mathbf{r}';t)\cdot\boldsymbol{\nabla}\Psi_H(\mathbf{r}';t)\Big] \tag{C-4}$$

In the term to be integrated on the right-hand side, a sign η is introduced each time we permute two field operators, or two adjoints; when we permute a field operator and an adjoint, we must add to the result the right-hand side of (C-3). Adding and subtracting two equal terms, we then obtain for the function to be integrated:

$$\begin{array}{l}\Psi_H(\mathbf{r};t)\ \boldsymbol{\nabla}\Psi_H^\dagger(\mathbf{r}';t)\cdot\boldsymbol{\nabla}\Psi_H(\mathbf{r}';t) - \eta\ \boldsymbol{\nabla}\Psi_H^\dagger(\mathbf{r}';t)\cdot(\Psi_H(\mathbf{r};t)\boldsymbol{\nabla}\Psi_H(\mathbf{r}';t)) \\ +\eta\ \boldsymbol{\nabla}\Psi_H^\dagger(\mathbf{r}';t)\cdot(\Psi_H(\mathbf{r};t)\boldsymbol{\nabla}\Psi_H(\mathbf{r}';t)) - \boldsymbol{\nabla}\Psi_H^\dagger(\mathbf{r}';t)\cdot\boldsymbol{\nabla}\Psi_H(\mathbf{r}';t)\Psi_H(\mathbf{r};t)\end{array} \tag{C-5}$$

that is:

$$\begin{array}{l}\Big[\Psi_H(\mathbf{r};t), \boldsymbol{\nabla}\Psi_H^\dagger(\mathbf{r}';t)\Big]_{-\eta}\cdot\boldsymbol{\nabla}\Psi_H(\mathbf{r}';t) \\ \qquad +\eta\ \boldsymbol{\nabla}\Psi_H^\dagger(\mathbf{r}';t)\cdot[\Psi_H(\mathbf{r};t), \boldsymbol{\nabla}\Psi_H(\mathbf{r}';t)]_{-\eta} \\ \qquad\qquad = \boldsymbol{\nabla}_{\mathbf{r}'}\ \delta(\mathbf{r}-\mathbf{r}')\cdot\boldsymbol{\nabla}\Psi_H(\mathbf{r}',t)\end{array} \tag{C-6}$$

The integration over d^3r' then yields the Laplacian at $\mathbf{r}$ of the field operator, and we finally get:

$$\Big[\Psi_H(\mathbf{r};t),\ \widehat{H}_{0H}(t)\Big] = -\frac{\hbar^2}{2m}\Delta\Psi_H(\mathbf{r};t) \tag{C-7}$$

C-2. Contribution of the potential energy

Instead of (C-4), it is now the commutator:

$$\Big[\Psi_H(\mathbf{r};t),\ \widehat{V}_{1H}(t)\Big] = \int \mathrm{d}^3r'\ V_1(\mathbf{r}')\ \Big[\Psi_H(\mathbf{r};t),\ \Psi_H^\dagger(\mathbf{r}';t)\Psi_H(\mathbf{r}';t)\Big] \tag{C-8}$$

which comes into play. The calculation is similar to the previous one, but without the gradients which were applied to the field operators depending on $\mathbf{r}'$. The right-hand side of (C-6) now becomes simply:

$$\delta(\mathbf{r}-\mathbf{r}')\Psi_H(\mathbf{r}') \tag{C-9}$$

and the integration over d^3r' is straightforward, so that:

$$\Big[\Psi_H(\mathbf{r};t),\ \widehat{V}_{1H}(t)\Big] = V_1(\mathbf{r})\ \Psi_H(\mathbf{r};t) \tag{C-10}$$

C-3. Contribution of the interaction energy

It is now the commutator of $\Psi_H(\mathbf{r};t)$ with a product of four field operators that will have to be integrated:

$$\Big[\Psi_H(\mathbf{r};t),\ \Psi_H^\dagger(\mathbf{r}';t)\Psi_H^\dagger(\mathbf{r}'';t)\Psi_H(\mathbf{r}'';t)\Psi_H(\mathbf{r}';t)\Big] \tag{C-11}$$

We will not go through the details of the calculation, a bit long but without any real difficulties; as in the previous two cases, it involves the repeated application of the commutation relations. The result is that the commutator of the field with the interaction energy can be written as:

$$\int \mathrm{d}^3 r' \; W_2(\mathbf{r}-\mathbf{r}')\Psi_H^\dagger(\mathbf{r}';t)\Psi_H(\mathbf{r}';t)\Psi_H(\mathbf{r};t) \tag{C-12}$$

C-4. Global evolution

Regrouping the three previous terms, we get the evolution equation for the field operator:

$$\left[i\hbar\frac{\partial}{\partial t}+\frac{\hbar^2}{2m}\Delta - V_1(\mathbf{r})\right]\Psi_H(\mathbf{r};t) = \int \mathrm{d}^3 r' \; W_2(\mathbf{r}-\mathbf{r}') \; \Psi_H^\dagger(\mathbf{r}';t)\Psi_H(\mathbf{r}';t)\Psi_H(\mathbf{r};t) \tag{C-13}$$

The left hand-side includes the differential operator of the usual Schrödinger equation for a single particle in a potential $V_1(\mathbf{r})$; however, as already pointed out, Ψ is not a simple function here, but an operator. The right-hand side includes the binary interaction effects; its presence implies that the evolution equation of the field operator is not "closed". Its evolution depends not only on the operator itself, but also on a term containing the product of three fields (or their conjugates).

Analyzing in a similar way the evolution of such a product of three factors, we see that it depends on that product, and also on the product of 5 fields (or their conjugates); in turn, the evolution of a product of 5 fields will involve 7 others, etc. We thus get a series of more and more complex equations, often called a "hierarchy"of equations. They are in general very hard to solve exactly. This is why it is frequent to use an approximation by truncating this hierarchy at a certain stage, or else by eliminating the coupling term at a certain order, or replacing it by a more convenient expression. Many different methods have been proposed to accomplish this, the most well-known being the mean field approximation (Complements $\mathrm{E_{XV}}$ and $\mathrm{G_{XV}}$).

D. Relation to field quantization

In conclusion, we make some remarks concerning the field quantization procedures, and their relation to the identical particle concept. Consider a single spinless particle in an external potential well, and call $\varphi_i(\mathbf{r})$ the wave functions associated with its stationary states in that potential (the subscript i runs from 1 to infinity; we assume, for the sake of simplicity, that the spectrum is entirely discrete). The $\varphi_i(\mathbf{r})$ form a basis on which we can expand any particle wave function $\psi(\mathbf{r})$:

$$\psi(\mathbf{r}) = \sum_i c_i \varphi_i(\mathbf{r}) \tag{D-1}$$

We already noted the similarity between this formula and equality (A-3), where the only difference is that the numbers c_i are replaced by the operators a_i. Along the same line, relations (B-8) and (B-15) are reminiscent of a particle's probability density of presence, and of its probability current. Finally, expression (B-42) is very similar to the average

value of the energy of a system of two particles, both placed in the same state described by the wave function $\psi(\mathbf{r})$, once we replace that usual wave function by an operator depending on the parameter $\mathbf{r}$. Consequently, the creation and annihilation operator method has an air of "second quantization": we start with the quantum wave function for one (or two) particle(s) ("first quantization"), and in a second stage, we replace the wave functions coefficients by operators ("second quantization"). However, we must keep in mind that we do not, in reality, quantize the same physical system twice; the main difference comes from the fact that we go from a very small number of particles, one or two, to a very large number of identical particles.

Field operators can also appear when quantizing a classical field, such as the electromagnetic field. This is the object of Chapters XIX and XX, where we will show how the concept of a photon emerges, as the elementary excitation of the electromagnetic field. We shall also see how the electric and magnetic fields, which were classical functions, become operators defined at each point in space, creating and annihilating photons.

Generally speaking, a system consisting of an ensemble of identical bosons and a system obtained by quantizing a classical field obey exactly the same equations. The particles of the first system play the role of field quanta for the second system, and the field operators then satisfy commutation relations. The two physical systems are therefore perfectly equivalent. In the case of the electromagnetic field, the particles in questions are the photons and they have a zero mass. However, this is not necessarily the case for all fields. Moreover, quantum fields associated with a system of identical fermions also exist. These fields do not have a direct classical correspondence and their operators obey anticommutation relations. In particle physics, one simultaneously takes into account fermonic and bosonic fields, associated in general with non-zero mass particles.

COMPLEMENTS OF CHAPTER XVI, READER'S GUIDE

A_{XVI} : SPATIAL CORRELATIONS IN A BOSON OR FERMION IDEAL GAS	In this complement we study the properties of the spatial correlation functions in systems of fermions or bosons. For fermions, we establish the existence of an "exchange hole" which corresponds to the impossibility for fermions with parallel spins to be found at the same point in space. For bosons, we discuss their tendency to *bunch* (group together). *Recommended in a first reading*
B_{XVI} : SPATIO-TEMPORAL COORELATION FUNCTIONS, GREEN'S FUNCTIONS	Green's functions are a very general tool for the theoretical study of N-body systems. In this complement, they are first introduced in ordinary space, then in reciprocal space (the Fourier momentum space). Knowledge of these functions allows calculation of numerous physical properties of the system. *Slightly more difficult than the previous complement*
C_{XVI} : WICK'S THEOREM	Wick's theorem permits calculating average values of any product of creation and annihilation operators, for an ideal gas system in thermal equilibrium. The calculation involves a very useful concept, the operator "contraction".

Complement A_{XVI}

Spatial correlations in an ideal gas of bosons or fermions

1 System in a Fock state 1769
1-a Two-point correlations 1770
1-b Four-point correlations 1771
2 Fermions in the ground state 1771
2-a Two-point correlations 1772
2-b Correlations between two particles 1774
3 Bosons in a Fock state 1775
3-a Ground state 1775
3-b Fragmented state 1776
3-c Other states 1777

In this complement, we establish a certain number of properties of the correlation functions, arising solely from the particle statistics (i.e. from the fact they are either bosons or fermions), and independent of their possible interactions. To keep the calculations simple, we assume that the N-particle system is described by a Fock state, characterized by the occupation numbers n_i of each individual state $|u_i\rangle$. We shall see that fermions and bosons behave very differently: whereas the latter tend to bunch, the former tend to avoid each other, as indicated by the existence of an "exchange hole".

We give in § 1 the general expression of these correlation functions, without making any hypothesis concerning the nature of the individual states; the physical system is not necessarily homogeneous in space. In §§ 2 and 3, we study successively bosons and fermions, assuming the physical system to be contained in a box of volume $\mathcal{V}$ in which the particles are free. The periodic boundary conditions (Complement C_{XIV}) allow taking into account the confinement while maintaining translation invariance (the system is perfectly homogeneous in space), which makes the calculations easier. For spinless particles, the individual states $|u_i\rangle$ correspond to plane waves normalized in the volume $\mathcal{V}$:

$$u_i(\mathbf{r}) = \frac{1}{\sqrt{\mathcal{V}}} e^{i\mathbf{k}_i \cdot \mathbf{r}} \tag{1}$$

where the $\mathbf{k}_i$ are chosen to satisfy the periodic boundary conditions.

1. System in a Fock state

We assume the state $|\Phi\rangle$ of the N-particle system to be a Fock state built from the basis of individual states $\{|u_i\rangle\}$, with occupation numbers n_i (not greater than 1 for fermions):

$$|\Phi\rangle = |n_1 : u_1;\ n_2 : u_2;\ ..;\ n_i : u_i;\ ..\rangle \tag{2}$$

This will be the case, for example, if the particles do not interact (ideal gas) and if the system is in a stationary state, such as its ground state.

1-a. Two-point correlations

For spinless particles, relations (A-3) and (B-21) of Chapter XVI lead to:

$$\widehat{D}(\mathbf{r},\mathbf{r}') = \sum_{i,j} u_i^*(\mathbf{r}) u_j(\mathbf{r}')\ a_i^\dagger a_j \tag{3}$$

Now the average value in a Fock state of the operator product $a_i^\dagger a_j$ is zero if $i \neq j$: the successive action of the two operators leads to another Fock state with the same particle number, but with two different occupation numbers – therefore to an orthogonal state. If, on the other hand, $i = j$, that product becomes the particle number operator $\widehat{n}_i$ that now acts on its eigenket, with eigenvalue n_i. We thus get:

$$\langle\Phi|\ a_i^\dagger a_j\ |\Phi\rangle = \delta_{ij}\ n_i \tag{4}$$

(where δ_{ij} is the Kronecker delta); this yields:

$$G_1(\mathbf{r},\mathbf{r}') = \left\langle \widehat{D}(\mathbf{r},\mathbf{r}') \right\rangle = \sum_i n_i\ u_i^*(\mathbf{r}) u_i(\mathbf{r}') \tag{5}$$

The physical interpretation of this result is the following: for a single particle in the individual state $|u_i\rangle$, the function G_1 would simply be $u_i^*(\mathbf{r})u_i(\mathbf{r}')$; for an N-particle system in a Fock state, each individual state gives a contribution, multiplied by a coefficient equal to its population.

Particles with non-zero spin:

For particles with spin, we define the $u_i^\nu(\mathbf{r})$ by:

$$u_i^\nu(\mathbf{r}) = \langle \mathbf{r}, \nu\ |u_i\rangle \tag{6}$$

When we add a discrete spin index ν to $\mathbf{r}$, formula (3) becomes:

$$\widehat{D}(\mathbf{r},\nu;\ \mathbf{r}',\nu') = \sum_{i,j} [u_i^\nu(\mathbf{r})]^*\ u_j^{\nu'}(\mathbf{r}')\ a_i^\dagger a_j \tag{7}$$

We obtain, with the same reasoning:

$$G_1(\mathbf{r},\nu;\mathbf{r}',\nu') = \left\langle \widehat{D}(\mathbf{r},\nu;\ \mathbf{r}',\nu') \right\rangle = \sum_i n_i\ \ [u_i^\nu(\mathbf{r})]^*\ u_i^{\nu'}(\mathbf{r}') \tag{8}$$

whose physical interpretation is similar to the previous one.

One often chooses a basis of individual states $|u_i\rangle$, such that each ket corresponds to a well defined value of the spin: each index i indicates both an individual orbital state and a value of ν (which then becomes a function ν_i of i). In that case, for a given i, the wave function $u_i^\nu(\mathbf{r})$ is only defined for a single value of the index ν; conversely, for a given ν, the wave functions are different from zero only if the index i (or j) belongs to a certain domain $D(\nu)$. In expression (8), ν and ν' are fixed, and the index i must necessarily belong to both $D(\nu)$ and $D(\nu')$, or else the result is zero. This leads to:

$$G_1(\mathbf{r},\nu;\mathbf{r}',\nu') = \delta_{\nu\nu'} \sum_{i\in D(\nu)} n_i\ \ [u_i^\nu(\mathbf{r})]^*\ u_i^{\nu'}(\mathbf{r}') \tag{9}$$

This correlation function is therefore zero if $\nu' \neq \nu$.

1-b. Four-point correlations

We limit ourselves to the calculation of $\left\langle \widehat{G}(\mathbf{r},\mathbf{r}',\mathbf{r}'',\mathbf{r}''') \right\rangle$ in the "diagonal case" $\mathbf{r}' = \mathbf{r}$ and $\mathbf{r}''' = \mathbf{r}''$; we call $\mathbf{r}_1$ the common value of $\mathbf{r}$ and $\mathbf{r}'$, and $\mathbf{r}_2$ the common value of $\mathbf{r}''$ and $\mathbf{r}'''$. Such a diagonal correlation function was already written in (B-20) of Chapter XVI and in § C-5-a of Chapter XV; this is the only one that plays a role in the particle interaction energy, as the associated operator is also diagonal in the position representation.

In the absence of spin, and for a Fock state, the calculation of the correlation function $G_2(\mathbf{r}_1,\mathbf{r}_2)$ was carried out in § C-5-b of Chapter XV, where relations (C-32) to (C-34) yield the value of this function:

$$\begin{aligned} G_2(\mathbf{r}_1,\mathbf{r}_2) = &\sum_i n_i\,(n_i-1)\ \ |u_i(\mathbf{r}_1)|^2\,|u_i(\mathbf{r}_2)|^2\ + \\ &+\sum_{i\neq j} n_i n_j\left[|u_i(\mathbf{r}_1)|^2\,|u_j(\mathbf{r}_2)|^2 + \eta\ u_i^*(\mathbf{r}_1)u_i(\mathbf{r}_2)u_j^*(\mathbf{r}_2)u_j(\mathbf{r}_1)\right] \end{aligned} \tag{10}$$

where $\eta = +1$ for bosons, $\eta = -1$ for fermions. The second line of this equation contains a "direct term" only involving the moduli squared of the wave functions; it also contains an "exchange term" where the phase of the wave functions come into play, and which changes sign depending on whether we are dealing with a system of bosons or fermions.

Particles with non-zero spin:

In the presence of spins, we must add, as previously, the corresponding spin indices ν, and the correlation function becomes:

$$\begin{aligned} &G_2(\mathbf{r}_1,\nu_1;\mathbf{r}_2,\nu_2) = \textstyle\sum_i n_i\,(n_i-1)\ \ |u_i^{\nu_1}(\mathbf{r}_1)|^2\,|u_i^{\nu_2}(\mathbf{r}_2)|^2\ + \\ &+\textstyle\sum_{i\neq j} n_i n_j\left[\left|u_i^{\nu_1}(\mathbf{r}_1)\right|^2\left|u_j^{\nu_2}(\mathbf{r}_2)\right|^2 + \eta\ \left[u_i^{\nu_1}(\mathbf{r}_1)\right]^* u_i^{\nu_2}(\mathbf{r}_2)\left[u_j^{\nu_2}(\mathbf{r}_2)\right]^* u_j^{\nu_1}(\mathbf{r}_1)\right] \end{aligned} \tag{11}$$

If we choose, as in (9), a basis of individual states $|u_i\rangle$ in which each ket has a well defined value of the spin, the summation is simpler and we get:

$$\begin{aligned} G_2(\mathbf{r}_1,\nu_1;\mathbf{r}_2,\nu_2) = &\,\delta_{\nu_1\nu_2}\textstyle\sum_{i\in D(\nu_1)} n_i\,(n_i-1)\ \ |u_i^{\nu_1}(\mathbf{r}_1)|^2\,|u_i^{\nu_2}(\mathbf{r}_2)|^2\ + \\ &+\textstyle\sum_{i\in D(\nu_1),\ j\in D(\nu_2);\ i\neq j} n_i n_j\left[\left|u_i^{\nu_1}(\mathbf{r}_1)\right|^2\left|u_j^{\nu_2}(\mathbf{r}_2\right|^2 + \right. \\ &\left. +\eta\delta_{\nu_1\nu_2}\left[u_i^{\nu_1}(\mathbf{r}_1)\right]^* u_i^{\nu_2}(\mathbf{r}_2)\left[u_j^{\nu_2}(\mathbf{r}_2)\right]^* u_j^{\nu_1}(\mathbf{r}_1)\right] \end{aligned} \tag{12}$$

As above, $D(\nu)$ corresponds to the domain of the index i for the wave function $u_i^\nu(\mathbf{r})$ to exist (otherwise, it is not defined). If the spin states are different ($\nu_1 \neq \nu_2$), the only contribution to the correlation function comes from the second line (the direct term). The exchange term which follows only concerns particles being in the same spin state; it changes sign depending on whether they are bosons or fermions. No exchange term exists for particles having orthogonal spins. This comes from the fact that to behave as strictly identical objects, two particles must occupy the same spin state; otherwise, their spin direction could, at least in principle, be used to distinguish them.

2. Fermions in the ground state

Consider an ideal gas of fermions, contained in a volume $\mathcal{V}$, and having a spin S equal to 1/2 (as for electrons); the index ν can only take on two values $\pm 1/2$. If we assume

the gas to be in its ground state, this corresponds, for an ideal gas, to a Fock state: for each of the two spin states, all the individual states with an energy lower than a certain value (called the Fermi energy) have an occupation number equal to 1, all the others being empty. We shall proceed as in Complement C_{XIV} (in particular for the study of the magnetic susceptibility) and attribute to each of the two spin states a different Fermi energy – this is useful to account for an average spin orientation (under the influence of a magnetic field for example). For all the values of the index i corresponding to $\nu = +1/2$, we thus assume that the occupation numbers are equal to 1 if they correspond to plane waves (1) having a wave vector smaller than the Fermi vector k_F^+, and zero otherwise. This Fermi vector is linked to the Fermi energy by the relation:

$$E_F^+ = \frac{\hbar^2 \left(k_F^+\right)^2}{2m} \tag{13}$$

where m is the mass of each particle. In a similar way, for all the spin values $\nu = -1/2$, we assume that only the states with wave vectors smaller than the Fermi vector k_F^- are occupied, with a relation similar to (13) in which the index + is replaced by −. The total particle number $N_\pm$ in each of the two spin states are then:

$$N_\pm = \sum_{|\mathbf{k}_i| \leq k_F^\pm} 1 \tag{14}$$

(the summation runs over all the states having a population equal to 1). In the limit of large volumes $\mathcal{V}$, this expression becomes an integral:

$$N_\pm = \frac{\mathcal{V}}{(2\pi)^3} \int_{|\mathbf{k}_i| \leq k_F^\pm} \mathrm{d}^3 k = \frac{\mathcal{V}}{2\pi^2} \int_0^{k_F^\pm} k^2 \mathrm{d}k = \frac{\left(k_F^\pm\right)^3}{6\pi^2} \mathcal{V} \tag{15}$$

Depending on whether k_F^+ is larger or smaller than k_F^-, the spins + or − will make up the majority, the populations being equal if $k_F^+ = k_F^-$.

2-a. Two-point correlations

Let us compute the average value of the operator $\widehat{D}_{(\mathbf{r},\nu;\ \mathbf{r}',\nu')}$ defined in (3), distinguishing the two cases where ν and ν' are equal or different.

(i) Same spin states

Taking (1) into account, relation (9) then yields:

$$G_1(\mathbf{r},\pm;\mathbf{r}',\pm) = \left\langle \widehat{D}(\mathbf{r},\pm;\ \mathbf{r}',\pm) \right\rangle = \frac{1}{\mathcal{V}} \sum_{|\mathbf{k}_i| \leq k_F^\pm} e^{i\mathbf{k}_i \cdot (\mathbf{r}' - \mathbf{r})} \tag{16}$$

where the notation for the spins is simplified from $\pm 1/2$ to $\pm$. In the limit of large volumes, the summation over $\mathbf{k}_i$ becomes an integral, and we get:

$$G_1(\mathbf{r},\pm;\mathbf{r}',\pm) = \frac{1}{(2\pi)^3} \int_{|\mathbf{k}_i| \leq k_F^\pm} \mathrm{d}^3 k\ e^{i\mathbf{k} \cdot (\mathbf{r}' - \mathbf{r})} \tag{17}$$

For $\mathbf{r} = \mathbf{r}'$, this function simply yields the particle density $N_\pm/\mathcal{V}$, already computed. For $\mathbf{r} \neq \mathbf{r}'$, the function to be computed is the Fourier transform of a function of $\mathbf{k}$ that only

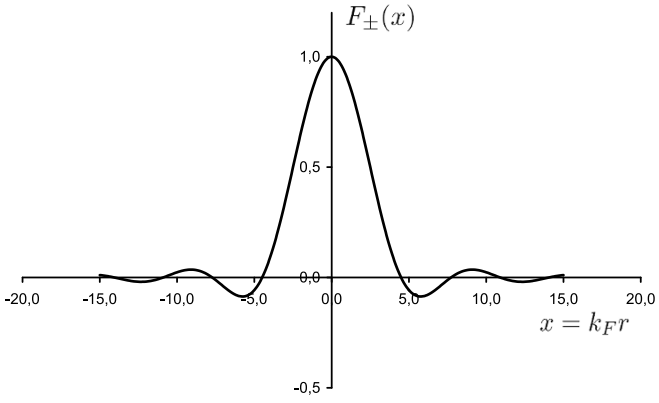


Figure 1: Plot of the function $F_{\pm}(x)$ as a function of the dimensionless variable $x = k_F r$.

depends on its modulus. Using relation (59) of Appendix I in volume II, we can write:

$$G_1(\mathbf{r}, \pm; \mathbf{r}', \pm) = \frac{N_{\pm}}{\mathcal{V}} \frac{F_{\pm}(|\mathbf{r}' - \mathbf{r}|)}{F_{\pm}(\mathbf{0})} \tag{18}$$

with[1]:

$$\begin{aligned} F_{\pm}(r) &= \frac{3}{\left(k_F^{\pm}\right)^3 r} \int_0^{k_F^{\pm}} k \mathrm{d}k \ \sin kr = \frac{3}{\left(k_F^{\pm}\right)^3 r} \left\{ -\frac{k}{r} \cos kr + \frac{1}{r^2} \sin kr \right\}_0^{k_F^{\pm}} \\ &= \frac{3}{\left(k_F^{\pm} r\right)^3} \left[\sin k_F^{\pm} r - \left(k_F^{\pm} r\right) \cos k_F^{\pm} r \right] \end{aligned} \tag{19}$$

Finally, we get:

$$G_1(\mathbf{r}, \pm; \mathbf{r}', \pm) = \frac{N_{\pm}}{\mathcal{V}} \ F_{\pm}(|\mathbf{r}' - \mathbf{r}|) \tag{20}$$

Figure 1 is a plot of $F_{\pm}(r)$ as a function of the mutual distance r between the two points. It shows that, for each spin state, the "non-diagonal" one-particle correlation function presents a maximum at $\mathbf{r} = \mathbf{r}'$, then rapidly decreases to zero over a distance of the order of a few Fermi wavelengths, $\lambda_F^{\pm} = 2\pi / k_F^{\pm}$. A system of free fermions, in its ground state, does not show any long-range "non-diagonal order".

(ii) Opposed spin states.

Relation (9) shows that the two-point correlation function is zero between two states of different spins; there is no non-diagonal order.

[1] An arbitrary coefficient $3/\left(k_F^{\pm}\right)^3$ has been introduced in the function F to make it tend towards 1 when its variable tends towards zero, which allows dropping the factor $F_{\pm}(\mathbf{0})$.

2-b. Correlations between two particles

We start from relation (12). In the second line, the condition $i \neq j$ may be ignored as, for fermions, the $i = j$ terms exactly cancel those on the third line; we can therefore consider the indices i and j as independent. Two cases must be distinguished:

(i) If $\nu_1 = \nu_2$, the three terms in (12) remain, but their behavior as a function of the volume $\mathcal{V}$ are different. This is because, in the limit of large volumes, each of the summations over i or j is proportional to the volume, whereas the moduli squared of the wave functions are each proportional to $1/\mathcal{V}$. For a large system, as the first of the three terms only contains a single sum over i, it varies as $1/\mathcal{V}$ and is thus negligible. We are left with the two other terms:

$$\begin{aligned} G_2\left(\mathbf{r},\pm;\mathbf{r}'',\pm\right) &= \left[\sum_{i\in D(\nu=\pm)} n_i \ \left|u_i^{\pm}(\mathbf{r})\right|^2\right]^2 - \left|\sum_{i\in D(\nu=\pm)} n_i \ \left[u_i^{\pm}(\mathbf{r})\right]^* u_i^{\pm}(\mathbf{r}'')\right|^2 \\ &= \left[\frac{N_\pm}{\mathcal{V}}\right]^2 - \left|\frac{1}{\mathcal{V}}\sum_{|\mathbf{k}_i|\leq k_F^{\pm}} e^{i\mathbf{k}_i\cdot\left(\mathbf{r}''-\mathbf{r}\right)}\right|^2 \end{aligned} \tag{21}$$

The same sum as (17) appears again in this relation. This leads to, in the limit of large volumes:

$$G_2\left(\mathbf{r},\pm;\mathbf{r}'',\pm\right) = \left[\frac{N_\pm}{\mathcal{V}}\right]^2 \left\{1 - \left[F_\pm(|\mathbf{r}''-\mathbf{r}|)\right]^2\right\} \tag{22}$$

The Pauli principle forbids particles in the same spin states to be at the same point in space; as expected, expression (22) goes to zero when $\mathbf{r} = \mathbf{r}''$. As the distance between particles increases, the function $F_\pm(|\mathbf{r}''-\mathbf{r}|)$ goes to zero, and the two-body correlation function tends towards the square of the one-body density $N_\pm/\mathcal{V}$, indicating that the long-range correlations disappear. This change of behavior occurs over a characteristic distance of the order of $\lambda_F^{\pm}$, comparable to the distance over which the non-diagonal order disappears. A plot of the spatial variations of the correlation function is given in Figure 2; it shows clearly the existence of an "exchange hole" corresponding to the mutual particle exclusion over this characteristic distance.

(ii) If $\nu_1 \neq \nu_2$, of the three terms of (12) only the second one (the direct term) is non-zero and yields a constant:

$$G_2\left(\mathbf{r},\pm;\mathbf{r}'',\mp\right) = \frac{N_+ N_-}{\mathcal{V}^2} \tag{23}$$

It is simply the product of the densities of the two kinds of spins; in the absence of interactions, particles with different spins do not show any correlation. This is because physically two particles at positions $\mathbf{r}$ and $\mathbf{r}''$ and in different spin states, can in principle be identified by the direction of their spin; consequently, they no longer behave as really indistinguishable quantum particles, and no Fermi statistical effects may be observed. As we assumed the particles did not interact with each other, no spatial correlations can develop.

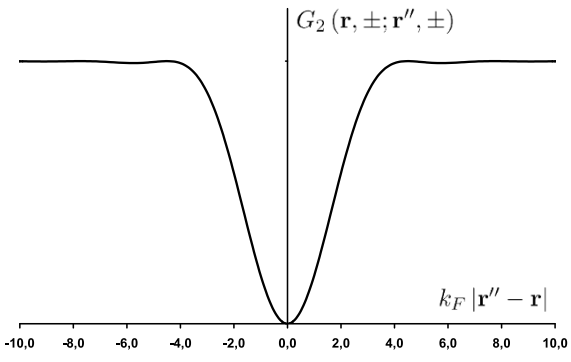


Figure 2: Plot, as a function of the dimensionless variable $k_F\,|\mathbf{r}'' - \mathbf{r}|$, of the correlation function $G_2(\mathbf{r}, \pm; \mathbf{r}'', \pm)$ between the positions $\mathbf{r}$ and $\mathbf{r}''$ of two particles in the same spin state, in a free fermion gas. As the Pauli principle forbids two particles to be at the same point in space, this function goes to zero at the origin, which creates an "exchange hole". As the distance increases, the function approaches 1 over a distance of the order of the inverse of the Fermi vector associated with this spin state.

Comment:

To keep the computation simple, we considered a system of non-interacting fermions, contained in a cubic box and in its ground state. The properties we discussed are, however, more general. In particular, it can be shown that a fermion system always exhibits an exchange hole for particles with identical spins, whether they interact or not; for a system at thermal equilibrium, the hole width gets smaller as the temperature increases, and goes from the Fermi wavelength at low temperature (degenerate system) to the thermal wavelength at high temperature (non-degenerate system).

3. Bosons in a Fock state

The situation is radically different for bosons, as there is no upper boundary for the occupation numbers n_i.

3-a. Ground state

For non-interacting spinless bosons in their ground state, the occupation number n_0 of the individual state $|u_0\rangle$ having the lowest energy is equal to the total particle number N, all the other occupation numbers being equal to zero. Relation (5) then yields:

$$G_1(\mathbf{r};\mathbf{r}') = \left\langle \widehat{D}_{(\mathbf{r},\mathbf{r}')} \right\rangle = N\; u_0^*(\mathbf{r}) u_0(\mathbf{r}') \tag{24}$$

As the wave function $u_0^*(\mathbf{r})$ extends over the entire volume $\mathcal{V}$, the modulus of this wave function does not decrease as the distance between $\mathbf{r}$ and $\mathbf{r}'$ increases and becomes comparable to the size of the system, as opposed to what occurs for fermions. This asymptotic behavior of $G_1(\mathbf{r};\mathbf{r}')$ has been used by Penrose and Onsager to define a general criterion for Bose-Einstein condensation, valid also for interacting bosons .

As for the two-particle correlation function, formula (10) yields:

$$G_2(\mathbf{r},\mathbf{r}'') = \left\langle \widehat{G}_{(\mathbf{r},\mathbf{r},\mathbf{r}'',\mathbf{r}'')} \right\rangle = N\,(N-1)\ \ |u_0(\mathbf{r})|^2\,|u_0(\mathbf{r}'')|^2 \tag{25}$$

If the ground state wave function is of the form $e^{-i\mathbf{k}_0\cdot\mathbf{r}}/\sqrt{\mathcal{V}}$, this function is simply equal to the constant $N\,(N-1)\,/\mathcal{V}^2$, independent of $\mathbf{r}$ and $\mathbf{r}'$. A system of bosons that are all in the same quantum state does not show any spatial correlations.

3-b. Fragmented state

We now assume the N-boson system to be in a "fragmented" state: instead of all the particles being in the same individual state, n_1 particles are in the state $|u_1\rangle$ and n_2 in the state $|u_2\rangle$, with $N = n_1 + n_2$. Relation (5) then yields:

$$G_1(\mathbf{r};\mathbf{r}') = n_1\ u_1^*(\mathbf{r})u_1(\mathbf{r}') + n_2\ u_2^*(\mathbf{r})u_2(\mathbf{r}') \tag{26}$$

When $\mathbf{r} = \mathbf{r}'$, expressions (24) and (26) contain the moduli squared of the wave functions $u_i(\mathbf{r}')$, which are all equal to $1/\mathcal{V}$ (for a system contained in a box of volume $\mathcal{V}$ with periodic boundary conditions); both expressions (24) and (26) are therefore equal. On the other hand, when $\mathbf{r}$ and $\mathbf{r}'$ are different, the phases of the two terms in (26) do not coincide any longer, and (destructive) interference effects can lower the modulus of $G_1(\mathbf{r};\mathbf{r}')$. Consequently, the fragmentation of a physical system into two states decreases the modulus of the non-diagonal terms of $G_1(\mathbf{r};\mathbf{r}')$. Obviously, the more fragmented states there are, the more noticeable the decrease.

Relation (10) now becomes:

$$\begin{aligned} G_2(\mathbf{r},\mathbf{r}'') \;=\; & n_1\,(n_1-1)\ \ |u_1(\mathbf{r})|^2\,|u_1(\mathbf{r}'')|^2 + n_2\,(n_2-1)\ \ |u_2(\mathbf{r})|^2\,|u_2(\mathbf{r}'')|^2 \\ & + n_1 n_2 \left[2\,|u_1(\mathbf{r})|^2\,|u_2(\mathbf{r}'')|^2 + u_1^*(\mathbf{r})u_1(\mathbf{r}'')u_2^*(\mathbf{r}'')u_2(\mathbf{r}) + \text{c.c.} \right] \end{aligned} \tag{27}$$

where the factor 2 in the second line comes from the fact that either $i = 1$ and $j = 2$, or the opposite; the last two terms of this expression correspond to the exchange term, and the notation c.c. indicates the complex conjugate of the previous term. Replacing the wave functions by $e^{-i\mathbf{k}_{1,2}\cdot\mathbf{r}}/\sqrt{\mathcal{V}}$, and assuming n_1 and n_2 to be very large compared to 1, we obtain the square of the sum $(n_1+n_2)^2 = N^2$, and we can write:

$$G_2(\mathbf{r},\mathbf{r}'') \simeq \frac{N^2}{\mathcal{V}^2} + 2\frac{n_1 n_2}{\mathcal{V}^2}\cos\left[(\mathbf{k}_2-\mathbf{k}_1)\cdot(\mathbf{r}-\mathbf{r}'')\right] \tag{28}$$

The first term is simply the square of the one-particle density $N/\mathcal{V}$; it does not have any spatial dependence and is what we expect in the absence of any particle correlation. On the other hand, the exchange term is position dependent; it presents a maximum when $\mathbf{r} = \mathbf{r}''$, and oscillates at the spatial frequency $|\mathbf{k}_2 - \mathbf{k}_1|$. This exchange term enhances the probability of finding two bosons close to one another (bunching effect

coming from the Bose-Einstein statistics); the probability of finding them at a greater distance is then lower, and then increases again, etc. For a short interaction range, only the first maximum plays a role, and increases the average value of the interaction. The consequences of that effect, in terms of the internal interaction energy, has been discussed in § 4-c of Complement C_{XV}.

3-c. Other states

We now consider situations described by Fock states where n_0 is still very large, but where other states $|u_i\rangle$ are also occupied, with populations n_i much smaller than n_0.

(i) One could for instance place a finite fraction n_0/N of particles in the ground state, and distribute the remaining fraction $1 - n_0/N$ among a large number of states, whose individual populations remain small and vary regularly with the index i. This leads to:

$$G_1(\mathbf{r};\mathbf{r}') = n_0 \; u_0^*(\mathbf{r})u_0(\mathbf{r}') + Q(\mathbf{r}-\mathbf{r}') \tag{29}$$

where $Q(\mathbf{r}-\mathbf{r}')$ is given by:

$$Q(\mathbf{r}-\mathbf{r}') = \frac{1}{\mathcal{V}}\sum_{i\neq 0} n_i \; e^{i\mathbf{k}_i\cdot(\mathbf{r}'-\mathbf{r})} = \frac{1}{(2\pi)^3}\int \mathrm{d}^3k_i \; n(\mathbf{k}_i) \; e^{i\mathbf{k}_i\cdot(\mathbf{r}'-\mathbf{r})} \tag{30}$$

This function is the Fourier transform of the distribution $n_i = n(\mathbf{k}_i)$ for $\mathbf{k}_i \neq 0$. As all the n_i are positive or zero, the function $Q(\mathbf{r}-\mathbf{r}')$ presents a maximum when $\mathbf{r} = \mathbf{r}'$, since this is where all the exponentials $e^{i\mathbf{k}_i\cdot(\mathbf{r}'-\mathbf{r})}$ are in phase; it then decreases when the difference $|\mathbf{r}-\mathbf{r}'|$ increases, as all the phases spread out. If the n_i distribution is a regular function of width Δk (a Gaussian for example), the function $Q(\mathbf{r}-\mathbf{r}')$ tends towards zero over a distance $\Delta l \simeq 1/\Delta k$, in general much smaller than the size of the system.

Figure 3 is a plot of the function $G_1(\mathbf{r};\mathbf{r}')$ when the particles are contained in a box, so that we can use (1). As we assumed $|u_0\rangle$ to be the ground state, the corresponding wave function is the inverse of the square root of the volume, and (29) becomes:

$$G_1(\mathbf{r};\mathbf{r}') = x_0 \; + Q(\mathbf{r}-\mathbf{r}') \tag{31}$$

where x_0 is the density of atoms in the ground state:

$$x_0 = \frac{n_0}{\mathcal{V}} \tag{32}$$

After the decrease linked to that of $Q(\mathbf{r}-\mathbf{r}')$, and occurring over the interval $\Delta l \simeq 1/\Delta k$, the function does not tend towards zero (as it would for fermions), but towards a constant proportional to the population n_0. As already mentioned, this particular behavior is the base for the Penrose and Onsager criterion that defines, in a general way, the appearance of Bose-Einstein condensation.

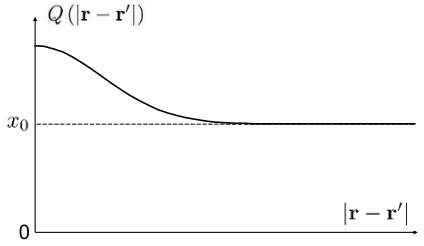


Figure 3: Plot of the function $G_1(\mathbf{r};\mathbf{r}') = x_0 + Q(\mathbf{r}-\mathbf{r}')$ for bosons, as a function of the distance $|\mathbf{r}-\mathbf{r}'|$. The function starts by decreasing over an interval of the order of the inverse of Δk; at a larger distance, it tends towards a constant x_0 proportional to the ground state population. The fact that it does not go to zero indicates the presence of a long-range non-diagonal order, and the existence of a highly populated individual level.

As for the two-body correlation function, relation (10) shows the existence of three kinds of terms for a system having only one single highly populated state $|u_0\rangle$:

– the terms corresponding to two particles in the highly populated state (condensate), which come from the first line[2] of (10), and yield again (25), replacing N by n_0.

– the crossed terms in $n_0 n_i$, which yield:

$$n_0 \sum_{i\neq 0} n_i \left[2\left|u_i(\mathbf{r})\right|^2 \left|u_0(\mathbf{r}'')\right|^2 + \; u_0^*(\mathbf{r})u_0(\mathbf{r}'')u_i^*(\mathbf{r}'')u_i(\mathbf{r}) + u_i^*(\mathbf{r})u_i(\mathbf{r}'')u_0^*(\mathbf{r}'')u_0(\mathbf{r})\right] \tag{33}$$

Inserting the value (1) for the wave functions (assuming $\mathbf{k}_0 = 0$), we get a contribution to $G_2(\mathbf{r},\mathbf{r}')$ equal to:

$$\frac{n_0}{\mathcal{V}^2} \sum_{i\neq 0} n_i \left[2 + e^{i\mathbf{k}_i\cdot\left(\mathbf{r}-\mathbf{r}''\right)} + e^{i\mathbf{k}_i\cdot\left(\mathbf{r}''-\mathbf{r}\right)}\right] \tag{34}$$

Using relation $\sum_{i\neq 0} n_i = N - n_0$, we can write this result in the form:

$$\frac{2n_0}{\mathcal{V}}\left[\frac{N-n_0}{\mathcal{V}} + \mathrm{Re}\left[Q(\mathbf{r}-\mathbf{r}')\right]\right] \tag{35}$$

[2]In the limit of large volumes, we have assumed that n_0 is the only population proportional to the volume. In the first line of (10), the term $i = 0$ then contains the product of n_0^2 and the two wave functions squared, each proportional to the inverse of the volume; this term is therefore independent of the volume. On the other hand, the terms $i \neq 0$ of the second line contains one summation over i, introducing a factor proportional to the volume (in the limit of large volumes), but also two wave functions squared, each inversely proportional to the volume. The net result is a contribution inversely proportional to the volume, hence negligible compared to the previous one.

where Re means "real part" and Q is the function defined in (30).

– the terms corresponding to two particles in states other than the ground state, and which yield:

$$\sum_{i \neq j} \frac{n_i n_j}{\mathcal{V}^2} \left[1 + e^{i(\mathbf{k}_i - \mathbf{k}_j)\cdot(\mathbf{r}'' - \mathbf{r})} \right] \tag{36}$$

In (34), as well as in (36), we notice that the contributions from all the states $i \neq 0$ have various phases in general; they are, however, all in phase when $\mathbf{r} = \mathbf{r}'$ and the correlation function then presents a maximum. The bosons have thus a tendency for bunching, and this effect is felt over a distance $\Delta l \simeq 1/\Delta k$, as if they were attracted to one another. It is, however, a purely statistical effect linked to the bosonic character of the particles, since we assumed there were no interactions between the particles.

(ii) One could also imagine the population distribution to be regular, without favoring any individual state, in which case the n_0 contribution vanishes; we are then left with the contribution from the $n_i n_j$ terms of (36), which is maximum when $\mathbf{r} = \mathbf{r}'$ for the same reasons as above. The general behavior of the two-body correlation function is shown in Figure 3: it presents a maximum at the origin, and then tends towards zero at large distance (in this case, $x_0 = 0$). Once again, identical bosons exhibit a bunching tendency.

Comment:

Suppose that, instead of assuming the bosonic system to be in a Fock state (a pure state), it is at thermal equilibrium, described by the thermodynamic equilibrium. This would lead to results similar to those we just derived, but with $\Delta l \simeq \lambda_T$, the thermal wavelength of the particles [24]. The boson bunching tendency is a quite general property.

Complement B_{XVI}

Spatio-temporal correlation functions, Green's functions

1 Green's functions in ordinary space 1781
1-a Spatio-temporal correlation functions 1782
1-b Two- and four-point Green's functions 1786
1-c An example, the ideal gas 1787
2 Fourier transforms 1790
2-a General definition 1790
2-b Ideal gas example 1791
2-c General expression in the presence of interactions 1792
2-d Discussion 1794
3 Spectral function, sum rule 1795
3-a Expression of the one-particle correlation functions 1795
3-b Sum rule 1796
3-c Expression of various physical quantities 1797

This complement discusses the properties of the spatio-temporal correlation functions of an ensemble of identical particles, generalizing the spatial correlation functions defined in § C-5-b of Chapter XV and § B-3 of Chapter XVI; the corresponding Green's functions shall also be introduced. We first study (§ 1) the normal and anti-normal spatio-temporal correlation functions, then the Green's function, and discuss some of their properties, illustrated with the example of an ideal gas. We then study in § 2 the Fourier transforms of these functions for physical systems that are translation invariant both in space and time; we shall write their general expression in the presence of interactions. In § 3, we finally introduce the "spectral function", which leads, for interacting particles, to very simple expressions for various physical quantities, in a form similar to the one used for an ideal gas.

1. Green's functions in ordinary space

In the previous complement, we studied the spatial dependence of the correlation functions, taken at a given time t. We now take into account the temporal dependence, using the Heisenberg picture (Complement G_{III}) where the operators are time-dependent. To keep the notation simple, we assume, from now on, that either the spin is zero for bosons ($S = 0$), or else, in the general case (fermions and bosons), that all the particles are in the same spin state. As mentioned before, the generalization to the case where S is non-zero would only require adding an index ν to all the field operators. In the Heisenberg picture, the field operator $\Psi(\mathbf{r})$ becomes a time-dependent operator $\Psi_H(\mathbf{r}, \mathbf{t})$:

$$\Psi_H(\mathbf{r}, t) = e^{iHt/\hbar}\ \Psi(\mathbf{r})\ e^{-iHt/\hbar} \tag{1}$$

where $e^{-iHt/\hbar}$ is the evolution operator, expressed as a function of the system Hamiltonian H (including the particle interactions when present), that we assume to be time-independent.

Consider a system of N identical particles, fermions or bosons, described by a density operator ρ. The spatio-temporal correlation functions and the Green's functions are defined as the average values, computed with ρ, of the products of a number of field operators $\Psi_H(\mathbf{r},t)$ and their Hermitian conjugates $\Psi_H^\dagger(\mathbf{r},t)$ taken at different space-time points $(\mathbf{r},t)$, $(\mathbf{r}',t')$,... etc.

1-a. Spatio-temporal correlation functions

The density operator ρ may contain very complex correlations between particles, and its time evolution can be very complicated in the presence of interactions. We are going to define a certain number of functions that characterize its most simple and useful properties, as they only pertain to a small number of particles.

α. *Two-point normal and anti-normal functions*

The one-particle spatio-temporal "normal" G_1^N and "anti-normal" G_1^{AN} correlation functions are defined by[1]:

$$\boxed{\begin{aligned} G_1^N(\mathbf{r},t;\mathbf{r}',t') &= \mathrm{Tr}\left\{\rho\ \Psi_H^\dagger(\mathbf{r},t)\Psi_H(\mathbf{r}',t')\right\} = \left\langle \Psi_H^\dagger(\mathbf{r},t)\Psi_H(\mathbf{r}',t')\right\rangle \\ G_1^{AN}(\mathbf{r},t;\mathbf{r}',t') &= \mathrm{Tr}\left\{\rho\ \Psi_H(\mathbf{r}',t')\Psi_H^\dagger(\mathbf{r},t)\right\} = \left\langle \Psi_H(\mathbf{r}',t')\Psi_H^\dagger(\mathbf{r},t)\right\rangle \end{aligned}} \tag{2}$$

The normal ordering is obtained when the creation operator is on the left and the annihilation on the right; it is the opposite for the anti-normal order. Note that it is only the order of the two operators that changes between G_1^N and G_1^{AN}; the position and time variables attributed to each of the two operators Ψ_H and $\Psi_H^\dagger$ remain the same.

In the particular case where $t' = t$, the normal correlation function simply yields the matrix elements of the one-particle density operator – see formulas (B-22) and (B-23) of Chapter XVI. The normal correlation function is therefore a generalization, at different times, of this matrix element, which will prove to be useful.

To understand the physical meaning of these two definitions in an intuitive way, we start with the anti-normal function and consider the simple case where an N-particle system is in a pure state $|\Phi_0\rangle$ (its ground state, for example). We then get:

$$\begin{aligned} G_1^{AN}(\mathbf{r},t;\mathbf{r}',t') &= \langle\Phi_0|\,\Psi_H(\mathbf{r}',t')\Psi_H^\dagger(\mathbf{r},t)\,|\Phi_0\rangle \\ &= \langle\Phi_0|\,e^{iHt'/\hbar}\Psi(\mathbf{r}')e^{iH\left(t-t'\right)/\hbar}\Psi^\dagger(\mathbf{r})e^{-iHt/\hbar}\,|\Phi_0\rangle \end{aligned} \tag{3}$$

The right-hand side reads as follows (from right to left): starting from the system initial state $|\Phi_0\rangle$, we let it evolve according to its own Hamiltonian until the time t, when we create a particle at point $\mathbf{r}$; we then let the $(N+1)$-particle system freely evolve until time t', and we finally annihilate a particle at point $\mathbf{r}'$. Consequently, the function G_1^{AN} is the scalar product of the ket thus obtained with the state $e^{-iHt'/\hbar}\,|\Phi_0\rangle$, result of the

[1] To be consistent with the notation of Chapter XVI (§ B), we choose a notation where, for the trace, the first group of variables $(\mathbf{r},t)$ of function G_1 is associated with the operator $\Psi^\dagger$, the second group $(\mathbf{r}',t')$ to the operator Ψ. Note however that the opposite convention can also be found in the literature.

free evolution of the initial ket over the same time interval (without the creation or annihilation of a particle). In other words, the perturbation created by the creation of a particle, followed by its later destruction, changes the state of the physical system; the value of G_1^{AN} is given by the probability amplitude of finding the system in the same state as the one it would have reached in the absence of this perturbation.

The previous interpretation is natural when $t' \geq t$; if this is not the case, the mathematical definition of G_1^{AN} is the same, but the intermediate evolution stage goes backward in time. In this process, and as expected, the dynamics of the system remains unchanged, including the particle interactions. Furthermore, the additional particle is not simply a particle juxtaposed to the pre-existing system, it is indistinguishable from the others and hence undergoes indistinguishability effects (more details on this point will be given in § 1-a-β).

For the normal function, $\Psi_H(\mathbf{r}', t')$ acts before $\Psi_H^\dagger(\mathbf{r}, t)$, which means this function has a natural interpretation if $t \geq t'$: the system evolves freely until t', at which time a particle is annihilated (which amounts to the creation of a "hole", see Complement $\mathrm{A_{XV}}$ $\mathrm{A_{XV}}$); the system, with $N-1$ particles, then evolves freely until time t, when a particle is created (which annihilates the hole). The normal function is therefore the analog of the anti-normal function, provided we replace the additional particle by a hole, and we invert the times.

β. *Physical discussion*

The definition of the correlation functions contains particle creation and annihilation operators, but this does not imply that such physical processes really occur in our system. Going from an N-particle system to another one with $N \pm 1$ particles can be mathematically useful but only plays an intermediate role since we finally return, via a second operator, to the same number of particles. Furthermore, the action of an operator $\Psi(\mathbf{r})$ is not merely a local destruction of a particle, neither is the action of $\Psi^\dagger(\mathbf{r})$ the simple creation of a particle at point $\mathbf{r}$, juxtaposed with the already present particles: the quantum indistinguishability that concerns all the particles (including the new one) plays an essential role.

A few simple examples

(i) For the normal correlation function, the perturbation starts with the annihilation of a particle (creation of a hole), followed by a later creation of a particle (destruction of a hole). At time $t = 0$, the annihilation operator $\Psi_H(\mathbf{r}', t)$ can be expanded according to formula (A-3) of Chapter XVI:

$$\Psi(\mathbf{r}') = \sum_i u_i(\mathbf{r}')\ a_i \tag{4}$$

where the effect of each operator a_i depends on the population n_i of the state $|u_i\rangle$ as it introduces the factor $\sqrt{n_i}$. Consider a very simple case: a gas of bosons, all in the same quantum state $|u_0\rangle$ (ideal gas of totally condensed bosons). Acting on a state $|\Phi_0\rangle$ where only the individual state $|u_0\rangle$ is occupied, all the terms of the sum (4) yield zero, except for the $i = 0$ term. The effect of the operator $\Psi(\mathbf{r}')$ on $|\Phi_0\rangle$ is to actually destroy a particle in the state $|u_0\rangle$; as $\mathbf{r}'$ varies, the result is still the same, simply multiplied by the coefficient $u_0(\mathbf{r}')$. If, for example, the ideal gas is placed in a trap where the individual ground state is $|u_0\rangle$, the operator $\Psi(\mathbf{r}')$ yields a ket with an appreciable norm only if

$\mathbf{r}'$ falls in a domain where $|u_0(\mathbf{r}')|$ is not negligible; if it falls outside this domain, the resulting ket is practically zero. In a general way, for fermions as for bosons, $\Psi(\mathbf{r}')$ can obviously destroy particles only while acting on an already occupied individual state. For bosons, in addition, the factor $\sqrt{n_i}$ means that the operator $\Psi(\mathbf{r}')$ gives more weight to the highly populated individual states rather than those with low occupation numbers. Consequently, the creation of a hole is not a local process at point $\mathbf{r}'$.

(ii) For the anti-normal correlation function, we start with the creation of a particle by the operator $\Psi^\dagger(\mathbf{r})$. In the case of bosons, and because of the factor $\sqrt{n_i+1}$ introduced by $a_i^\dagger$, $\Psi^\dagger(\mathbf{r})$ tends to preferentially create particles in states $|u_i\rangle$ having a high population n_i. Let us go back to the previous example of a large number of bosons in a trap, all in the same individual ground state $|u_0\rangle$. When $|u_0(\mathbf{r})|$ is not negligible, the supplementary boson is created in the same ground state. If, on the other hand, $\mathbf{r}$ is far away from the trap center and falls in a domain where $|u_0(\mathbf{r})|$ is practically zero, the boson is actually created at point $\mathbf{r}$ but without perturbing very much the bosons already present.

For fermions, on the contrary, it is impossible to create a particle in an already occupied state; in a Fock state, an additional fermion can only be created in a state orthogonal to all the initially occupied states. Let us assume the ideal gas of fermions is in its ground state, and contained in a harmonic trap; the energy levels, up to the Fermi level, are all occupied. The effect of the creation operator $\Psi^\dagger(\mathbf{r})$ at a point close to the trap center is to create an additional particle in a state that can be expanded on all the individual stationary states of the trap; as this state must be orthogonal to all the already occupied states, it only has components on the non-occupied states, which have an energy higher than the Fermi level. Now the corresponding wave functions take on small values at the center of the trap, and are maximum in the classical turning point regions[2], which means, in this case, at the edge of the existing fermion cloud (or even further). If position $\mathbf{r}$ is close to the center of the trap, the additional fermion will be added on the periphery or outside the cloud of fermions. On the other hand, if $\mathbf{r}$ falls outside, in a region of space where all the wave functions of the occupied states are practically zero, the additional particle will be created practically right at point $\mathbf{r}$. Examples of these various situations will be given in § 1-c.

γ. *Properties*

The complex conjugate of G_1^N is obtained by changing the field operators' order in (2) and replacing them by their Hermitian conjugates:

$$\left[G_1^N(\mathbf{r},t;\ \mathbf{r}',t')\right]^* = \left\langle \Psi_H^\dagger(\mathbf{r}',t')\Psi_H(\mathbf{r},t)\right\rangle = G_1^N(\mathbf{r}',t';\ \mathbf{r},t) \tag{5}$$

The complex conjugation of the function G_1^N is therefore equivalent to exchanging the two points $(\mathbf{r},t)$ and $(\mathbf{r}',t')$, which amounts to a parity operation on the variables $\mathbf{r}-\mathbf{r}'$ and $t-t'$. A similar property is easily demonstrated for G_1^{AN}. As a result, the Fourier transforms of these functions with respect to the variables $\mathbf{r}-\mathbf{r}'$ and $t-t'$ must be real.

[2]The modulus squared of a stationary state wave function yields, at each point, the probability density of presence for this state. This probability is maximal in the regions of space where, classically, the particle spends the most time, i.e. where its velocity is small, as is the case for the classical turning point regions. Figure 6 of Chapter V gives an example of such a situation.

In addition, when the system is in a state that is translation invariant in space and time[3], the correlation functions only depend on the differences $\mathbf{r}-\mathbf{r}'$ and $t-t'$.

We note that the linear combination $G_1^{AN}(\mathbf{r},t;\mathbf{r}',t')-\eta G_1^N(\mathbf{r},t;\mathbf{r}',t')$ involves the average value of the operator:

$$\left[\Psi_H(\mathbf{r}',t')\ ,\ \Psi_H^\dagger(\mathbf{r},t)\right]_{-\eta} \tag{6}$$

where $\eta=+1$ for bosons, $\eta=-1$ for fermions. If $t=t'$, and taking into account relation (A-19) of Chapter XVI, this equality becomes:

$$G_1^{AN}(\mathbf{r},t;\mathbf{r}',t)-\eta G_1^N(\mathbf{r},t;\mathbf{r}',t)=\mathrm{Tr}\left\{\rho\ \delta\left(\mathbf{r}-\mathbf{r}'\right)\right\}=\delta\left(\mathbf{r}-\mathbf{r}'\right) \tag{7}$$

δ. Temporal evolution, four-point functions

We now use, in definition (2) of G_1^N, the Hermitian conjugate of the evolution equation for $\Psi_H(\mathbf{r},t)$, written in (C-13) of Chapter XVI; this leads to:

$$\begin{aligned}\left[-i\hbar\frac{\partial}{\partial t}+\frac{\hbar^2}{2m}\Delta_{\mathbf{r}}-V_1(\mathbf{r})\right]&G_1^N(\mathbf{r},t;\mathbf{r}',t')\\ &=\int \mathrm{d}^3r''\ W_2(\mathbf{r}-\mathbf{r}'')\ \left\langle\Psi_H^\dagger(\mathbf{r};t)\Psi_H^\dagger(\mathbf{r}'';t)\Psi_H(\mathbf{r}'';t)\Psi_H(\mathbf{r}';t')\right\rangle\\ &=\int \mathrm{d}^3r''\ W_2(\mathbf{r}-\mathbf{r}'')\ G_2^N(\mathbf{r},t;\mathbf{r}',t;\mathbf{r}'',t;\mathbf{r}'',t)\end{aligned} \tag{8}$$

which involves the normal two-particle (or four-point) correlation function, whose general expression is:

$$G_2^N(\mathbf{r},t;\mathbf{r}',t';\mathbf{r}'',t'';\mathbf{r}''',t''')=\mathrm{Tr}\left\{\rho\ \Psi_H^\dagger(\mathbf{r},t)\Psi_H^\dagger(\mathbf{r}'',t'')\Psi_H(\mathbf{r}''',t''')\Psi_H(\mathbf{r}',t')\right\} \tag{9}$$

Taking into account a time dependence generalizes formula (B-30) of Chapter XVI (or more precisely its average value). We obtain, in a similar way, the equation giving the variation of G_1^N with respect to the variables $\mathbf{r}'$ and t', or that giving the evolution of the anti-normal function G_1^{AN}. The four groups of space-time variables the function G_2^N depends on, are, in general, independent. Most of the time, however, we only need the "diagonal part" $G_2^N(\mathbf{r},t;\mathbf{r}'',t'')$ of the correlation function, obtained for $\mathbf{r}'=\mathbf{r}$, $t'=t$ and $\mathbf{r}'''=\mathbf{r}''$, $t'''=t''$; this diagonal part is analogous to the two-particle correlation function (two positions, two times) in classical statistical mechanics. When the system is translation-invariant both in space and time, this function only depends on the differences $\mathbf{r}-\mathbf{r}''$ and $t-t''$.

Whereas, for the function G_1^N, a hole is created and then destroyed, for the function G_2^N there are now two holes first created and then destroyed; the natural order of the increasing times is given by the time variables of the operators in (9), taken from right to left. One could define, in a similar way, a function G_2^{AN} where two particles would be created at the beginning and then destroyed at later times (compared to the normal correlation function, the role of particles and holes are interchanged). Consequently, when the particles interact, the evolution equation of G_1^N involves another higher order

[3]This is the case if the system Hamiltonian is translation invariant, and if the system is in an eigenstate of H, or described by a density operator ρ that is a function of H.

correlation function, G_2^N. In turn, the evolution of G_2^N involves correlation functions of an even higher order, G_3^N etc. This means that, because of the interactions, the set of equations is not "closed", but includes a complete "hierarchy" of a large number of equations, involving correlations of higher and higher order.

1-b. Two- and four-point Green's functions

Equation (8) is a linear partial differential equation, with a right-hand side sometimes called a "source term". In our case, this right-hand side does not contain any singular function. But when this right-hand side is modified to include a delta function, the new solutions of the equation are called the "Green's functions" $\mathcal{G}$. We now show how to introduce Green's functions in the problem we are concerned with.

α. *Two-point Green's function*

The two-point Green's function $\mathcal{G}_1$ is obtained by a combination of the two correlation functions of § 1-a. We saw in § 1-a that, when $t' > t$, the anti-normal correlation function is the most natural as it includes the propagation of a particle from t to t'. On the other hand, for $t' < t$, the most natural one is the normal function that involves the propagation of a hole from t' to t. We can combine those two possibilities into one by setting:

$$\boxed{\mathcal{G}_1(\mathbf{r},t;\mathbf{r}',t') = Y\,(t'-t)\,G_1^{AN}(\mathbf{r},t;\mathbf{r}',t') + \eta Y\,(t-t')\,G_1^{N}(\mathbf{r},t;\mathbf{r}',t')} \tag{10}$$

where $Y\,(t-t')$ is the Heaviside function of the variable $t'-t$ (equal to 1 if $t'-t>0$, zero otherwise), and where η equals $+1$ for bosons, -1 for fermions; we shall see later on, for example in § 1-c-γ, that the introduction of the η factor simplifies the following computations.

When we take the derivative of the two-point Green's function $\mathcal{G}_1$ with respect to time, the discontinuities introduced by the Heaviside functions yield delta functions; the precise calculation will be done in § 1-c-γ for an ideal gas, allowing us to verify that $\mathcal{G}_1$ is indeed a Green's function. Using this type of function is quite useful in a number of problems, for example those involving Fourier transforms, or in perturbation calculations.

β. *Four-point Green's function*

By analogy with (9) we define a two-particle (or four-point) Green's function:

$$\boxed{\mathcal{G}_2(\mathbf{r},t;\mathbf{r}',t';\mathbf{r}'',t'';\mathbf{r}''',t''') = \mathrm{Tr}\left\{\rho\ T\ \Psi_H(\mathbf{r}',t')\Psi_H(\mathbf{r}''',t''')\Psi_H^\dagger(\mathbf{r}'',t'')\Psi_H^\dagger(\mathbf{r},t)\right\}} \tag{11}$$

where T is the time ordering operator, which orders the 4 times by decreasing values from left to right (by definition, this operator also includes a factor η^P, where P is the parity of the permutation necessary for this time ordering; this may result in a change of sign for fermions).

1-c. An example, the ideal gas

When the system considered is an ideal gas, it is possible to get explicit values of the previous functions. As before, we assume the gas to be contained in a box of volume $\mathcal{V}$, with periodic boundary conditions. Using relation (A-3) of Chapter XVI in the case where the $u_i(\mathbf{r})$ are plane waves, we can write the field operator $\Psi(\mathbf{r})$ as:

$$\Psi(\mathbf{r}) = \sum_i \frac{e^{i\mathbf{k}_i\cdot\mathbf{r}}}{\sqrt{\mathcal{V}}}\ a_{\mathbf{k}_i} \tag{12}$$

where the sum covers all the wave vectors $\mathbf{k}_i$ satisfying the periodic boundary conditions. This expansion is convenient as, for an ideal gas, the time dependence of the operators in the Heisenberg picture $a^H_{\mathbf{k}}$ is particularly simple, as we now show. The Hamiltonian can be written as:

$$H = \sum_j\ \hbar\omega_{\mathbf{k}_j}\left(a^\dagger_{\mathbf{k}_j}a_{\mathbf{k}_j}\right) \tag{13}$$

with:

$$\omega_{\mathbf{k}_j} = \frac{\hbar\,(\mathbf{k}_j)^2}{2m} \tag{14}$$

As the operator $\left(a^\dagger_{\mathbf{k}_j}a_{\mathbf{k}_j}\right)$ commutes with any annihilation operator $a_{\mathbf{k}_i}$ pertaining to a different momentum[4], and since $\left[a_{\mathbf{k}_i}, a^\dagger_{\mathbf{k}_i}a_{\mathbf{k}_i}\right] = a_{\mathbf{k}_i}$, we get:

$$[a_{\mathbf{k}_i}, H] = \hbar\omega_{\mathbf{k}_i}\ a_{\mathbf{k}_i} \tag{15}$$

This corresponds, in the Heisenberg picture, to an evolution:

$$a^H_{\mathbf{k}_i}(t) = e^{-i\omega_{\mathbf{k}_i}t}\ a_{\mathbf{k}_i} \tag{16}$$

The time evolution of $\Psi_H(\mathbf{r},t)$ is then written as:

$$\Psi_H(\mathbf{r},t) = \sum_i \frac{e^{i\left(\mathbf{k}_i\cdot\mathbf{r}-\omega_{\mathbf{k}_i}t\right)}}{\sqrt{\mathcal{V}}}a_{\mathbf{k}_i} \tag{17}$$

α. Normal correlation function

Inserting this result in (2), we get:

$$G^N_1(\mathbf{r},t;\mathbf{r}',t') = \sum_{i,j}\frac{e^{-i\left(\mathbf{k}_j\cdot\mathbf{r}-\omega_{\mathbf{k}_j}t\right)}e^{i\left(\mathbf{k}_i\cdot\mathbf{r}'-\omega_{\mathbf{k}_i}t'\right)}}{\mathcal{V}}\left\langle a^\dagger_{\mathbf{k}_j}a_{\mathbf{k}_i}\right\rangle \tag{18}$$

We now show that, because of the translation invariance, the average value $\left\langle a^\dagger_{\mathbf{k}_j}a_{\mathbf{k}_i}\right\rangle$ must be zero whenever $\mathbf{k}_j \neq \mathbf{k}_i$. Assume, for example, that the system density operator

[4]For fermions, two minus signs are introduced because of the anticommutations, but they cancel each other. If the momentum is the same, we have $a_{\mathbf{k}_i}a^\dagger_{\mathbf{k}_i}a_{\mathbf{k}_i} = -a^\dagger_{\mathbf{k}_i}a_{\mathbf{k}_i}a_{\mathbf{k}_i} + a_{\mathbf{k}_i} = 0 + a_{\mathbf{k}_i}$ and $a^\dagger_{\mathbf{k}_i}a_{\mathbf{k}_i}a_{\mathbf{k}_i} = 0$, so that $\left[a_{\mathbf{k}_i}, a^\dagger_{\mathbf{k}_i}a_{\mathbf{k}_i}\right] = a_{\mathbf{k}_i}$.

is the canonical thermal equilibrium operator $\rho_{eq} = e^{-\beta H}/Z$. This operator is diagonal in the basis of the Fock states, and the trace of the product $\rho a^\dagger_{\mathbf{k}_j} a_{\mathbf{k}_i}$ will be zero unless both annihilation and creation operators act on the same individual state: the non-zero condition is therefore $\mathbf{k}_j = \mathbf{k}_i$. In the same way, if we are now using the grand canonical equilibrium $\rho_{eq} = e^{-\beta(H-\mu N)}/Z$, the operator ρ is still diagonal in the same basis, and the same rule applies. In a general way, one can see[5] that the translation invariance of ρ requires:

$$\left\langle a^\dagger_{\mathbf{k}_j} a_{\mathbf{k}_i} \right\rangle = \mathrm{Tr}\left\{ \rho a^\dagger_{\mathbf{k}_j} a_{\mathbf{k}_i} \right\} = \delta_{ij} \times n_i \tag{19}$$

where n_i is the average value of the population operator $\widehat{n}_i$:

$$n_i = \langle \widehat{n}_i \rangle \tag{20}$$

We can then write:

$$G_1^N(\mathbf{r}, t; \mathbf{r}', t') = \frac{1}{\mathcal{V}} \sum_i n_i \; e^{i\left[\mathbf{k}_i \cdot (\mathbf{r}' - \mathbf{r}) - \omega_{\mathbf{k}_i}(t' - t)\right]} \tag{21}$$

The normal correlation function is simply the sum of all the contributions from the individual states occupied by free particles, labeled by the index $\mathbf{k}_i$. Each contributes proportionally to its average population n_i, and has a spatio-temporal dependence given by the progressive wave $e^{i\left[\mathbf{k}_i \cdot (\mathbf{r}' - \mathbf{r}) - \omega_{\mathbf{k}_i}(t' - t)\right]}$ it is associated with.

As expected for a translation invariant (both in space and time) system, this normal function only depends on the differences $\mathbf{r} - \mathbf{r}'$ and $t - t'$. Taking (14) into account, it obeys the partial differential equation:

$$\left[i \frac{\partial}{\partial t'} + \frac{\hbar}{2m} \Delta_{\mathbf{r}'} \right] G_1^N(\mathbf{r}, t; \mathbf{r}', t') = 0 \tag{22}$$

which corresponds to the free propagation of particles in an ideal gas (a similar equation exists for the variables t and $\mathbf{r}$, with a change of sign for the time derivative term).

Expression (21) allows shedding new light on the physical interpretation of the normal correlation function given in § 1-a, in terms of the creation, in the N-particle system, of a "hole", which then propagates until it is annihilated at a later time. In an ideal gas, each term of the sum in (21) is a free wave plane: in the absence of interactions, the particles can freely propagate along straight lines. The n_i appearing in the formula shows that the hole can only propagate along already populated individual states in the N-particle state: a hole can only be created in an already occupied quantum state, as pointed out already in § 1-a-β. As a result, the created hole is not a point-like object actually localized at point $\mathbf{r}'$: it is only built from superpositions of free occupied states, whereas for a truly point-like excitation, one would have to combine values of $\mathbf{k}_i$ extending to infinity.

[5] To prove this in a general way, we simply have to note that the operator $a^\dagger_{\mathbf{k}_j} a_{\mathbf{k}_i}$ is translation invariant only if $\mathbf{k}_j = \mathbf{k}_i$. To show this, one can use the expression of the translation operator as an exponential of the operator associated with the total momentum (Complement E_{II}, § 3).

β. Anti-normal correlation function

For the anti-normal function $G_1^{AN}(\mathbf{r},t;\mathbf{r}',t')$, the calculation is practically the same, the only difference being the inversion of the order of the operators $a_{\mathbf{k}}$ and $a_{\mathbf{k}}^{\dagger}$, which leads to:

$$G_1^{AN}(\mathbf{r},t;\mathbf{r}',t') = \frac{1}{\mathcal{V}}\sum_i \left[1+\eta n_i\right]\ e^{i\left[\mathbf{k}_i\cdot(\mathbf{r}'-\mathbf{r})-\omega_{\mathbf{k}_i}(t'-t)\right]} \tag{23}$$

It obeys the same partial differential equation as the normal function. It follows from (21) and (23) that the linear combination:

$$G_1^{AN}(\mathbf{r},t;\mathbf{r}',t') - \eta G_1^{N}(\mathbf{r},t;\mathbf{r}',t') = \frac{1}{\mathcal{V}}\sum_i e^{i\left[\mathbf{k}_i\cdot(\mathbf{r}'-\mathbf{r})-\omega_{\mathbf{k}_i}(t'-t)\right]} \tag{24}$$

is independent of the state of the system; for $t=t'$, it is simply equal to the function $\delta(\mathbf{r}-\mathbf{r}')$. We shall see later the relation between this expression and the spectral function.

Let us come back, here again, to the interpretation of the effect of the operator $\Psi^{\dagger}(\mathbf{r})$, discussed in § 1-a-β. For a fermion system ($\eta=-1$), relation (23) clearly shows that the particle creation does not involve any of the already occupied individual states $\mathbf{k}_i$, with $n_i=1$, since the corresponding term is zero. The object created by the excitation cannot have any component on the already occupied individual states, which simply means that its wave function must remain orthogonal to those of all the fermions already present. For a boson system, the effect is just the opposite: if, for instance, an individual state of bosons is highly populated compared to all the others, the term in (23) corresponding to this large n_i will be dominant: the additional particle will be mainly in the same individual state as all the other particles already present.

γ. Two-point Green's function

Using definition (10), we obtain the Green's function $\mathcal{G}_1(\mathbf{r},t;\mathbf{r}',t')$:

$$\mathcal{G}_1(\mathbf{r},t;\mathbf{r}',t') = \frac{1}{\mathcal{V}}\sum_i e^{i\left[\mathbf{k}_i\cdot(\mathbf{r}'-\mathbf{r})-\omega_{\mathbf{k}_i}(t'-t)\right]}\ \left\{Y\left(t'-t\right)\left[1+\eta n_i\right]+\eta Y\left(t-t'\right)n_i\right\} \tag{25}$$

If we take the derivative with respect to the time t', the two Heaviside step functions yield delta functions with opposite signs. We then get the partial differential equation:

$$\left[i\frac{\partial}{\partial t'}+\frac{\hbar}{2m}\Delta_{\mathbf{r}'}\right]\mathcal{G}_1(\mathbf{r},t;\mathbf{r}',t') = \frac{i\delta\left(t-t'\right)}{\mathcal{V}}\sum_i e^{i\mathbf{k}_i\cdot(\mathbf{r}'-\mathbf{r})}\left[1+\eta n_i-\eta n_i\right] \tag{26}$$

that is:

$$\left[i\frac{\partial}{\partial t'}+\frac{\hbar}{2m}\Delta_{\mathbf{r}'}\right]\mathcal{G}_1(\mathbf{r},t;\mathbf{r}',t') = i\ \delta\left(t-t'\right)\ \delta\left(\mathbf{r}'-\mathbf{r}\right) \tag{27}$$

As expected, the right-hand side is a product of delta functions of the set of variables characterizing a Green's function.

In the presence of interactions, this partial differential equation is no longer valid; we must add to its right-hand side the interaction contributions, which involve Green's functions of a higher order.

δ. Four-point functions

The computation of the four-point correlation functions is very similar to the one we just explained; we simply use again relation (17) to get their explicit expressions for an ideal gas contained in a box. The results are the same as those obtained in Complement A_{XVI}, as for example in relation (21), except for the fact that we must now multiply each spatial plane wave $e^{i\mathbf{k}_i\cdot\mathbf{r}}$ by the associated time evolution factor $e^{-i\omega_{\mathbf{k}_i}t}$.

2. Fourier transforms

From now on, we shall only study systems that are translation-invariant in space and time. Consequently, the correlation functions only depend on the difference of the variables $\mathbf{r}'-\mathbf{r}$ and $t'-t$, so that we can choose to cancel both $\mathbf{r}$ and t.

2-a. General definition

Let us introduce the two (double) Fourier transforms with respect to time and space:

$$\boxed{\overline{G}_1^{N,AN}(\mathbf{k},\omega)=\int \mathrm{d}^3r'\int \mathrm{d}t'\ e^{i(\omega t'-\mathbf{k}\cdot\mathbf{r}')}\ G_1^{N,AN}(\mathbf{0},0;\mathbf{r}',t')} \tag{28}$$

(we have set $\mathbf{r}=0$ and $t=0$ in the function $G_1^{N,AN}$). These functions are real, as shown by the parity relation (5).

The Fourier transform $\overline{\mathcal{G}}_1$ of the Green's function $\mathcal{G}_1$ introduced in (10) is called the "one-particle propagator"; it is defined by:

$$\boxed{\overline{\mathcal{G}}_1(\mathbf{k},\omega)=\int \mathrm{d}^3r'\int \mathrm{d}t'\ e^{i(\omega t'-\mathbf{k}\cdot\mathbf{r}')}\mathcal{G}_1(\mathbf{0},0;\mathbf{r}',t')} \tag{29}$$

For a system contained in a volume $\mathcal{V}$ with periodic boundary conditions (Complement F_{XI}), the integrals over d^3r' in the above formulas must be taken over the volume $\mathcal{V}$. They yield the coefficients of a Fourier series where the vectors $\mathbf{k}$ take the discrete values $\mathbf{k}_i$ corresponding to the boundary conditions. This series characterizes the spatial dependence. As for the time dependence, the Fourier transform is a continuous function[6]. The inverse transformation relations are:

$$G_1^{N,AN}(\mathbf{0},0;\mathbf{r}',t')=\frac{1}{\mathcal{V}}\sum_{\mathbf{k}_i}\int\frac{\mathrm{d}\omega}{2\pi}e^{i(\mathbf{k}_i\cdot\mathbf{r}'-\omega t')}\overline{G}_1^{N,AN}(\mathbf{k}_i,\omega) \tag{30}$$

where, in the limit of very large volumes, the discrete summation becomes an integral with the coefficient $\mathcal{V}/(2\pi)^3$:

$$G_1^{N,AN}(\mathbf{0},0;\mathbf{r}',t')=\int\frac{\mathrm{d}^3k}{(2\pi)^3}\int\frac{\mathrm{d}\omega}{2\pi}e^{i(\mathbf{k}\cdot\mathbf{r}'-\omega t')}\overline{G}_1^{N,AN}(\mathbf{k},\omega) \tag{31}$$

[6] As opposed to the space variables, the time variable is not confined to a finite variation domain, and time Fourier transforms may be singular; such an example, concerning an ideal gas, will be given in § 2-b, where we shall introduce, as a convergence factor, a decreasing exponential $e^{\pm\varepsilon t}$ (where ε tends toward zero through positive values).

Comment:

One can also express the functions $\overline{G}_1^{N,AN}(k,\omega)$, not as Fourier transforms as in (28), but directly from the average values of products of creation $a^\dagger_{\mathbf{k}_j}$ and annihilation $a_{\mathbf{k}_j}$ operators in the individual states $|\mathbf{k}\rangle$. Formulas (A-3) and (A-6) of Chapter XVI tell us that:

$$\Psi(\mathbf{r}) = \frac{1}{\sqrt{\mathcal{V}}} \sum_j e^{i\mathbf{k}_j\cdot\mathbf{r}} a_{\mathbf{k}_j} \qquad \text{and} \qquad \Psi^\dagger(\mathbf{r}) = \frac{1}{\sqrt{\mathcal{V}}} \sum_{j'} e^{-i\mathbf{k}_{j'}\cdot\mathbf{r}} a^\dagger_{\mathbf{k}_{j'}} \tag{32}$$

Inserting these relations in the definitions (2), then in (28), we get:

$$\overline{G}_1^N(\mathbf{k}_i,\omega) = \frac{1}{\mathcal{V}} \int \mathrm{d}^3r' \int \mathrm{d}t'\ e^{i(\omega t' - \mathbf{k}_i\cdot\mathbf{r}')} \sum_{j,j'} e^{i\mathbf{k}_j\cdot\mathbf{r}'} \left\langle a^\dagger_{\mathbf{k}_{j'}} a^H_{\mathbf{k}_j}(t') \right\rangle \tag{33}$$

where $a^H_{\mathbf{k}_j}(t')$ is the operator $a_{\mathbf{k}_j}$ in the Heisenberg picture. The integral over the volume selects a single term, $j = i$, and cancels the volume $\mathcal{V}$. Furthermore, the translation invariance means that neither the density operator ρ, nor the system Hamiltonian, have matrix elements between state vectors with different total momentum. Now the operator $a^\dagger_{\mathbf{k}_{j'}}$ increases that total momentum by $\hbar k_{j'}$ and $a^H_{\mathbf{k}_j}(t')$ decreases it by $\hbar k_j$. The average value $\left\langle a^\dagger_{\mathbf{k}_{j'}} a^H_{\mathbf{k}_j}(t') \right\rangle$ is therefore zero unless j' is equal to j, and that eliminates the summation over j'. Finally:

$$\overline{G}_1^N(\mathbf{k}_i,\omega) = \int \mathrm{d}t'\ e^{i\omega t'} \left\langle a^\dagger_{\mathbf{k}_i} a^H_{\mathbf{k}_i}(t') \right\rangle \tag{34}$$

In a similar way, we can also show that:

$$\overline{G}_1^{AN}(\mathbf{k}_i,\omega) = \int \mathrm{d}t'\ e^{i\omega t'} \left\langle a^H_{\mathbf{k}_i}(t') a^\dagger_{\mathbf{k}_i} \right\rangle \tag{35}$$

2-b. Ideal gas example

For an ideal gas contained in a box (with periodic boundary conditions), the $\mathbf{k}$ are discrete. Replacing $\mathbf{k}$ by $\mathbf{k}_i$ in (28) and using expression (21), after replacing the dummy index $\mathbf{k}_i$ by $\mathbf{k}_j$, we get a product of exponentials to be integrated. The integral over d^3r', combined with the factor $1/\mathcal{V}$, yields a Kronecker delta $\delta_{\mathbf{k}_j,\mathbf{k}_i}$ that eliminates the summation over $\mathbf{k}_j$; the integral over $\mathrm{d}t'$ yields $2\pi\delta(\omega - \omega_{\mathbf{k}_i})$, and we get:

$$\overline{G}_1^N(\mathbf{k}_i,\omega) = 2\pi\ n_i\ \delta(\omega - \omega_{\mathbf{k}_i}) \tag{36}$$

In a similar way:

$$\overline{G}_1^{AN}(\mathbf{k}_i,\omega) = 2\pi\ [1 + \eta n_i]\ \delta(\omega - \omega_{\mathbf{k}_i}) \tag{37}$$

and we can finally write:

$$\overline{G}_1^{AN}(\mathbf{k}_i,\omega) - \eta\overline{G}_1^N(\mathbf{k}_i,\omega) = 2\pi\delta(\omega - \omega_{\mathbf{k}_i}) \tag{38}$$

These expressions are particularly simple, but no longer valid when the particles interact. They are, however, useful as a reference point to understand the interaction effects.

The one-particle propagator $\overline{\mathcal{G}}_1(\mathbf{k},\omega)$ defined in (29) can be obtained in a similar way; the integral over d^3r' is unchanged, but the integral over time now yields:

$$\int \mathrm{d}t'\ e^{i(\omega-\omega_{\mathbf{k}_i})t'}\ \{Y(t')\ [1+\eta n_i]+\eta Y(-t')\ n_i\} \tag{39}$$

which does not converge. For the term in $Y(t')$, a classical method is to introduce a convergence factor by changing ω into $\omega+i\varepsilon$, with $\varepsilon\to 0$ through positive values; for the $Y(-t')$ term, the change is to $\omega-i\varepsilon$. We get:

$$\begin{aligned}\overline{\mathcal{G}}_1(\mathbf{k}_i,\omega) &= -\frac{1+\eta n_i}{i\left[\omega-\omega_{\mathbf{k}_i}+i\varepsilon\right]}+\frac{\eta n_i}{i\left[\omega-\omega_{\mathbf{k}_i}-i\varepsilon\right]}\\ &= \frac{i}{\left[\omega-\omega_{\mathbf{k}_i}+i\varepsilon\right]}+\eta n_i\frac{2\varepsilon}{\left[\omega-\omega_{\mathbf{k}_i}\right]^2+\varepsilon^2}\end{aligned} \tag{40}$$

In the limit where $\varepsilon\to 0$, the two fractions on the right-hand side yield principal parts and delta functions – see relation (12) of Appendix II. The first term on the right-hand side yields both a principal part $\mathcal{P}\left[1/\left(\omega-\omega_{\mathbf{k}_i}\right)\right]$ and a delta function $\delta\left(\omega-\omega_{\mathbf{k}_i}\right)$, whereas the second one (in ηn_i) yields only a delta function $\delta\left(\omega-\omega_{\mathbf{k}_i}\right)$. If the system is diluted, it is in the classical regime (i.e. non-degenerate), where all the occupation numbers are small compared to 1 and where the exchange effects are weak; the second term, associated with the indistinguishability between the particles, is then negligible.

2-c. General expression in the presence of interactions

A translation invariant system has a Hamiltonian that commutes with its total momentum. We can then build a basis with state vectors that are, for each particle number N and for each value of the total momentum $\hbar\mathbf{K}$, eigenvectors of H with energies E; we shall note $E_{N,\mathbf{K}}$ these energies as it is often useful to explicitly keep track of the values of N and $\mathbf{K}$ that define the subspace corresponding to that eigenvalue E in the spectrum of H. We call $\left|\Phi^N_{E,K,\alpha}\right\rangle$ the corresponding eigenvectors, where the index α accounts for possible degeneracies of these eigenvalues.

We assume the system to be in a stationary state, and translation invariant. This means that its density operator ρ cannot have non-diagonal matrix elements between eigenvectors corresponding to different momenta or energies; in each of the subspaces common to both of these quantities, we can choose the basis $\left|\Phi^N_{E,K,\alpha}\right\rangle$ that diagonalizes the density operator and set:

$$\rho^N_{E,\mathbf{K},\alpha}=\left\langle\Phi^N_{E,\mathbf{K},\alpha}\right|\rho\left|\Phi^N_{E,\mathbf{K},\alpha}\right\rangle \tag{41}$$

We now use this basis to compute the trace appearing in (2). We first insert expression (12) for the field operators (and their conjugates) as a function of the operators $a_{\mathbf{k}_i}$; taking into account the exponentials introduced by the operators in the Heisenberg picture, we get:

$$\begin{aligned}&G^N_1(0,0;\mathbf{r}',t')=\tfrac{1}{\mathcal{V}}\textstyle\sum_i\sum_j e^{i\mathbf{k}_j\cdot\mathbf{r}'}\sum_{N,\mathbf{K},E,\alpha}\sum_{N',\mathbf{K}',E',\alpha'}\rho^N_{E,\mathbf{K},\alpha}\\ &\times\left\langle\Phi^N_{E,\mathbf{K},\alpha}\right|a^\dagger_{\mathbf{k}_i}\left|\Phi^{N'}_{E',\mathbf{K}',\alpha'}\right\rangle\left\langle\Phi^{N'}_{E',\mathbf{K}',\alpha'}\right|a_{\mathbf{k}_j}\left|\Phi^N_{E,\mathbf{K},\alpha}\right\rangle\ \mathrm{e}^{i\left[E'_{N',\mathbf{K}'}-E_{N,\mathbf{K}}\right]t'/\hbar}\end{aligned} \tag{42}$$

Several simplifications can be made on the right-hand side of this equation. First of all, as the operator $a_{\mathbf{k}_j}$ destroys a particle, we necessarily have $N'=N-1$, or else the

matrix element of $a_{\mathbf{k}_j}$ would be zero; the sum over N' disappears. Along the same line, as this operator decreases the total momentum by $\hbar\mathbf{k}_j$, we must also have $\mathbf{K}' = \mathbf{K} - \mathbf{k}_j$, and the sum over $\mathbf{K}'$ is also eliminated. Now, for the matrix element of $a^\dagger_{\mathbf{k}_i}$ to be non-zero, and since this operator adds $\hbar\mathbf{k}_i$ to the momentum $\hbar\,(\mathbf{K} - \mathbf{k}_j)$, to recover the initial momentum $\hbar\mathbf{K}$, we necessarily have $\mathbf{k}_i = \mathbf{k}_j$. Once these simplifications have been made, we insert the result in definition (28) of $\overline{G}_1^N(\mathbf{k},\omega)$, and get:

$$\begin{array}{l}\overline{G}_1^N(\mathbf{k},\omega) = \frac{1}{\mathcal{V}}\int \mathrm{d}^3r' \int \mathrm{d}t'\ e^{i\left(\omega t' - \mathbf{k}\cdot\mathbf{r}'\right)} \sum_j e^{i\mathbf{k}_j\cdot\mathbf{r}'} \\ \sum_{N,\mathbf{K}} \sum_{E,E',\alpha,\alpha'} \rho^N_{E,\mathbf{K},\alpha}\ \left|\left\langle \Phi^N_{E,\mathbf{K},\alpha}\right| a^\dagger_{\mathbf{k}_i}\left|\Phi^{N-1}_{E',\mathbf{K}-\mathbf{k},\alpha'}\right\rangle\right|^2\ \mathrm{e}^{i\left[E'_{N-1,\mathbf{K}-\mathbf{k}} - E_{N,\mathbf{K}}\right]t'/\hbar}\end{array} \tag{43}$$

On the first line, the integral over d^3r' yields a Kronecker delta that forces $\mathbf{k}_j = \mathbf{k}$; the integral then yields $\mathcal{V}$, which cancels the volume appearing in the denominator. The time integral in relation (43) yields a 2π coefficient multiplied by a delta function of the variable $\omega + \left[E'_{N-1,\mathbf{K}-\mathbf{k}} - E_{N,\mathbf{K}}\right]/\hbar$. We finally get:

$$\begin{array}{l}\overline{G}_1^N(\mathbf{k},\omega) = \\ 2\pi \sum_{N,\mathbf{K}} \sum_{E,E',\alpha,\alpha'} \rho^N_{E,\mathbf{K},\alpha}\ \left|\left\langle \Phi^N_{E,\mathbf{K},\alpha}\right| a^\dagger_{\mathbf{k}}\left|\Phi^{N-1}_{E',\mathbf{K}-\mathbf{k},\alpha'}\right\rangle\right|^2\ \delta\left[\omega - \frac{E_{N,\mathbf{K}} - E'_{N-1,\mathbf{K}-\mathbf{k}}}{\hbar}\right]\end{array} \tag{44}$$

The same type of calculation also yields:

$$\begin{array}{l}\overline{G}_1^{AN}(\mathbf{k},\omega) = \\ 2\pi \sum_{N,\mathbf{K}} \sum_{E,E',\alpha,\alpha'} \rho^N_{E,\mathbf{K},\alpha}\ \left|\left\langle \Phi^N_{E,\mathbf{K},\alpha}\right| a_{\mathbf{k}}\left|\Phi^{N+1}_{E',\mathbf{K}+\mathbf{k},\alpha'}\right\rangle\right|^2\ \delta\left[\omega - \frac{E'_{N+1,\mathbf{K}+\mathbf{k}} - E_{N,\mathbf{K}}}{\hbar}\right]\end{array} \tag{45}$$

Let us assume, in addition, that our system is at thermal equilibrium, and described by the grand canonical density operator:

$$\rho = \frac{1}{Z} e^{-\beta\left(H - \mu\widehat{N}\right)} \tag{46}$$

with the classical notation: $\widehat{N}$ is the total particle number, $\beta = 1/k_BT$ (k_B is the Boltzmann constant) and $Z = \mathrm{Tr}\left\{e^{-\beta\left(H-\mu\widehat{N}\right)}\right\}$ is the grand canonical partition function. The two functions $\overline{G}_1^N$ and $\overline{G}_1^{AN}$ then obey a simple relation, often called "boundary condition":

$$\boxed{\overline{G}_1^{AN}(\mathbf{k},\omega) = e^{\beta(\hbar\omega-\mu)}\overline{G}_1^N(\mathbf{k},\omega)} \tag{47}$$

This relation turns out to be crucial in many calculations involving Green's functions; we shall use it in § 3.

Demonstration:

To establish this relation, we rewrite equality (45) using the fact that $a_{\mathbf{k}}$ and $a_{\mathbf{k}}^{\dagger}$ are Hermitian conjugates. We then get:

$$\overline{G}_1^{AN}(\mathbf{k},\omega) = 2\pi \sum_{N,\mathbf{K}} \sum_{E,E',\alpha,\alpha'} \rho^N_{\mathbf{K},E,\alpha} \times \\ \times \left| \left\langle \Phi^{N+1}_{E',\mathbf{K}+\mathbf{k},\alpha'} \right| a_{\mathbf{k}}^{\dagger} \left| \Phi^N_{E,\mathbf{K},\alpha} \right\rangle \right|^2 \delta \left[\omega - \frac{E'_{N+1,\mathbf{K}+\mathbf{k}} - E_{N,\mathbf{K}}}{\hbar} \right] \tag{48}$$

We now permute the indices E' and E, as well as the indices α and α', we change the dummy summation index N to $N' = N + 1$, and finally replace the dummy variable K by $K' = K + k$. This yields:

$$\overline{G}_1^{AN}(\mathbf{k},\omega) = 2\pi \sum_{N',\mathbf{K}'} \sum_{E,E',\alpha,\alpha'} \rho^{N'-1}_{\mathbf{K}'-\mathbf{k},E',\alpha'} \times \\ \times \left| \left\langle \Phi^{N'}_{E,\mathbf{K}',\alpha} \right| a_{\mathbf{k}}^{\dagger} \left| \Phi^{N'-1}_{E',\mathbf{K}'-\mathbf{k},\alpha'} \right\rangle \right|^2 \delta \left[\omega - \frac{E_{N',\mathbf{K}'} - E'_{N'-1,\mathbf{K}'-\mathbf{k}}}{\hbar} \right] \tag{49}$$

In this summation, just as in (44), the lowest value of the index N' that gives a contribution is $N' = 1$; we therefore get the same expression as (44), with (aside from the irrelevant change of the dummy index N into N') just one modification, the replacement of $\rho^N_{\mathbf{K},E,\alpha}$ by $\rho^{N-1}_{\mathbf{K}'-\mathbf{k},E',\alpha'}$. However, since:

$$\rho^N_{\mathbf{K},E,\alpha} = \frac{1}{Z} e^{-\beta\left[E_{N,\mathbf{K}} - \mu N\right]} \tag{50}$$

the ratio of these two diagonal elements $\rho^{N-1}_{\mathbf{K}'-\mathbf{k},E',\alpha'} / \rho^N_{\mathbf{K},E,\alpha}$ introduces in the integral a factor:

$$e^{\beta\left[-E'_{N-1,\mathbf{K}'-\mathbf{k}} + \mu(N-1) + E_{N,\mathbf{K}} - \mu N\right]} = e^{\beta\left[E_{N,\mathbf{K}} - E'_{N-1,\mathbf{K}'-\mathbf{k}} - \mu\right]} \tag{51}$$

Now, the delta function in (49) allows replacing, in the exponent of this factor, the energy difference by $\hbar\omega$:

$$e^{\beta\left[E_{N,\mathbf{K}} - E_{N-1,\mathbf{K}'-\mathbf{k}} - \mu\right]} \Longrightarrow e^{\beta[\hbar\omega - \mu]} \tag{52}$$

This factor comes out of the summation and relation (47) is established.

2-d. Discussion

Expressions (44) and (45) give an idea of the behavior of the functions $\overline{G}_1^N$ and $\overline{G}_1^{AN}$. For an ideal gas, relations (36) and (37) show that they are singular functions, actually delta functions forcing the energy $\hbar\omega$ to take on exactly the one-particle kinetic energy $\hbar^2 k^2/2m$. This is because, in an ideal gas, one can choose Fock states as stationary states $\left|\Phi^N_{E,\mathbf{K},\alpha}\right\rangle$, where each individual state, with a given momentum $\hbar\mathbf{k}$, has a well defined population. In such a case, the operator $a_{\mathbf{k}}$ in expression (45) of $\overline{G}_1^{AN}$ can only couple, via its matrix element $\left\langle \Phi^N_{E,\mathbf{K},\alpha} \right| a_{\mathbf{k}} \left| \Phi^{N+1}_{E',\mathbf{K}+\mathbf{k},\alpha'} \right\rangle$, one state $\left|\Phi^{N+1}_{E',\mathbf{K}+\mathbf{k},\alpha'}\right\rangle$ to the initial state $\left|\Phi^N_{E,\mathbf{K},\alpha}\right\rangle$: the Fock state, where the occupation number $n_{\mathbf{k}}$ is larger by one unit. Consequently, the energy difference $E'_{N+1,\mathbf{K}+\mathbf{k}} - E_{N,\mathbf{K}}$ always takes the same value, the energy $\hbar^2 k^2/2m$ of an individual particle with momentum $\hbar\mathbf{k}$. We find again the results of § 2-b.

Let us assume now that the system includes mutual interactions. The matrix element $\left\langle \Phi^N_{E,\mathbf{K},\alpha} \right| a_{\mathbf{k}} \left| \Phi^{N+1}_{E',\mathbf{K}+\mathbf{k},\alpha'} \right\rangle$ then yields the scalar product of an N-particle system

stationary state with another state where a particle with a well defined momentum $\hbar\mathbf{k}$ has been removed from a stationary state of a system with $N+1$ particles. This new state has no reason to also be stationary, because of the particle interactions. It is therefore expected to have a non-zero scalar product with a whole series of kets $\left|\Phi^N_{E,\mathbf{K},\alpha}\right\rangle$ with different energies, which can become more numerous as the system enlarges. The sum over E no longer reduces to a single value of the energy; in the limit of large systems, it becomes a continuous sum, which absorbs the delta function, so that the function $\overline{G}_1^{AN}(\mathbf{k},\omega)$ takes, for a given $\mathbf{k}$, non-negligible values in a whole ω domain. A priori, its variations can take any value in this domain; however, if the interactions are not too strong and by comparison with the ideal gas, we expect the scalar product to have non-negligible values mostly in an E' energy domain close to $E_{N,\mathbf{K}}+\hbar^2k^2/2m$. For each $\mathbf{k}$, the interaction effect is to enlarge the energy peak, infinitely narrow for an ideal gas, by giving it a width that increases as the interaction becomes stronger. This width is interpreted in terms of a finite lifetime of the excitation created when a free particle is removed from a stationary state of the $(N+1)$-particle system; as the moduli of the two matrix elements $\left\langle\Phi^N_{E,\mathbf{K},\alpha}\right| a_{\mathbf{k}} \left|\Phi^{N+1}_{E',\mathbf{K}+\mathbf{k},\alpha'}\right\rangle$ and $\left\langle\Phi^{N+1}_{E',\mathbf{K}+\mathbf{k},\alpha'}\right| a^\dagger_{\mathbf{k}} \left|\Phi^N_{E,\mathbf{K},\alpha}\right\rangle$ are equal, this lifetime can also be interpreted as that of the excitation created when a free particle is added to a stationary state of the N-particle system.

3. Spectral function, sum rule

In view of formula (7), it is natural to introduce the real function $A(\mathbf{k},\omega)$ as:

$$\boxed{\begin{aligned} A(\mathbf{k},\omega) &= \overline{G}_1^{AN}(\mathbf{k},\omega) - \eta\overline{G}_1^{N}(\mathbf{k},\omega) && \text{(53a)}\\ &= \int \mathrm{d}t'\ e^{i\omega t'} \left\langle \left[a^H_{\mathbf{k}_i}(t'), a^\dagger_{\mathbf{k}_i}\right]_{-\eta} \right\rangle && \text{(53b)} \end{aligned}}$$

where we have used (34) and (35). We call $A(\mathbf{k},\omega)$ the "spectral function"; it is real since both functions $\overline{G}_1^N$ and $\overline{G}_1^{AN}$ are real.

For an ideal gas, formula (38) shows that we simply have:

$$A(\mathbf{k},\omega) = 2\pi\ \delta(\omega-\omega_k) \tag{54}$$

but this equality is no longer valid as soon as the particles interact. We are going to show, however, that the spectral function allows expressing the correlation functions with a general formula, reminiscent of that established for an ideal gas, even in the presence of interactions.

3-a. Expression of the one-particle correlation functions

Inserting (47) in the definition of $A(\mathbf{k},\omega)$, we get:

$$A(\mathbf{k},\omega) = \overline{G}_1^N \left[e^{\beta(\hbar\omega-\mu)} - \eta\right] \tag{55}$$

that is:

$$\overline{G}_1^N(\mathbf{k},\omega) = A(\mathbf{k},\omega)\ f^\eta_\beta(\hbar\omega-\mu) \tag{56}$$

where $f_\beta^\eta(\hbar\omega - \mu)$ is the usual Bose-Einstein distribution for bosons, and the Fermi-Dirac distribution for fermions:

$$f_\beta^\eta(\hbar\omega - \mu) = \frac{1}{e^{\beta(\hbar\omega-\mu)} - \eta} \tag{57}$$

Using again (47), we get:

$$\overline{G}_1^{AN}(\mathbf{k},\omega) = e^{\beta(\hbar\omega-\mu)} A(\mathbf{k},\omega)\ f_\beta^\eta(\hbar\omega-\mu) = A(\mathbf{k},\omega)\ \left[1 + \eta f_\beta^\eta(\hbar\omega-\mu)\right] \tag{58}$$

Knowing the spectral function $A(\mathbf{k},\omega)$, we can determine the one-particle correlation functions with formulas similar to those of the ideal gas, containing the same quantum distribution functions f_β^η. Note, however, that in the presence of interactions, the energy $\hbar\omega$ and momentum $\hbar\mathbf{k}$ variables are independent, whereas for an ideal gas, they are constrained by relation (54).

3-b. Sum rule

We now insert relations (28) and (2) in (53a); since the operators Ψ and Ψ_H coincide at $t = 0$, we get:

$$A(\mathbf{k},\omega) = \int \mathrm{d}^3 r' \int \mathrm{d}t'\ e^{i(\omega t' - \mathbf{k}\cdot\mathbf{r}')}\ \mathrm{Tr}\left\{\rho\ \left[\Psi_H(\mathbf{r}',t')\ ,\ \Psi^\dagger(\mathbf{r}=\mathbf{0})\right]_{-\eta}\right\} \tag{59}$$

Taking a summation over ω, a time delta function is introduced:

$$\int \mathrm{d}\omega\ A(\mathbf{k},\omega) = 2\pi \int \mathrm{d}^3 r'\ e^{-i\mathbf{k}\cdot\mathbf{r}'}\ \mathrm{Tr}\left\{\rho\ \left[\Psi(\mathbf{r}')\ ,\ \Psi^\dagger(\mathbf{r}=\mathbf{0})\right]_{-\eta}\right\} \tag{60}$$

that is:

$$\int \frac{\mathrm{d}\omega}{2\pi}\ A(\mathbf{k},\omega) = \int \mathrm{d}^3 r'\ e^{-i\mathbf{k}\cdot\mathbf{r}'}\ \delta(\mathbf{r}') = 1 \tag{61}$$

In the presence of interactions, we do not know, a priori, how the spectral function depends on $\mathbf{k}$ and ω. However, for each value of $\mathbf{k}$, the interactions can only modify the frequency distribution, but not its integral over ω.

We have seen that for a gas of interacting particles, the effect of the interactions is to "spread" the spectral function over a certain domain of frequencies, all the while obeying the sum rule (61). There is no reasons for the spectral function to present any particular shape, or still contain delta functions of frequency. It often happens, however, that it presents marked peaks, whose narrowness signals the existence of excitations in the system, behaving almost like free particles (with long lifetimes). These excitations are called "quasi-particles" as they are, in a way, an extension – once the interactions have been introduced – of the independent particles of the ideal gas. The peak associated with a particle of momentum $\hbar\mathbf{k}$ is not, in general, centered on the energy $e_k = \hbar^2 k^2/2m$ of a free particle: in addition to the spreading, the energies of the quasi-particles are shifted by the interactions. This spreading and shifting is reminiscent of the results obtained in Complement D_{XIII}, where we studied, to lowest order, the coupling of a discrete state with a continuum. In other words, one can say that the interactions couple the state of a free particle with momentum $\hbar\mathbf{k}$, to a continuum of states having different momenta,

which explains the analogy. Note, however, that the present results concern not a single but an ensemble of identical particles, and its properties at thermal equilibrium; the physical situation is therefore different.

To sum up, to go from the ideal gas (where ω is necessarily equal to ω_k) to an interacting gas, we simply introduce in the Green's function, and for each value of $\mathbf{k}$, a weighting by a spectral function $A(\mathbf{k},\omega)$; this function distributes the ω dependence over a certain frequency domain. We are going to show that from the spectral function we can infer many properties of an interacting physical system. But this obviously does not mean that the spectral function is easy to compute! On the contrary, in most interacting physical systems, we do not know its exact value. Its very existence, however, independently of its precise mathematical form, is a very useful conceptual tool.

3-c. Expression of various physical quantities

The spectral function contains information on a great number of physical properties of interacting systems, in a form much more concise than the density operator of the N-body system: that operator contains everything but is mathematically far more complicated than a simple function.

Consider first the particle density, given by:

$$n(\mathbf{r}) = \left\langle \Psi^\dagger(\mathbf{r})\Psi(\mathbf{r}) \right\rangle = G_1^N(\mathbf{r},0;\mathbf{r},0) \tag{62}$$

according to the definition (2) of the normal function G_1^N. For a translation invariant system, this density is independent of $\mathbf{r}$, and we can set $\mathbf{r}=0$. The density is then the value, at the origin, of the normal function (31), that is, taking (56) into account:

$$n = \int \frac{\mathrm{d}^3k}{(2\pi)^3} \int \frac{\mathrm{d}\omega}{2\pi}\, \overline{G}_1^N(\mathbf{k},\omega) = \int \frac{\mathrm{d}^3k}{(2\pi)^3} \int \frac{\mathrm{d}\omega}{2\pi}\, A(\mathbf{k},\omega)\; f_\beta^\eta(\hbar\omega-\mu) \tag{63}$$

Let us now study a quantity furnishing more precise information, the particle momentum distribution, and compute the average number $\langle n_\mathbf{k}\rangle$ of particles having a momentum $\hbar\mathbf{k}$. We assume the system is contained in a cubic box of volume $\mathcal{V}$. Relations (A-9) and (A-10) of Chapter XVI, applied to the case where the basis wave functions are given by (12), yield:

$$\langle n_\mathbf{k}\rangle = \left\langle a_\mathbf{k}^\dagger a_\mathbf{k} \right\rangle = \frac{1}{\mathcal{V}} \int \mathrm{d}^3r' \int \mathrm{d}^3r\; e^{i\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r})} \left\langle \Psi^\dagger(\mathbf{r}')\Psi(\mathbf{r}) \right\rangle \tag{64}$$

Replacing the integral variable $\mathbf{r}$ by $\mathbf{s}=\mathbf{r}-\mathbf{r}'$, the average value $\left\langle \Psi^\dagger(\mathbf{r}')\Psi(\mathbf{r}'+\mathbf{s}) \right\rangle$ appears, which is independent of $\mathbf{r}'$ because of the translation invariance; we can then replace, in this average value, $\mathbf{r}'$ by zero, and the integral over d^3r' simply yields the volume $\mathcal{V}$. We are left with:

$$\langle n_\mathbf{k}\rangle = \int \mathrm{d}^3s\; e^{-i\mathbf{k}\cdot\mathbf{s}} \left\langle \Psi^\dagger(\mathbf{0})\Psi(\mathbf{s}) \right\rangle \tag{65}$$

Now, performing the integral over ω of the definition (28) of $\overline{G}_1^N(\mathbf{k},\omega)$, and taking (2)

into account, we get:

$$\int \mathrm{d}\omega\ \overline{G}_1^N(\mathbf{k},\omega) = 2\pi \int \mathrm{d}^3 r' e^{-i\mathbf{k}\cdot\mathbf{r}'} G_1^N(0,0;\mathbf{r}',0)$$
$$= 2\pi \int \mathrm{d}^3 r' e^{-i\mathbf{k}\cdot\mathbf{r}'} \left\langle \Psi^\dagger(\mathbf{0})\Psi(\mathbf{r}')\right\rangle \tag{66}$$

which is identical to (65) within a factor of 2π. It then follows that[7]:

$$\langle n_{\mathbf{k}}\rangle = \frac{1}{2\pi}\int \mathrm{d}\omega\ \overline{G}_1^N(\mathbf{k},\omega) = \int \frac{\mathrm{d}\omega}{2\pi} A(\mathbf{k},\omega)\ f_\beta^\eta(\hbar\omega-\mu) \tag{67}$$

In the same way, one can show that the system average energy[8] per unit volume $\langle H\rangle/\mathcal{V}$ is given by:

$$\frac{\langle H\rangle}{\mathcal{V}} = \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\int \frac{\mathrm{d}\omega}{2\pi}\left[\hbar\frac{\omega+\omega_k}{2}\right]\ A(\mathbf{k},\omega) f_\beta^\eta(\hbar\omega-\mu) \tag{68}$$

where ω_k is defined in (14). Once this function is known, one can get, by integration over β, the logarithm of the partition function, which in turn yields, by derivation, all the thermodynamic quantities (particle density, pressure, etc.). It is remarkable that the spectral function, whose definition comes from the one-particle Green's functions and could therefore be expected to only contain information on the one-particle density operator, actually allows computing all these physical quantities that depend on the correlations between the particles, and hence on their interactions. With this method, the study of N-body properties is reduced to the computation of functions mathematically defined for a single particle. It generalizes, in a way, the ideal gas equations, while taking rigorously into account the presence of an ensemble of indistinguishable particles at thermal equilibrium. It is therefore quite powerful.

Nevertheless, this obviously does not solve the problem of calculating the equilibrium properties of an interacting system; in practice, getting precise values for the spectral function can pose a very difficult mathematical problem. Numerous approximation methods have been developed to try and resolve this, using in particular the concept of "self energy" and of perturbation diagrams, but this is beyond the scope of this complement.

[7] For an ideal gas of bosons, the chemical potential is always below the lowest individual energy, so that the distribution function f_β^η never diverges. As in relation (67) ω is integrated from $-\infty$ to $+\infty$, this divergence now seems unavoidable. Relation (55), however, shows that the spectral function $A(\mathbf{k},\omega)$ for bosons goes to zero for $\hbar\omega=\mu$, as long as the function $\overline{G}_1^N$ remains regular at this point. Consequently, integrals (63), (67) and (68) do not present any divergences associated with that ω value

[8] Relation (68) can be demonstrated by first computing the time evolution of $\Psi_H(\mathbf{r},t)$ and $\Psi_H^\dagger(\mathbf{r},t')$ to get the expression for $i\hbar\left(\frac{\partial}{\partial t}-\frac{\partial}{\partial t'}\right)\Psi_H^\dagger(\mathbf{r},t')\Psi_H(\mathbf{r},t)$; one then takes its average value and performs a Fourier transformation to get the average value of the energy – see for example § 2.2 of reference [7].

Complement C_{XVI}

Wick's theorem

1 Demonstration of the theorem 1799
1-a Statement of the problem 1799
1-b Recurrence relation 1800
1-c Contractions 1802
1-d Statement of the theorem 1804
2 Applications: correlation functions for an ideal gas 1804
2-a First order correlation function 1805
2-b Second order correlation functions 1806
2-c Higher order correlation functions 1808

For an ideal gas at thermal equilibrium, we computed in Complement B_{XV} the average values of one- and two-particle operators, and showed they could all be expressed in terms of the one-body quantum distributions f_β (the Fermi-Dirac distribution for fermions, and Bose-Einstein for bosons). In this complement we establish a theorem that allows generalizing those results to operators involving any number of particles. The demonstration of Wick's theorem is explained in § 1, and will be applied in § 2 to the calculation of correlation functions in an ideal gas.

1. Demonstration of the theorem

Let us consider an ideal gas at thermal equilibrium, described by the grand canonical ensemble (Appendix VI, § 1-c), with the density operator:

$$\rho_{eq} = \frac{1}{Z} e^{-\beta\left(\widehat{H}-\mu\widehat{N}\right)} \tag{1}$$

where $\beta = 1/k_B T$ is the inverse of the temperature multiplied by the Boltzmann constant k_B, μ the chemical potential, and $\widehat{H}$ the Hamiltonian:

$$\widehat{H} = \sum_{\mathbf{k}} e_k \; a_k^\dagger a_k \qquad \text{with:} \qquad e_k = \frac{\hbar^2 k^2}{2m} \tag{2}$$

The grand canonical partition function Z is defined by:

$$Z = \mathrm{Tr}\left\{ e^{-\beta\left(\widehat{H}-\mu\widehat{N}\right)} \right\} \tag{3}$$

1-a. Statement of the problem

We wish to calculate the average value $\langle A_P \rangle$ of a product of operators:

$$A_P = b_1 b_2 .... b_P \tag{4}$$

where each of the operators b_k is, either an annihilation operator a_k, or a creation operator $a_k^\dagger$:

$$b_k = \begin{cases} a_k \\ a_k^\dagger \end{cases} \tag{5}$$

Taking into account relations (A-48) and (A-49) of Chapter XV, we have:

$$\begin{aligned} [b_i, b_j]_{-\eta} &= b_i b_j - \eta b_j b_i \\ &= \begin{cases} 0 \text{ if } b_i \text{ and } b_j \text{ are both creation or annihilation operators} \\ \delta_{ij} \text{ if } b_i \text{ is an annihilation, and } b_j \text{ a creation operator} \\ -\eta\delta_{ij} \text{ if } b_i \text{ is a creation, and } b_j \text{ an annihilation operator} \end{cases} \end{aligned} \tag{6}$$

with $\eta = +1$ for bosons and $\eta = -1$ for fermions.

Assuming the quantum state is described by the density operator ρ_{eq} given in (1), the average value of A_P is:

$$\langle A_P \rangle = \mathrm{Tr}\,\{\rho_{eq}\ b_1 b_2 .... b_P\} \tag{7}$$

As ρ_{eq} is diagonal in the basis of the Fock states associated with the a_k, this average value will be different from zero only if the series of operators b contains, for each creation operator $a_k^\dagger$, an annihilation operator a_k in that same individual state; they must exactly balance one another and must therefore appear the same number of times. In particular, the average value $\langle A_P \rangle$ will always be zero if P is odd; from now on, we shall assume that $P = 2n$, n being an integer.

1-b. Recurrence relation

We have to compute:

$$\langle A_{2n} \rangle = \mathrm{Tr}\,\{\rho_{eq}\ b_1 b_2 .... b_{2n}\} \tag{8}$$

We first start by changing the order of b_1 and b_2, using one of the relations (6); we will then continue to progressively shift b_1 towards the right, by permuting it first with b_3, then with b_4, etc. until the permutation with b_{2n} brings it to the very last position. As a trace is invariant under a circular permutation, the operator b_1 can then be moved back all the way to the first position, ahead of ρ_{eq}; a last commutation with ρ_{eq}, that we compute just below, returns it to its initial position, and allows computing the value of $\langle A_{2n} \rangle$ as a function of the average values of a product of $2(n-1)$ operators. The computation goes as follows:

$$\begin{aligned} \langle A_{2n} \rangle &= \mathrm{Tr}\left\{\rho_{eq}\ \ [b_1, b_2]_{-\eta}\, b_3 b_4 ... b_{2n}\right\} + \eta\ \mathrm{Tr}\,\{\rho_{eq}\ b_2 b_1 b_3 b_4 ... b_{2n}\} \\ &= [b_1, b_2]_{-\eta}\ \mathrm{Tr}\,\{\rho_{eq}\ b_3 b_4 ... b_{2n}\} + \eta\,[b_1, b_3]_{-\eta}\ \mathrm{Tr}\,\{\rho_{eq}\ b_2 b_4 ... b_{2n}\} \\ &\qquad + \mathrm{Tr}\,\{\rho_{eq}\ b_2 b_3 b_1 b_4 ... b_{2n}\} \\ &= [b_1, b_2]_{-\eta}\ \mathrm{Tr}\,\{\rho_{eq}\ b_3 b_4 ... b_{2n}\} + \eta\,[b_1, b_3]_{-\eta}\ \mathrm{Tr}\,\{\rho_{eq}\ b_2 b_4 ... b_{2n}\} + ... \\ &\qquad ... + \eta^{2n-2}\,[b_1, b_{2n}]_{-\eta}\ \mathrm{Tr}\,\{\rho_{eq}\ b_2 b_3 b_4 ... b_{2n-1}\} \\ &\qquad\qquad + \eta^{2n-1}\ \mathrm{Tr}\,\{\rho_{eq}\ b_2 b_3 b_4 ... b_{2n} b_1\} \end{aligned} \tag{9}$$

Most of the terms on the right-hand side are in general zero: the first is non-zero only if b_1 and b_2 are two conjugated operators (an annihilation and a creation operator associated with the same individual quantum state); the second is non-zero if this is also the case for operators b_1 and b_3, etc.

After a circular permutation under the trace, the last term of the sum can be written as:

$$\eta^{2n-1}\mathrm{Tr}\left\{b_1\rho_{eq}\ b_2b_3....b_{2n}\right\} \tag{10}$$

We shall now relate both operators $b_1\rho_{eq}$ and $\rho_{eq}b_1$, showing that they are proportional to each other. Assume, for example, that b_1 is a creation operator $a_q^\dagger$; in the operator ρ_{eq} defined in (1), all the terms $k \neq q$ commute with $a_q^\dagger$ and the change of order for the operators leads to two expressions, to be compared:

$$a_q^\dagger\ e^{-\beta(e_q-\mu)\ a_q^\dagger a_q} \qquad \text{and} \quad e^{-\beta(e_q-\mu)\ a_q^\dagger a_q}\ a_q^\dagger \tag{11}$$

By action on the Fock vectors, it is easy to check (as we assumed the system was in thermal equilibrium) that:

$$a_q^\dagger\ e^{-\beta(e_q-\mu)\ a_q^\dagger a_q} = e^{\beta(e_q-\mu)} \times e^{-\beta(e_q-\mu)\ a_q^\dagger a_q}\ a_q^\dagger \tag{12}$$

which leads to:

$$a_q^\dagger\ \rho_{eq} = e^{\beta(e_q-\mu)}\rho_{eq}\ a_q^\dagger \tag{13}$$

If b_1 is an annihilation operator, the same reasoning shows that the change of order introduces the inverse factor: $e^{-\beta(e_n-\mu)}$. To sum up:

$$b_1\rho_{eq} = e^{\pm\beta(e_1-\mu)}\rho_{eq}b_1 \tag{14}$$

with a + sign in the exponential if b_1 is a creation operator, and a − sign if b_1 is an annihilation operator. Consequently, the last term in (9) is equal to $e^{\pm\beta(e_1-\mu)}\left\langle A_{2n}\right\rangle$, with a factor $\eta^{2n-1} = \eta$ since $\eta = \pm 1$.

Moving this last term to the left-hand side, we get:

$$\begin{aligned}\left\langle A_{2n}\right\rangle\left[1-\eta e^{\pm\beta(e_1-\mu)}\right] &= \left[b_1,b_2\right]_{-\eta}\mathrm{Tr}\left\{\rho_{eq}\ b_3...b_{2n}\right\}\\ &+\eta\left[b_1,b_3\right]_{-\eta}\mathrm{Tr}\left\{\rho_{eq}\ b_2b_4...b_{2n}\right\}+...\\ &+\eta^{2n-2}\left[b_1,b_{2n}\right]_{-\eta}\mathrm{Tr}\left\{\rho_{eq}\ b_2b_3...b_{2n-1}\right\}\end{aligned} \tag{15}$$

On the right-hand side of this equality, all the (anti)commutators $\left[b_1,b_2\right]_{-\eta}$ are actually numbers, and many of them are zero: as before, the only non-zero ones are those for which the two concerned operators b are conjugates of each other (a creation and an annihilation operator for the same individual quantum state). The average value of the product of $2n$ operators we are looking for can therefore be expressed as a linear combination of average values of products of $2n-2$ operators.

1-c. Contractions

We now define the "contraction" of two operators b_i and b_j as the number, written $\overbrace{b_i b_j}$, defined by:

$$\overbrace{b_i b_j} = \frac{1}{1-\eta e^{\pm\beta(e_i-\mu)}} \times [b_i, b_j]_{-\eta} = -\eta f_\beta\left(\pm(e_i-\mu)\right) \times [b_i, b_j]_{-\eta} \tag{16}$$

where, as above, in the denominator a + sign is chosen in the exponential if b_1 is a creation operator, and a − sign if b_1 is an annihilation operator. The function f_β is the Fermi-Dirac distribution f_β^{FD} for fermions, and that of Bose-Einstein f_β^{BE} for bosons:

$$f_\beta(e-\mu) = \frac{1}{e^{\beta(e-\mu)}-\eta} \tag{17}$$

The contraction is zero if it concerns two operators b_i and b_j acting on different individual quantum states; it is also zero if the operators are both creation or both annihilation operators in the same individual quantum state. If b_i is the creation operator, and b_j the annihilation operator in the same individual state, the contraction is simply equal to the distribution function $f_\beta(e_i-\mu)$ since:

$$\overbrace{b_i b_j} = \overbrace{a_i^\dagger a_i} = \frac{-\eta}{1-\eta e^{+\beta(e_i-\mu)}} = f_\beta(e_i-\mu) \tag{18}$$

In the opposite case (antinormal order), the contraction is given by:

$$\overbrace{b_i b_j} = \overbrace{a_i a_i^\dagger} = \frac{1}{1-\eta e^{-\beta(e_i-\mu)}} = 1+\eta f_\beta(e_i-\mu) \tag{19}$$

Relation (15) can thus be rewritten as:

$$\langle A_{2n}\rangle = \overbrace{b_1 b_2}\langle\, b_3 ... b_{2n}\rangle + \eta\, \overbrace{b_1 b_3}\langle b_2 b_4 ... b_{2n}\rangle + ... + \eta^{2n-2}\overbrace{b_1 b_{2n}}\langle b_2 b_3 ... b_{2n-1}\rangle \tag{20}$$

where the traces have been replaced with quantum averages.

We shall then reason by recurrence: each of the average values on the right-hand side of (20) is of the same type as $\langle A_{2n}\rangle$ written in (8), except for the fact that n has been lowered by one unit. Dealing with each of the average values $\langle A_{2n-2}\rangle$ as we did for $\langle A_{2n}\rangle$, that last average value now appears as a double sum of terms containing two contractions and average values $\langle A_{2n-4}\rangle$. Continuing as many times as necessary, we end up with an average value $\langle A_{2n}\rangle$ expressed as the sum of diverse products of n contractions.

As an illustration, let us consider a few simple examples. If $n=1$, we get directly:

$$\boxed{\langle b_1 b_2\rangle = \overbrace{b_1 b_2}} \tag{21}$$

This simple relation can actually be used as a definition of contractions, instead of (16). If b_1 is a creation operator and b_2 the corresponding annihilation operator, we get the result (18), equivalent to relations (19) and (23) of Complement B_{XV}; if the operator's order is reversed, we get (19) which comes directly from the previous result and from the

commutation or anticommutation relation (6). For all the other cases, we find zero on each side of the equality.

If $n = 2$, we use a first time relation (20), and obtain:

$$\langle b_1 b_2 b_3 b_4 \rangle = \overbrace{b_1 b_2} \langle b_3 b_4 \rangle + \eta\, \overbrace{b_1 b_3} \langle b_2 b_4 \rangle + \overbrace{b_1 b_4} \langle b_2 b_3 \rangle \tag{22}$$

Using again this same relation, we compute each of the average values of the product of two operators b, which yields:

$$\begin{aligned} \langle b_1 b_2 b_3 b_4 \rangle &= \overbrace{b_1 b_2}\, \overbrace{b_3 b_4} + \eta\, \overbrace{b_1 b_3}\, \overbrace{b_2 b_4} + \overbrace{b_1 b_4}\, \overbrace{b_2 b_3} \\ &= \overbrace{b_1 b_2 b_3 b_4} + \overbrace{b_1 \underbrace{b_2 b_3} b_4} + \overbrace{b_1 \underbrace{b_2 b_3 b_4}} \end{aligned} \tag{23}$$

In the second line, we have used a generalization of the notation of products of contractions. When two operators inside a contraction are separated, a permutation is needed to group them. For fermions, this introduces a sign given by the parity of the required permutation, but no sign change for bosons. When two contractions are embedded, we group together all pairs of operators belonging to the same contraction and, for fermions, we multiply by the parity of the corresponding permutation[1]; for bosons, no sign change is introduced. In the present case, we therefore have :

$$\overbrace{b_1 \underbrace{b_2 b_3} b_4} = \eta\, \overbrace{b_1 b_3}\, \overbrace{b_2 b_4} \qquad \text{and} \qquad \overbrace{b_1 \underbrace{b_2 b_3 b_4}} = \overbrace{b_1 b_4}\, \overbrace{b_2 b_3} \tag{24}$$

The final result (23) only contains products of contractions, i.e. of distribution functions. One can easily check that, among those three products, a maximum of two are non-zero.

Comments:

(i) Another notation is frequently used, where operators and contractions are embedded in the same average value, for instance:

$$\langle b_1 b_2 ... \overbrace{b_i ... b_j ... b_l} ... b_{2n} \rangle = (\varepsilon_\alpha)\, \overbrace{b_i b_l} \langle b_1 b_2 ... b_j ... b_{2n} \rangle \tag{25}$$

where, for fermions, ε_α is the parity of the permutation needed to bring operator b_l next to b_i; for bosons, $\varepsilon_\alpha = 1$. This can be generalized to cases where several contraction appear, embedded or not.

(ii) In the limit of zero temperature where $\beta \to \infty$, relation (16) simplifies into:

$$\begin{aligned} &\overbrace{b_i b_j} = 0 \text{ if the two operators are of the same nature (both creation,} \\ &\qquad\qquad \text{or both annihilation operators)} \end{aligned} \tag{26}$$

as well as:

$$\overbrace{a_i a_j^\dagger} = \delta_{ij} \times \begin{cases} 1 \text{ si } e_i > \mu \\ 0 \text{ si } e_i < \mu \end{cases} \tag{27}$$

[1]We multiply by -1 if, when writing the permutation, the number of crossings between brackets is odd. This is for instance the case in the permutation in the left of (24), but not that on the right.

and:

$$\overbrace{a_i^\dagger a_j} = \delta_{ij} \times \begin{cases} 0 \text{ si } e_i > \mu \\ 1 \text{ si } e_i < \mu \end{cases} \tag{28}$$

(the second lines ($e_i < \mu$) of these relations are useful only for fermions, since for bosons μ cannot be larger than e_i).

1-d. Statement of the theorem

The recurrence over n we have been using leads to Wick's theorem:

"The average value $\langle b_1 b_2 ... b_i .. b_l ... b_{2n} \rangle$ is the sum of all the complete systems of contractions that can be made on the string of operators $b_1 b_2 ... b_i .. b_l ... b_{2n}$. Each system is the product of binary contractions (16); for fermions, this product is multiplied by parity factors ε_α associated with each of them."

The word "complete" means in this case that in every considered system of contractions, each operator listed in the string of b_i operators is taken in one and only one contraction. The parity factor first includes the parity ε_{α_1} of the permutation that brings right after b_1 the operator it is contracted with; these two operators are then taken out of the list of the b_i. In the remaining list, we again compute the parity ε_{α_2} of the permutation needed to bring together the next two operators to be contracted, and it is multiplied by ε_{α_1}. We continue this until all the contractions have been taken into account, and obtain the product $\varepsilon_{\alpha_1} \varepsilon_{\alpha_2} ... \varepsilon_{\alpha_n}$ of all the parities involved. Among all the system of contractions, a very large number yield zero. The only non-zero ones are those for which every contraction contains a creation and an annihilation operator in the same individual quantum state. This rule significantly limits the number of contractions involved in the final result.

As seen above, the theorem yields again the results of Complement B_{XV}. For example, if (as is the case in the formula for the two-particle symmetric operators) the first two operators b are creation operators, and the last two annihilation operators, the first system of contraction in (23) yields zero, and we are left with the last two, corresponding to the two terms of equation (43) in Complement B_{XV}. The main interest of the theorem is, however, that it allows getting, almost without calculations, the average value of the product of any number of operators.

Comment:

Until now, we assumed that the operators b were creation or annihilation operators associated with the basis of individual states formed by the one-particle Hamiltonian eigenvectors. If this is not the case, and we wish to compute the average value of the product of creation $a_{v_s}^\dagger$ and annihilation a_{v_s} operators associated with any other basis, we first use formulas (A-51) and (A-52) of Chapter XV to express those operators in terms of the operators associated with the eigenbasis of the one-particle Hamiltonian, and then use Wick's theorem.

2. Applications: correlation functions for an ideal gas

As an illustration of the use of Wick's theorem, we now compute the N-order correlation functions in an ideal gas at thermal equilibrium. Thanks to Wick's theorem, they can

each be expressed as simple products of first order correlation functions. As a first step, we will derive, in a simpler way, a number of results already obtained in § 3 of Complement B_{XV}; these will then be generalized to correlation functions of a higher order.

Consider a gas of spinless particles, confined by a one-body potential inside a cubic box of edge length L; this potential is zero inside the box, and becomes infinite outside. We use periodic boundary conditions to account for this confinement (Complément C_{XIV}, § 1-c); the normalized eigenfunctions $u_{\mathbf{k}}(\mathbf{r})$ of the kinetic energy are then written:

$$u_{\mathbf{k}}(\mathbf{r}) = \frac{1}{L^{3/2}} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{29}$$

where the possible wave vectors $\mathbf{k}$ are those whose three components are integer multiples of $2\pi/L$.

2-a. First order correlation function

Relation (B-21) of Chapter XVI defines the first order correlation function G_1, which depends on the two positions $\mathbf{r}_1$ and $\mathbf{r}'_1$:

$$G_1(\mathbf{r}_1, \mathbf{r}'_1) = \left\langle \Psi^\dagger(\mathbf{r}_1)\Psi(\mathbf{r}'_1) \right\rangle \tag{30}$$

Using relations (A-3) and (A-6) of Chapter XVI, the field operator can be expressed as a function of the annihilation operators $a_{\mathbf{k}}$ in the state (29), and its adjoint, as a function of the creation operators $a^\dagger_{\mathbf{k}}$ in that same state. Taking into account (29), this leads to:

$$G_1(\mathbf{r}_1, \mathbf{r}'_1) = \frac{1}{L^3} \sum_{\mathbf{k}} \sum_{\mathbf{k}'} e^{i\left(\mathbf{k}'\cdot\mathbf{r}'_1 - \mathbf{k}\cdot\mathbf{r}_1\right)} \left\langle a^\dagger_{\mathbf{k}} a_{\mathbf{k}'} \right\rangle \tag{31}$$

At thermal equilibrium, all the average values $\langle A \rangle$ of operators A are taken in the state described by the density operator ρ_{eq} written in (1):

$$\langle A \rangle = \mathrm{Tr}\left\{\rho_{eq} A\right\} \tag{32}$$

We can then use Wick's theorem in a particularly simple case, since in (31) the only contraction that comes into play is the one containing $\left[a^\dagger_{\mathbf{k}}, a_{\mathbf{k}'}\right]_{-\eta}$. Relation (18) thus applies and we get:

$$G_1(\mathbf{r}_1, \mathbf{r}'_1) = \frac{1}{L^3} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\left(\mathbf{r}'_1 - \mathbf{r}_1\right)} \; f_\beta\left(e_{\mathbf{k}} - \mu\right) \tag{33}$$

The correlation function $G_1(\mathbf{r}, \mathbf{r}')$ is therefore directly (to within a constant factor) the Fourier transform of the distribution function $f_\beta(e_{\mathbf{k}} - \mu)$ itself.

The definition of G_1 can be generalized, using the expressions of the field operators in the Heisenberg picture; this leads to a correlation function depending on space and time:

$$G_1(\mathbf{r}_1, t; \mathbf{r}'_1, t') = \left\langle \Psi^\dagger_H(\mathbf{r}_1, t)\Psi_H(\mathbf{r}'_1, t') \right\rangle \tag{34}$$

For free particles (ideal gas), we have (§ 1-c of Complement B_{XV}):

$$\Psi_H(\mathbf{r},t) = \frac{1}{L^{3/2}} \sum_{\mathbf{k}} e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} a_{\mathbf{k}} \tag{35}$$

where ω_k is the (angular) Bohr frequency associated with the energy of a particle of mass m, with wave vector $\mathbf{k}$:

$$\omega_k = \frac{\hbar k^2}{2m} \tag{36}$$

For an ideal gas, we simply multiply each exponential $e^{i\mathbf{k}\cdot\mathbf{r}}$ by $\mathrm{e}^{-i\omega_k t}$ to go from the Schrödinger to the Heisenberg representation. Expression (33) is then generalized as:

$$G_1(\mathbf{r}_1,t;\mathbf{r}'_1,t') = \frac{1}{L^3} \sum_{\mathbf{k}} e^{i\left[\mathbf{k}\cdot\left(\mathbf{r}'_1-\mathbf{r}_1\right)-\omega_k\left(t'-t\right)\right]}\ f_\beta\left(e_{\mathbf{k}}-\mu\right) \tag{37}$$

Note that this correlation function only depends on the differences in positions (space homogeneity) and times (time translation invariance).

2-b. Second order correlation functions

α. Application of Wick's theorem

The second order correlation function is defined as:

$$G_2(\mathbf{r}_1,\mathbf{r}'_1;\mathbf{r}_2,\mathbf{r}'_2) = \left\langle \Psi^\dagger(\mathbf{r}_1)\Psi^\dagger(\mathbf{r}_2)\Psi(\mathbf{r}'_2)\Psi(\mathbf{r}'_1)\right\rangle \tag{38}$$

Here again, the average value is computed with the density operator at thermal equilibrium. The same calculation as in § 2-a leads to:

$$G_2(\mathbf{r}_1,\mathbf{r}'_1;\mathbf{r}_2,\mathbf{r}'_2) = \frac{1}{L^6} \sum_{\mathbf{k}}\sum_{\mathbf{k}'}\sum_{\mathbf{k}''}\sum_{\mathbf{k}'''} e^{i\left(\mathbf{k}'\cdot\mathbf{r}'_1+\mathbf{k}'''\cdot\mathbf{r}'_2-\mathbf{k}\cdot\mathbf{r}_1-\mathbf{k}''\cdot\mathbf{r}_2\right)} \left\langle a^\dagger_{\mathbf{k}} a^\dagger_{\mathbf{k}''} a_{\mathbf{k}'''} a_{\mathbf{k}'}\right\rangle \tag{39}$$

As we already saw in (23), using Wick's theorem yields two contraction systems, one where $a^\dagger_{\mathbf{k}}$ is contracted with $a_{\mathbf{k}'}$ (and hence $a^\dagger_{\mathbf{k}''}$ with $a_{\mathbf{k}'''}$), and another one where $a^\dagger_{\mathbf{k}}$ is contracted with $a_{\mathbf{k}'''}$ (and hence $a^\dagger_{\mathbf{k}''}$ with $a_{\mathbf{k}'}$):

$$\overbrace{a^\dagger_{\mathbf{k}}\underbrace{a^\dagger_{\mathbf{k}''}a_{\mathbf{k}'''}}a_{\mathbf{k}'}} \qquad \text{et} \qquad \overbrace{a^\dagger_{\mathbf{k}}\underbrace{a^\dagger_{\mathbf{k}''}a_{\mathbf{k}'''}a_{\mathbf{k}'}}}$$

The second contraction involves an odd permutation, and introduces a factor η. We therefore obtain (after changing the dummy index $\mathbf{k}''$ into $\mathbf{k}'$):

$$\begin{aligned} &G_2(\mathbf{r}_1,\mathbf{r}'_1;\mathbf{r}_2,\mathbf{r}'_2) \\ &\quad= \frac{1}{L^6}\sum_{\mathbf{k}}\sum_{\mathbf{k}'}\left[e^{i\left[\mathbf{k}\cdot\left(\mathbf{r}'_1-\mathbf{r}_1\right)+\mathbf{k}'\cdot\left(\mathbf{r}'_2-\mathbf{r}_2\right)\right]} + \eta e^{i\left[\mathbf{k}\cdot\left(\mathbf{r}'_2-\mathbf{r}_1\right)+\mathbf{k}'\cdot\left(\mathbf{r}'_1-\mathbf{r}_2\right)\right]}\right] \\ &\qquad\qquad \times f_\beta\left(e_{\mathbf{k}}-\mu\right) f_\beta\left(e_{\mathbf{k}'}-\mu\right) \end{aligned} \tag{40}$$

that is, taking (33) into account:

$$G_2(\mathbf{r}_1,\mathbf{r}'_1;\mathbf{r}_2,\mathbf{r}'_2) = G_1(\mathbf{r}_1,\mathbf{r}'_1)\times G_1(\mathbf{r}_2,\mathbf{r}'_2) + \eta\ G_1(\mathbf{r}_1,\mathbf{r}'_2)\times G_1(\mathbf{r}_2,\mathbf{r}'_1) \tag{41}$$

This means that the second order correlation function is simply expressed as the sum of two products of first order correlation functions. The first is the direct term, and corresponds to totally uncorrelated particles. The second is the exchange term, a consequence of the quantum indistinguishability of the particles; it has an opposite sign for fermions and bosons. As in Complement A_{XVI}, we will show that this term introduces correlations between the particles.

β. *Double density*

Of particular interest is the "diagonal" case where $\mathbf{r}_1 = \mathbf{r}'_1$ and $\mathbf{r}_2 = \mathbf{r}'_2$, as the function $G_2(\mathbf{r}_1, \mathbf{r}'_1; \mathbf{r}_2, \mathbf{r}'_2)$ becomes very simple to interpret: it is the "double density" $G_2(\mathbf{r}_1, \mathbf{r}_2)$ characterizing the probability of finding a particle at point $\mathbf{r}_1$ and another one at point $\mathbf{r}_2$. The above relation takes on the simplified form:

$$G_2(\mathbf{r}_1, \mathbf{r}_2) = G_1(\mathbf{r}_1, \mathbf{r}_1) \times G_1(\mathbf{r}_2, \mathbf{r}_2) + \eta\ G_1(\mathbf{r}_1, \mathbf{r}_2) \times G_1(\mathbf{r}_2, \mathbf{r}_1) \tag{42}$$

If, in addition, $\mathbf{r}_1 = \mathbf{r}_2$, this function indicates the probability of finding two particles at the same point. We then get:

$$\begin{cases} \text{for fermions} & G_2(\mathbf{r}; \mathbf{r}) = 0 \\ \text{for bosons} & G_2(\mathbf{r}; \mathbf{r}) = 2\ \left[G_1(\mathbf{r}, \mathbf{r})\right]^2 \end{cases} \tag{43}$$

For fermions, we see as expected that one can never find two of them at the same point in space, a consequence of Pauli exclusion principle. For bosons, we find that the double density is twice the square of the one-body density. Now, if both particles were uncorrelated, this double density should simply be equal to the square, without the factor two. This factor two thus indicates an increase in the probability of finding two bosons at the same point in space; it expresses the bunching tendency of identical bosons, a tendency that comes from a pure quantum statistical effect since we assumed the particle's interactions to be zero. These results were already discussed in Complement A_{XVI} – see in particular Figure 3.

The Hartree-Fock method (mean field approximation), presented in Complements E_{XV} and F_{XV}, uses a variational ket (or a density operator) such that the binary correlation function $G_2(\mathbf{r}_1, \mathbf{r}_2)$ is given by the sum of products of functions G_1, written in (42), even in the presence of interactions. Moreover, another way of introducing the Hartree-Fock approximation is to assume directly that the binary correlation function keeps this form even in the presence of interactions, which then allows a simple calculation of the interaction energy. Even though this method has numerous applications, and may be quite precise in certain cases, it does rely on an approximation: when the particles interact with each other, there is no general reason for G_2 to remain linked to G_1 by this relation, obtained with the assumption that the gas was ideal.

γ. *Time-dependent correlation function*

As we did for the first order correlation function, we can include time dependence in the field operators, and define:

$$G_2(\mathbf{r}_1, t_1; \mathbf{r}'_1, t'_1; \mathbf{r}_2, t_2; \mathbf{r}'_2, t'_2) = \left\langle \Psi_H^\dagger(\mathbf{r}_1, t_1) \Psi_H^\dagger(\mathbf{r}_2, t_2) \Psi_H(\mathbf{r}'_2, t'_2) \Psi_H(\mathbf{r}'_1, t'_1) \right\rangle \tag{44}$$

To include the time dependence, we simply add, as above, in each spatial exponential with wave vector $\mathbf{k}$, a time exponential with the corresponding angular frequency $-\omega_k$, which leads to:

$$\begin{aligned} G_2(\mathbf{r}_1,t_1;\mathbf{r}'_1,t'_1;\mathbf{r}_2,t_2;\mathbf{r}'_2,t'_2) = \\ \frac{1}{L^6}\sum_{\mathbf{k}}\sum_{\mathbf{k}'}\Big\{ \mathrm{e}^{\mathrm{i}[\mathbf{k}\cdot(\mathbf{r}'_1-\mathbf{r}_1)-\omega_k(t'_1-t_1)+\mathbf{k}'\cdot(\mathbf{r}'_2-\mathbf{r}_2)-\omega_{k'}(t'_2-t_2)]} \\ + \eta\, \mathrm{e}^{\mathrm{i}[\mathbf{k}\cdot(\mathbf{r}'_2-\mathbf{r}_1)-\omega_k(t'_2-t_1)+\mathbf{k}'\cdot(\mathbf{r}'_1-\mathbf{r}_2)-\omega_{k'}(t'_1-t_2)]}\Big\} f_\beta(e_{\mathbf{k}}-\mu)\, f_\beta(e_{\mathbf{k}'}-\mu) \\ = G_1(\mathbf{r}_1,t_1;\mathbf{r}'_1,t'_1)\, G_1(\mathbf{r}_2,t_2;\mathbf{r}'_2,t'_2) + \eta\, G_1(\mathbf{r}_1,t_1;\mathbf{r}'_2,t'_2)\, G_1(\mathbf{r}_2,t_2;\mathbf{r}'_1,t'_1) \end{aligned} \tag{45}$$

Hence, when time dependence is included, we get a factored relation similar to (41). As before, because of the space homogeneity and the time translation invariance, only the differences in the space and time variables appear in the correlation function expression.

2-c. Higher order correlation functions

In a more general way[2], the N-order correlation function G_N is defined by:

$$G_N(\mathbf{r}_1,\mathbf{r}'_1;\mathbf{r}_2,\mathbf{r}'_2;...;\mathbf{r}_N,\mathbf{r}'_N) = \left\langle \Psi^\dagger(\mathbf{r}_1)\Psi^\dagger(\mathbf{r}_2)...\Psi^\dagger(\mathbf{r}_N)\Psi(\mathbf{r}'_N)...\Psi(\mathbf{r}'_2)\Psi(\mathbf{r}'_1)\right\rangle \tag{46}$$

These functions give information on the correlated behavior of groups of N particles in an ideal gas at equilibrium. Using Wick's theorem, each of them can be expressed in terms of the first order correlation function $G_1(\mathbf{r}_1,\mathbf{r}'_1)$. As an example, let us study the correlation function of order three:

$$\begin{aligned} G_3(\mathbf{r}_1,\mathbf{r}'_1;\mathbf{r}_2,\mathbf{r}'_2;\mathbf{r}_3,\mathbf{r}'_3) &= \left\langle \Psi^\dagger(\mathbf{r}_1)\Psi^\dagger(\mathbf{r}_2)\Psi^\dagger(\mathbf{r}_3)\Psi(\mathbf{r}'_3)\Psi(\mathbf{r}'_2)\Psi(\mathbf{r}'_1)\right\rangle \\ &= \frac{1}{L^9}\sum_{\mathbf{k}}\sum_{\mathbf{k}'}\sum_{\mathbf{k}''}\sum_{\mathbf{k}'''}\sum_{\mathbf{k}''''}\sum_{\mathbf{k}'''''} \\ &\qquad e^{i\left(\mathbf{k}'\cdot\mathbf{r}'_1+\mathbf{k}'''\cdot\mathbf{r}'_2+\mathbf{k}'''''\cdot\mathbf{r}'_3-\mathbf{k}\cdot\mathbf{r}_1-\mathbf{k}''\cdot\mathbf{r}_2-\mathbf{k}''''\cdot\mathbf{r}_3\right)} \\ &\qquad \times \left\langle a^\dagger_{\mathbf{k}} a^\dagger_{\mathbf{k}''} a^\dagger_{\mathbf{k}''''} a_{\mathbf{k}'''''} a_{\mathbf{k}'''} a_{\mathbf{k}'}\right\rangle \end{aligned} \tag{47}$$

Six contraction systems must be considered to compute the average values. In the first system, $\mathbf{k}$ and $\mathbf{k}'$ are associated, $\mathbf{k}''$ with $\mathbf{k}'''$, and finally $\mathbf{k}''''$ with $\mathbf{k}'''''$. One can then permute the three vectors $\mathbf{k}'$, $\mathbf{k}'''$ and $\mathbf{k}'''''$ in 5 different ways, with odd or even permutations. In each of the terms thus obtained, the sixfold summation on the wave vectors is reduced to a triple sum, which yields a product of functions G_1. This leads to:

$$\begin{aligned} G_3(\mathbf{r}_1,\mathbf{r}'_1;\mathbf{r}_2,\mathbf{r}'_2;\mathbf{r}_3,\mathbf{r}'_3) &= G_1(\mathbf{r}_1,\mathbf{r}'_1)\, G_1(\mathbf{r}_2,\mathbf{r}'_2)\, G_1(\mathbf{r}_3,\mathbf{r}'_3) \\ &+G_1(\mathbf{r}_1,\mathbf{r}'_2)\, G_1(\mathbf{r}_2,\mathbf{r}'_3)\, G_1(\mathbf{r}_3,\mathbf{r}'_1) + G_1(\mathbf{r}_1,\mathbf{r}'_3)\, G_1(\mathbf{r}_2,\mathbf{r}'_1)\, G_1(\mathbf{r}_3,\mathbf{r}'_2) \\ &+\eta\, G_1(\mathbf{r}_1,\mathbf{r}'_1)\, G_1(\mathbf{r}_2,\mathbf{r}'_3)\, G_1(\mathbf{r}_3,\mathbf{r}'_2) + \eta\, G_1(\mathbf{r}_1,\mathbf{r}'_3)\, G_1(\mathbf{r}_2,\mathbf{r}'_2)\, G_1(\mathbf{r}_3,\mathbf{r}'_1) \\ &+\eta\, G_1(\mathbf{r}_1,\mathbf{r}'_2)\, G_1(\mathbf{r}_2,\mathbf{r}'_1)\, G_1(\mathbf{r}_3,\mathbf{r}'_3) \end{aligned} \tag{48}$$

The computation can be generalized in the same way to correlation functions of any order; in an ideal gas, they are not independent since they are all simple products

[2]We only consider here the so-called "normal"correlation functions, those where the $\Psi^\dagger$ come before the Ψ. In Complement B_{XVI} we introduce more general correlation functions.

of first order correlation functions. In other words, the function G_1 contains all the information necessary for computing correlations of any order.

Finally, we can compute the triple density $G_3(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ by setting $\mathbf{r}_1 = \mathbf{r}'_1$, $\mathbf{r}_2 = \mathbf{r}'_2$ and $\mathbf{r}_3 = \mathbf{r}'_3$ in (48). The particular case $\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}_3$ where all the positions are identical is interesting. For fermions, the triple density is zero, for the same reason as with the double density: Pauli principle does not allow several fermions to occupy the same point in space. For bosons, we find:

$$G_3(\mathbf{r}, \mathbf{r}; \mathbf{r}) = 6 \left[G_1 (\mathbf{r}, \mathbf{r})\right]^3 \tag{49}$$

For three identical bosons, the bunching tendency under the effect of their quantum statistics is even higher than for two bosons, introducing a factor 6 instead of 2.

Comment:

The results we have obtained are valid when the system's density operator is that of an ideal gas at thermal equilibrium as in relation (1), but they could be quite different if the system is in another state. If, for example, we assume (as in Complement $\mathrm{A_{XVI}}$) that the system is described by a Fock state, the relations between correlation functions can be totally different. The simplest case is that of an ideal gas of bosons in its ground state, where all the bosons occupy the same individual state; relations (24) and (25) of that complement indicate that:

$$G_2(\mathbf{r}_1, \mathbf{r}_2) = \frac{N-1}{N} G_1(\mathbf{r}_1, \mathbf{r}_1) \times G_1(\mathbf{r}_2, \mathbf{r}_2) \simeq G_1(\mathbf{r}_1, \mathbf{r}_1) \times G_1(\mathbf{r}_2, \mathbf{r}_2) \tag{50}$$

G_2 is thus simply the product of two functions G_1, without the exchange term of (42); consequently, the factor 2 in the second line of (43) no longer exists. In a similar way, one can show that the factor 6 of relation (49) is no longer present. In a general way, for an ensemble of bosons all in the same individual state, the bunching effects related to the indistinguishability of the particles are not present.

Following this line of thought, note that it is not possible to get the projector onto a Fock state (other than the vacuum) such as the one discussed above, by using the density operator (1) at thermal equilibrium, and taking its limit as the temperature T goes to zero, i.e. as $\beta \to \infty$. This is because this density operator associates with each individual state an occupation number distribution that is always a decreasing exponential, and never a narrow curve centered around a high value of the particle number. Consequently, there exist large fluctuations of the particle number in each mode, and hence the presence of the factors 2 in (43) and 6 in (49), whatever the value of β.

To conclude, let us mention that Wick's theorem can take on diverse forms, in particular at zero or non-zero temperadure (see for instance Chapter 4 of Reference [5]). As we saw, thanks to this theorem, and when dealing with independent particles, the computation of correlation functions of any order, time-dependent or not, can be reduced to computing the product of first order correlation functions. It is obviously a great simplification. This property is reminiscent of random Gaussian variables in classical statistics: for such variables, all moments of any order can be expressed in terms of products of the lowest order moment. These properties are characteristic of an ideal gas: in a system where particles interact, the correlation functions of successive orders remain independent in general. Nevertheless, the use of Wick's theorem is not limited to ideal gases; its range of application is much more general, and it is very useful in perturbation calculations where power series of the interaction potential are derived [5].

Chapter XVII

Paired states of identical particles

A Creation and annihilation operators of a pair of particles . 1813
A-1 Spinless particles, or particles in the same spin state 1813
A-2 Particles in different spin states 1816
B Building paired states 1818
B-1 Well determined particle number 1818
B-2 Undetermined particle number 1820
B-3 Pairs of particles and pairs of individual states 1822
C Properties of the kets characterizing the paired states . . . 1822
C-1 Normalization 1822
C-2 Average value and root mean square deviation of particle number 1825
C-3 "Anomalous" average values 1828
D Correlations between particles, pair wave function 1830
D-1 Particles in the same spin state 1831
D-2 Fermions in a singlet state 1834
E Paired states as a quasi-particle vacuum; Bogolubov-Valatin transformations 1836
E-1 Transformation of the creation and annihilation operators . . 1836
E-2 Effect on the kets $|\overline{\varphi}_{\mathbf{k}}\rangle$ 1838
E-3 Basis of excited states, quasi-particles 1840

Introduction

Fock states were introduced in Chapter XV by the action on the vacuum of a product of individual state creation operators. A certain number of their properties were studied in § C-5-b-β of Chapter XV and in Complement $\mathrm{A_{XVI}}$ (exchange hole for fermions, bunching

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

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

effect for bosons). We also used Fock states in Complements C_{XV}, D_{XV} and F_{XV} as variational kets to account, approximately, for the interactions between the particles. This led us, both for fermions and bosons, to a mean field theory where each particle can be seen as propagating in the mean field created by all the others.

We now introduce a larger class of variational states to improve the accuracy of these results, allowing us to study many more properties of physical systems of identical particles. It concerns the "paired states" obtained by the action on the vacuum of a product of creation operators, no longer of individual particles but rather of pairs of particles (if these particles form a molecule, we are dealing with molecule creation operators). As we shall see in the course of this chapter, these paired states are more general than the Fock states, since they can be reduced to Fock states for certain values of the parameters characteristic of the pair[1]. What is sought is an improvement of the variational method allowing us to ameliorate our treatment of the interactions, compared to that based on Fock states.

The additional flexibility introduced by the paired states plays an essential role for the following simple reason: changing the properties of the pair wave function $\chi(\mathbf{r}_1 - \mathbf{r}_2)$ used to build them, we modify the binary correlation function of the N-particle system. We therefore take advantage of the power of the mean field method, while retaining the possibility of taking into account any binary correlations. Whereas using variational Fock states allows taking into account only statistical correlations (due to particle indistinguishability), the paired states enable us to add dynamic correlations (due to interactions). These latter correlations are essential: when dealing with binary interactions (as is the case with a standard Hamiltonian), these correlations actually determine the average value of the potential energy (Chapter XV, § C-5-b). Three-body, four-body, etc.. correlations are indeed present in the system, playing their role; they are not, however, directly involved in the energy. This explains why the optimization with paired states of only the binary correlations can lead to fairly good results in the study of N-body systems. These possibilities have a wide range of applications for both fermions and bosons, which will be discussed in the complements.

This chapter is centered on the study of the general properties of paired states, and introduces the tools for handling such states. We study, in parallel, fermions and bosons to highlight the numerous analogies between results obtained for both cases. We first introduce (§ A) the creation and annihilation operators for pairs of particles. We then build (§ B) the paired states and discuss some of their properties; this permits introducing (§ C) the concept of "normal average values" (average values of operators conserving the particle number) or "anomalous average values" (average values of operators changing the particle number). We then show in § D how the paired states allow us to actually vary the spatial correlation functions of a system of identical particles. This will lead us to introduce a function playing an important role in what follows (in particular in the complements of this chapter), the pair wave function ϕ_{pair}, which is related to the anomalous average values. We then study in § E another interesting property of the paired states: they can be related to the concept of "quasi-particle" thanks to the introduction of new creation and annihilation operators resulting from a linear transformation of the initial operators (Bogolubov transformation). As the paired states are eigenkets of the new annihilation operators with a zero eigenvalue, they behave as

[1]For example, we shall clarify at the end of § C-1-a why the Hartree-Fock method can be viewed as a particular case of the pairing method.

a "quasi-particle vacuum". Furthermore, the creation operators can associate with each paired state an entire basis of other orthogonal states, which are interpreted as states occupied by quasi-particles.

This study of the necessary tools for handling the paired states will be continued in the first two complements, $\mathrm{A_{XVII}}$ and $\mathrm{B_{XVII}}$. Complement$\mathrm{A_{XVII}}$ discusses a complementary aspect of pairing, the introduction of the pair field operators. These operators have a non-zero average value in paired states, and highlight the cooperative effects existing in those states. This can lead to the spontaneous appearance in the system of an order parameter, described by the same pair wave function ϕ_{pair} as the one appearing in the computations of correlation functions in a paired state. In addition, Complement $\mathrm{A_{XVII}}$ will show that the commutation properties of these operators are reminiscent of those of a boson field: in a certain sense, a composite object built from two identical particles (whether they are bosons or fermions) behaves as a boson. It is, however, only an approximation, as can be inferred from the corrective terms appearing in the computation of the commutators, which can sometimes play an important role. Complement $\mathrm{B_{XVII}}$ discusses the computation of the energy average value in a paired state, whose expression is the basis of the following complements; it gives an example of how to deal with normal and anomalous average values in these computations.

The last three complements apply these results to the variational study of interacting boson or fermion systems. For fermions, the paired states play an essential role in the BCS (Bardeen-Cooper-Schrieffer, theory of supraconductivity) theory of supraconductivity (Complement $\mathrm{C_{XVII}}$), and explain the appearance of a pair field as a collective effect; paired states also come into play noticeably in nuclear physics, and in the study of ultra-cold fermionic atomic gases. For repulsive bosons (Complement $\mathrm{E_{XVII}}$), paired states can be quite useful for studying the ground state properties, and to obtain, for example, the Bogolubov linear spectrum. In that case, the paired state is associated with another state (a coherent state, for example), whose role is to describe the condensate as an accumulation of a notable fraction of particles in a single individual quantum state.

A. Creation and annihilation operators of a pair of particles

Let us introduce the creation or annihilation operators, no longer of a single particle, but of two identical particles in a bound state. We first assume the particles have no spin (or are both in the same spin state, so that no spin variable is needed).

A-1. Spinless particles, or particles in the same spin state

Consider two identical particles (bosons or fermions in the same spin state), with positions $\mathbf{r}_1$ and $\mathbf{r}_2$; the system is contained in a cubic box of edge length L and volume $\mathcal{V} = L^3$. These two particles occupy a bound state, characterized by a normalized wave function $\chi(\mathbf{r}_1 - \mathbf{r}_2)$, forming a kind of binary "molecule". The state of the system is defined by this wave function (as far as its internal orbital variables are concerned), by spin variables identical for both particles (since those spin variables are of no importance here, they need not be written explicitly in what follows), and finally by its external orbital variables (center of mass). The normalized wave function of a "molecule" having

a total momentum $\hbar\mathbf{K}$ is then:

$$\begin{aligned}\phi_{\mathbf{K}}(\mathbf{r}_1,\mathbf{r}_2) &= (L)^{-3/2}\, e^{i\mathbf{K}\cdot(\mathbf{r}_1+\mathbf{r}_2)/2}\chi(\mathbf{r}_1-\mathbf{r}_2) \\ &= (L)^{-3}\sum_{\mathbf{k}} g_{\mathbf{k}}\; e^{i\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right)\cdot\mathbf{r}_1} e^{i\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right)\cdot\mathbf{r}_2}\end{aligned} \tag{A-1}$$

where $g_{\mathbf{k}}$ is the Fourier transform of χ:

$$\begin{aligned} g_{\mathbf{k}} &= \frac{1}{L^{3/2}}\int_{L^3} d^3r\; e^{-i\mathbf{k}\cdot\mathbf{r}}\chi(\mathbf{r}) \\ \chi(\mathbf{r}) &= \frac{1}{L^{3/2}}\sum_{\mathbf{k}}\; g_{\mathbf{k}}\; e^{i\mathbf{k}\cdot\mathbf{r}}\end{aligned} \tag{A-2}$$

We assume that the individual wave functions of the particles obey the periodic boundary conditions (Complement C_{XIV}, § 1-c); in (A-1), each component of the wave vector of particle 1 or 2 can therefore take only the values $2\pi n_x/L$, $2\pi n_y/L$ and $2\pi n_z/L$, where n_x, n_y and n_z are any integer number (positive, negative, or zero). The normalization of the functions χ and g is written:

$$\int_{L^3} \mathrm{d}^3r\;\; |\chi(\mathbf{r})|^2 = \sum_{\mathbf{k}}\;\; |g_{\mathbf{k}}|^2 = 1 \tag{A-3}$$

Moreover, for identical particles, the symmetrization (or antisymmetrization) requires the function $\chi(\mathbf{r})$ and its Fourier transform $g(\mathbf{k})$ to have the parity η:

$$g_{-\mathbf{k}} = \eta\; g_{\mathbf{k}} \tag{A-4}$$

($\eta = +1$ for bosons, $\eta = -1$ for fermions).

In terms of kets, relation (A-1) becomes:

$$\begin{aligned}|\phi_{\mathbf{K}}(1,2)\rangle &= (L)^{-3}\sum_{\mathbf{k}}\int \mathrm{d}^3r_1\int \mathrm{d}^3r_2\; g_{\mathbf{k}}\; e^{i\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right)\cdot\mathbf{r}_1} e^{i\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right)\cdot\mathbf{r}_2}\,|1:\mathbf{r}_1;2:\mathbf{r}_2\rangle \\ &= \sum_{\mathbf{k}}\; g_{\mathbf{k}}\;\left|1:\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right);2:\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right)\right\rangle\end{aligned} \tag{A-5}$$

which, taking (A-4) into account, and changing the sign of the sum variable $\mathbf{k}$, can also be written as:

$$\begin{aligned}|\phi_{\mathbf{K}}(1,2)\rangle = \frac{1}{2}\sum_{\mathbf{k}}\; g_{\mathbf{k}}\;\Bigg[&\left|1:\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right);2:\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right)\right\rangle \\ &+\eta\left|1:\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right);2:\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right)\right\rangle\Bigg]\end{aligned} \tag{A-6}$$

The expression between brackets in the summation is simply the (anti)symmetrized ket of two particles, the first one of momentum $\hbar(\mathbf{k}+\mathbf{K}/2)$, and the other one of momentum $\hbar(-\mathbf{k}+\mathbf{K}/2)$. Two cases must be distinguished:

(i) If $\mathbf{k} \neq 0$, to normalize the ket between brackets, we divide it by $\sqrt{2}$; we then get a Fock state where two individual states with different momenta are occupied (see

the general definition of the Fock states in Chapter XV). The ket between brackets is thus equal to:

$$\sqrt{2}\ a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}}\ |0\rangle \tag{A-7}$$

(ii) If $\mathbf{k} = 0$ and in the case of bosons, the ket between brackets is equal to twice the Fock state where a single individual level is occupied by two particles; this ket is equal to:

$$\sqrt{2}\left(a^\dagger_{\frac{\mathbf{K}}{2}}\right)^2 |0\rangle \tag{A-8}$$

For fermions, the ket between brackets must be zero, which is indeed the case of the ket in (A-8). To sum up, whether we are dealing with fermions or bosons, and whether $\mathbf{k}$ is zero or not, the ket between brackets can always be expressed as (A-7). This leads to:

$$|\phi_{\mathbf{K}}\rangle = \frac{1}{\sqrt{2}} \sum_{\mathbf{k}}\ g_{\mathbf{k}}\ a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}}\ |0\rangle \tag{A-9}$$

If the particles are all in the same spin state, remember that in this expression the spin index is implicit: each creation operator is associated with an individual state whose momentum is specified by the operator index, and whose spin state is the common spin state of all the particles.

The creation operator $A^\dagger_{\mathbf{K}}$ of a "molecule" having a total momentum $\hbar\mathbf{K}$ can therefore be written as:

$$A^\dagger_{\mathbf{K}} = \frac{1}{\sqrt{2}} \sum_{\mathbf{k}}\ g_{\mathbf{k}}\ a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}} \tag{A-10}$$

Its action is to create two particles of momenta $\hbar\left[(\mathbf{K}/2) \pm \mathbf{k}\right]$, with amplitudes given by the function $g_{\mathbf{k}}$. As this function has the parity η, we note that:

$$g_{-\mathbf{k}}\ a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} = \eta g_{\mathbf{k}}\ a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} = g_{\mathbf{k}}\ a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}} \tag{A-11}$$

Accordingly, the contributions of opposite values of $\mathbf{k}$ double each other in (A-10). Such a redundancy will cause a problem in § B-2, when we write a tensor product. It is thus preferable to eliminate it right now and this is why we restrict the summation over $\mathbf{k}$ to half the wave vector space. Calling D this half space, we shall write $A^\dagger_{\mathbf{K}}$ in the form:

$$A^\dagger_{\mathbf{K}} = \sqrt{2} \sum_{\mathbf{k}\in D}\ g_{\mathbf{k}}\ a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}} \tag{A-12}$$

For a "molecule" having a zero total momentum, this relation becomes:

$$A^\dagger_{\mathbf{K}=\mathbf{0}} = \sqrt{2} \sum_{\mathbf{k}\in D}\ g_{\mathbf{k}}\ a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} \tag{A-13}$$

As for the annihilation operator of a molecule with total momentum $\hbar\mathbf{K}$, it is simply the adjoint of (A-12):

$$A_{\mathbf{K}} = \sqrt{2} \sum_{\mathbf{k}\in D}\ g^*_{\mathbf{k}}\ a_{\frac{\mathbf{K}}{2}-\mathbf{k}} a_{\frac{\mathbf{K}}{2}+\mathbf{k}} \tag{A-14}$$

We have reasoned in terms of "molecules" being created or annihilated, but the wave function $\chi(\mathbf{r})$ and its Fourier transform $g_{\mathbf{k}}$ are not related to any particular bound state, and do not imply the existence of any attraction potential between the two components of this "molecule". Actually, in what follows, the $g_{\mathbf{k}}$ will play the role of freely adjustable parameters, for example when using a variational method. To illustrate this generality, we shall, from now on, talk about "pairs".

Comment:

If we choose for $g_{\mathbf{k}}$ as in a Kronecker delta, with the symmetrization required by (A-4):

$$g_{\mathbf{k}} = \frac{1}{\sqrt{2}} \left[\delta_{\mathbf{k},\mathbf{k}_0} + \eta \delta_{\mathbf{k},-\mathbf{k}_0}\right] \tag{A-15}$$

we get, according to (A-12):

$$A^\dagger_{\mathbf{K}} = \frac{1}{2}\left[a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}_0} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}_0} + \eta\, a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}_0} a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}_0}\right] = a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}_0} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}_0} \tag{A-16}$$

In the right-hand side of this relation, the momenta appearing as indices of the creation operators can take any given values, obtained by varying $\mathbf{K}$ and $\mathbf{k}_0$. It is therefore possible, by a suitable choice of the pair's parameters, to create two particles in individual states having any given momenta, and thereby obtain a Fock state. Successive applications of P operators $A^\dagger_{\mathbf{K}}$ (having, in general, different values of $\mathbf{K}$ and $\mathbf{k}_0$) can thus yield a Fock state with $2P$ particles whose momenta can take on any values.

A-2. Particles in different spin states

We assume the internal state of the pair is a tensor product of an orbital state depending on $\mathbf{r}_1 - \mathbf{r}_2$ and a spin state $|\chi_S\rangle$. Equation (A-1) must then be replaced by:

$$\begin{aligned}
\phi^{\nu_1,\nu_2}_{\mathbf{K}}(\mathbf{r}_1,\mathbf{r}_2) &= \langle 1:\mathbf{r}_1,\nu_1;\ 2:\mathbf{r}_2,\nu_2\ |\Phi_{\mathbf{K}}\rangle \\
&= (L)^{-3/2}\, e^{i\mathbf{K}\cdot(\mathbf{r}_1+\mathbf{r}_2)/2} \chi(\mathbf{r}_1-\mathbf{r}_2)\, \langle \nu_1,\nu_2\ |\chi_S\rangle \\
&= (L)^{-3} \sum_{\mathbf{k}} g_{\mathbf{k}}\, e^{i\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right)\cdot\mathbf{r}_1} e^{i\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right)\cdot\mathbf{r}_2}\, \langle \nu_1,\nu_2\ |\chi_S\rangle
\end{aligned} \tag{A-17}$$

This means that relation (A-1) is to be multiplied by $\langle \nu_1,\nu_2\ |\chi_S\rangle$; relation (A-5) is now written:

$$\begin{aligned}
|\phi_{\mathbf{K}}(1,2)\rangle &= (L)^{-3} \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2 \sum_{\mathbf{k}} g_{\mathbf{k}}\, e^{i\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right)\cdot\mathbf{r}_1} e^{i\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right)\cdot\mathbf{r}_2} \\
&\qquad \times \sum_{\nu_1,\nu_2} \langle \nu_1,\nu_2\ |\chi_S\rangle\ \ |1:\mathbf{r}_1,\nu_1;\ 2:\mathbf{r}_2,\nu_2\rangle \\
&= \sum_{\mathbf{k}} g_{\mathbf{k}} \sum_{\nu_1,\nu_2} \langle \nu_1,\nu_2\ |\chi_S\rangle \left|1:\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right),\nu_1;\ 2:\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right),\nu_2\right\rangle
\end{aligned} \tag{A-18}$$

The function $\chi(\mathbf{r}_1 - \mathbf{r}_2)$ is supposed to have an orbital parity equal to η_o, and the spin ket $|\chi_S\rangle$, a parity with respect to the exchange of spins equal to η_s, with, obviously:

$$\eta_o \ \eta_s \ = \eta \tag{A-19}$$

Hence:

$$|\phi_{\mathbf{K}}(1,2)\rangle = \frac{1}{2}\sum_{\mathbf{k}} \ g_{\mathbf{k}} \ \sum_{\nu_1,\nu_2} \langle \nu_1, \nu_2 \ |\chi_S\rangle \left\{ \left| 1 : \frac{\mathbf{K}}{2} + \mathbf{k}, \nu_1; 2 : \frac{\mathbf{K}}{2} - \mathbf{k}, \nu_2 \right\rangle \right.$$
$$\left. + \eta \left| 1 : \frac{\mathbf{K}}{2} - \mathbf{k}, \nu_2; 2 : \frac{\mathbf{K}}{2} + \mathbf{k}, \nu_1 \right\rangle \right\} \tag{A-20}$$

which shows that the creation operator of a pair is:

$$A^{\dagger}_{\mathbf{K}} = \frac{1}{\sqrt{2}}\sum_{\mathbf{k}} \ g_{\mathbf{k}} \ \sum_{\nu_1,\nu_2} \langle \nu_1, \nu_2 \ |\chi_S\rangle \ \ a^{\dagger}_{\frac{\mathbf{K}}{2}+\mathbf{k},\nu_1} a^{\dagger}_{\frac{\mathbf{K}}{2}-\mathbf{k},\nu_2} \tag{A-21}$$

As an example, for two fermions of spin 1/2 in a singlet state:

$$A^{\dagger}_{\mathbf{K}} = \frac{1}{2}\sum_{\mathbf{k}} \ g_{\mathbf{k}} \ \left[a^{\dagger}_{\frac{\mathbf{K}}{2}+\mathbf{k},\nu=+} a^{\dagger}_{\frac{\mathbf{K}}{2}-\mathbf{k},\nu=-} - a^{\dagger}_{\frac{\mathbf{K}}{2}+\mathbf{k},\nu=-} a^{\dagger}_{\frac{\mathbf{K}}{2}-\mathbf{k},\nu=+} \right] \tag{A-22}$$

Since $\eta_s = -1$, the functions $\chi(\mathbf{r})$ and $g_{\mathbf{k}}$ are even. Using this parity, we can exchange the dummy indices $\mathbf{k}$ and $-\mathbf{k}$ in the second term on the right-hand side, and change the order of the two creation operators, with a sign change (anticommutation of fermionic operators). This second term then doubles the first one, and we get:

$$A^{\dagger}_{\mathbf{K}} = \sum_{\mathbf{k}} \ g_{\mathbf{k}} \ a^{\dagger}_{\frac{\mathbf{K}}{2}+\mathbf{k},\uparrow} a^{\dagger}_{\frac{\mathbf{K}}{2}-\mathbf{k},\downarrow} \tag{A-23}$$

with the simplified notation we shall use from now on:

$$\begin{array}{lll} a_{\mathbf{k},\nu=+} & \text{noted:} & a_{\mathbf{k},\uparrow} \\ a_{\mathbf{k},\nu=-} & \text{noted:} & a_{\mathbf{k},\downarrow} \end{array} \tag{A-24}$$

(and, of course, a similar notation for the creation operators $a^{\dagger}$). Note in passing that, because of the presence of spins, no redundancy is present in the summation appearing in (A-23), and there is no need to restrict it to a half-space.

Comments:

(i) Taking for $g_{\mathbf{k}}$ a (symmetrized) delta function, as in (A-15), it is possible, as we pointed out before, to construct any Fock state with arbitrary momenta by successive application of operators $A^{\dagger}_{\mathbf{K}}$ on the vacuum; note, however, that the total occupation numbers of the two spin states must remain equal.

(ii) Choosing in (A-22) a function $g_{\mathbf{k}}$ that is even instead of odd for fermions, the operator written in (A-23) creates a fermion pair with a total spin state $S = 1$, and a $M = 0$ component. This is because replacing the minus sign by a plus sign in the middle of the bracket of relation (A-22) yields a triplet spin state; using the fact that $g_{\mathbf{k}}$ is now an odd function, the same reasoning leads to (A-23).

B. Building paired states

To avoid complex formulas, we will build the simplest possible paired states $|\Psi_P\rangle$. We shall be guided by the Gross-Pitaevskii variational method (Complement C_{XV}), where we assumed that the state of the N-particle system could be obtained from the vacuum by creating N particles in the same individual state. However, instead of applying many times the creation operator $a^\dagger_{\mathbf{k}}$ of a single particle to the particle vacuum, we shall now use the pair creation operator $A^\dagger_{\mathbf{K}}$. This difference is essential, in particular for fermions. As we know, it is impossible to create several fermions in the same individual quantum state, since the square, cube, etc. of any creation operator acting on a given individual state yields zero. We shall see, however, that the creation of P pairs of fermions, all in the same quantum state, does not lead to a zero state vector.

B-1. Well determined particle number

We define the paired state $|\Psi_P(\mathbf{K})\rangle$ as the (non-normalized) state vector where $N = 2P$ particles form P pairs, each having a total momentum $\hbar\mathbf{K}$:

$$|\Psi_P(\mathbf{K})\rangle = \left[A^\dagger_{\mathbf{K}}\right]^P |0\rangle \tag{B-1}$$

where $A^\dagger_{\mathbf{K}}$ has been defined in (A-12) or (A-21), depending on the case. To keep things simple, we assume in what follows that all the created pairs have zero total momentum; if this is not the case, we can change the reference frame and choose the one where the common value of the total momentum of all the pairs of particles is zero. The paired state $|\Psi_P\rangle$ is then written:

$$|\Psi_P\rangle = \left[A^\dagger_{\mathbf{K}=0}\right]^P |0\rangle \tag{B-2}$$

We shall first study (as in § A-1) the case of bosons or fermions in the same spin state. As for the case of particles in several spin states (as in § A-2), we shall, from now on, limit our study to fermions in a singlet state; this will allows exposing the general principle while avoiding more complex calculations. In both cases, the $2P$-particle state only depends on the values of the parameters $g_{\mathbf{k}}$. As soon as $P > 1$, we will see that the normalization of the ket $|\Psi_P(\mathbf{K})\rangle$ does not reduce to the simple condition (A-3), which required the sum of the $|g_{\mathbf{k}}|^2$ to be equal to unity. This is why we shall consider from now on that the $g_{\mathbf{k}}$ are totally free variational parameters. For example, multiplying them all by the same constant, one can choose to vary at will the norm of $|\Psi_P(\mathbf{K})\rangle$. This will offer a flexibility simplifying the computations.

B-1-a. Particles in the same spin state

For particles in the same spin state, we can use (A-13), which leads to:

$$|\Psi_P\rangle = \left\{\sum_{\mathbf{k}\in D} \sqrt{2} g_{\mathbf{k}}\, a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}}\right\}^P |0\rangle \tag{B-3}$$

where D is the summation domain defined previously (half of the $\mathbf{k}$ space); remember that the physical system is assumed to be contained in a cubic box of side length L and

volume $\mathcal{V} = L^3$; the periodic boundary conditions then fixes all the possible values for the summation over $\mathbf{k}$. Note also that the spin index is implicit: $a^\dagger_{\mathbf{k}}$ is the creation operator in the individual state defined by the momentum $\hbar\mathbf{k}$ and the unique spin state we are concerned with.

Initially, the parameters $g_{\mathbf{k}}$ were introduced as the Fourier components of the normalized pair wave function $\chi(\mathbf{r})$; the sum of their moduli squared was fixed to unity. This condition, however, does not ensure the normalization of $|\Psi_P\rangle$, as we now show. Because of the power P of the operator appearing in (B-3), factors containing square roots of occupation numbers will be introduced for each occurrence of the index $\mathbf{k}$; the ket $|\Psi_P\rangle$ is therefore not a simple tensor product, and its norm is not simply the sum of the squared moduli of the $g_{\mathbf{k}}$ raised to the power P. It will be simpler for the following computations to consider the $g_{\mathbf{k}}$ as entirely free parameters, and hence not impose a normalization of the state $|\Psi_P\rangle$.

On can choose to take a finite or infinite number of non-zero $g_{\mathbf{k}}$. The simplest case is the one already discussed above, where $g_{\mathbf{k}} \propto \delta_{\mathbf{k},\mathbf{k}_0}$; the ket $|\Psi_P\rangle$ then becomes proportional to a simple Fock state where only two states of opposite momenta are occupied. For other functions $g_{\mathbf{k}}$, the structure of the paired state will be more complex; adjusting those parameters allows a fine tuning of the particle correlation properties, which is not possible with a simple Fock state.

B-1-b. Fermions in a singlet state

Another frequently encountered case concerns fermionic particles in a singlet state; we must then use operator (A-23). The paired state is then:

$$|\Psi_P\rangle = \left[A^\dagger_{\mathbf{K}=0}\right]^P |0\rangle \quad = \left\{\sum_{\mathbf{k}} g_{\mathbf{k}}\, a^\dagger_{\mathbf{k},\uparrow} a^\dagger_{-\mathbf{k},\downarrow}\right\}^P |0\rangle \qquad \text{with } N = 2P \tag{B-4}$$

The summation over $\mathbf{k}$ runs over all the non-zero wave vectors, without the restriction (B-3) to the half-space D (because of the spins, the pairs of states $\mathbf{k},\uparrow$, $-\mathbf{k},\downarrow$ and $-\mathbf{k},\uparrow$, $\mathbf{k},\downarrow$ are different).

Here again we see that the normalization of $|\Psi_P\rangle$ does not simply reduce to condition (A-3). When $P > 1$, the same index $\mathbf{k}$ may appear twice (or more) in the expansion of the power P of the operator on the right-hand side of (B-4); the corresponding component cancels out since the square of any fermionic creation operator is zero. The norm of the ket $|\Psi_P\rangle$ is therefore a complex expression. Rather than imposing this norm to be equal to one, it is easier to let it vary, and consider the $g_{\mathbf{k}}$ to be totally free variational parameters.

B-1-c. Consequences of the symmetrization

The state vectors (B-3) for bosons, and (B-4) for fermions, are not simple juxtapositions of P pairs of particles, each being described by the relative wave function $\chi(\mathbf{r})$, with $g_{\mathbf{k}}$ as its Fourier transform, according to (A-2). As we already saw, the symmetrization or antisymmetrization of the $2P$-particle paired states strongly affect their norm; it also affects the very structure of these states, which are not merely the tensor product of P pair states. This is particularly obvious for fermions: expanding the sum of operators to the power P in the curly brackets of (B-4), we get the product of P sums

over indices $\mathbf{k}_1$, $\mathbf{k}_2$,.., $\mathbf{k}_P$, and many terms will cancel out: all those for which two (or more) summation indices $\mathbf{k}$ are equal (in which case we get squares of creation operators, which are zero for fermions).

There exists, however, a limiting case where the paired state practically describes the juxtaposition of P binary molecules. It occurs when the wave function $\chi(\mathbf{r}_1 - \mathbf{r}_2)$ varies over a very short range, and hence has a large number of significant Fourier components $g_{\mathbf{k}}$. When this number is much larger than the number of pairs P, most of the terms do not contain several occurrences of the same summation index $\mathbf{k}$, and consequently the paired state vector is very close to the tensor product of the pairs of particles. This state vector describes a gas formed by very strongly bound binary molecules, each moving in the mean field created by all the others (it is, in a certain sense, a "molecule Fock state"). This is actually a very special case; in general, when the pair wave function does not obey that criterion, we can study many other physical situations, hence the interest for introducing paired states.

Even though the values of $g_{\mathbf{k}}$ or the wave function $\chi(\mathbf{r})$ of a "molecule" are mathematically the starting ingredient that allows building $|\Psi_P\rangle$, the resulting state after symmetrization has a complex structure, hard to describe in terms of molecules. On the other hand, this state has a simple property: it contains exactly $N = 2P$ particles, since this is the case for each of its non-zero components; as all the particles are paired, it contains exactly P pairs.

B-2. Undetermined particle number

Computations with the ket $|\Psi_P\rangle$ written above (and in particular its normalization) are not easy: a great number of individual states $\mathbf{k}$ appear inside the curly brackets, which must be raised to a very large power P. This practical difficulty leads us to introduce another variational state $|\Psi_{\text{paired}}\rangle$ where the total number of particles is no longer fixed. This new state, which leads to simpler calculations[2], is defined, starting with (B-2), by:

$$|\Psi_{\text{paired}}\rangle = \sum_{P=0}^{\infty} \frac{1}{P!} |\Psi_P\rangle = \sum_{P=0}^{\infty} \frac{1}{P!} \left[A^{\dagger}_{\mathbf{K}=0}\right]^P |0\rangle \tag{B-5}$$

The $|\Psi_P\rangle$ are not normalized; multiplying all the $g_{\mathbf{k}}$, and hence $A_{\mathbf{K}=0}$, by the same constant α, changes their norm by the factor α^P. This results in varying the relative weights of the terms in the serie (B-5). The larger α, the more weight is placed on the high values of P, which is a way, for example, of modifying the average particle number.

In (B-5) we recognize the series expansion of an exponential, so that:

$$|\Psi_{\text{paired}}\rangle = \exp\left\{A^{\dagger}_{\mathbf{K}=0}\right\} |0\rangle \tag{B-6}$$

This property will greatly simplify the following calculations and is the major reason for letting the total particle number fluctuate.

Writing (B-5), we chose a state vector that is the superposition of states corresponding to different total particle numbers N; there are actually no physical processes taken into account in our approach that could create such a coherent superposition. This

[2]This does not mean that computations with a variational state having a fixed number of particles are always impossible, as shown for example in the treatment of the BCS theory in § 5.4 and Appendix 5C of the book [8].

operation reminds us of the passage from the canonical to the grand canonical ensemble where one introduces, for mathematical convenience, an (incoherent) statistical mixture of different N values. In our present case, however, we are dealing with a coherent superposition, introduced arbitrarily as we just did, and we may wonder whether it might radically change the physics of our problem. This is actually not the case for two reasons. The first is that, for very large values of N, we are going to show that the components of $|\Psi_{\text{paired}}\rangle$ are only important in a domain of N whose width is very small compared to the average value of the particle number; the distribution of the possible values for N is thus very narrow, in relative value, and the particle number remains quite well defined. The second reason is that we shall compute average values of operators that, such as $\widehat{H}$, conserve the total particle number, and for which the coherence of the state vector between kets of different N values is irrelevant. The average value in the coherent state $|\Psi_{\text{paired}}\rangle$ is therefore the weighted average of the average values obtained for each N which, when the average value of the particle number is very large, are approximately the same (since the N distribution is very narrow). In other words, the average values we are going to compute are good approximations of those we would obtain by projecting $|\Psi_{\text{paired}}\rangle$ onto one of its main components with fixed N; using the coherent superposition (B-5) is thus very convenient from a mathematical point of view, without greatly perturbing the results from a physical point of view. A more detailed discussion of this question will be presented in § 1 of Complement B_{XVII}.

B-2-a. Particles in the same spin state

When all the particles are in the same spin state, inserting (A-13) in (B-6) leads to:

$$|\Psi_{\text{paired}}\rangle = \exp\left\{\sum_{\mathbf{k}\in D} \sqrt{2} g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k}} a^{\dagger}_{-\mathbf{k}}\right\}|0\rangle \tag{B-7}$$

The operators $a^{\dagger}_{\mathbf{k}} a^{\dagger}_{-\mathbf{k}}$ and $a^{\dagger}_{\mathbf{k}'} a^{\dagger}_{-\mathbf{k}'}$ commute with each other (for fermions in the same spin state, two minus signs cancel each other as we commute products of two operators). It then follows that the exponential of the sum is a product of exponentials, and we can write:

$$\begin{aligned} |\Psi_{\text{paired}}\rangle &= \left\{\prod_{\otimes \mathbf{k}\in D} \exp\left[\sqrt{2} g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k}} a^{\dagger}_{-\mathbf{k}}\right]\right\}|0\rangle \\ &= \prod_{\otimes \mathbf{k}\in D} |\varphi_{\mathbf{k}}\rangle \end{aligned} \tag{B-8}$$

The state vector $|\Psi_{\text{paired}}\rangle$ is then simply a tensor product[3] of state vectors $|\varphi_{\mathbf{k}}\rangle$:

$$|\varphi_{\mathbf{k}}\rangle = \exp\left\{\sqrt{2} g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k}} a^{\dagger}_{-\mathbf{k}}\right\}|0\rangle \tag{B-9}$$

[3]The Fock space is the tensor product of the states associated with all the individual quantum states $|\mathbf{k}\rangle$, each having any positive occupation number. One can regroup those spaces in pairs corresponding to opposite values of $\mathbf{k}$, and introduce spaces $\mathcal{E}(\mathbf{k})$ whose tensor product is also the Fock space. To build a basis in those spaces, one must vary two occupation numbers.

The restriction of the summation over $\mathbf{k}$ to a half-space D, introduced above, prevents each component of the tensor product from appearing twice in (B-8).

For fermions in a single spin state, the square of any creation operator is zero; the exponential reduces to the sum of the first two terms of its expansion:

$$|\varphi_{\mathbf{k}}\rangle = \left[1 + \sqrt{2} g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k}} a^{\dagger}_{-\mathbf{k}}\right] |0\rangle \qquad \text{(fermions only)} \tag{B-10}$$

B-2-b. Fermions in a singlet state

For paired fermions in a singlet state, the state $|\Psi_{\text{paired}}\rangle$ will be called the "BCS state" (Complement C_{XVII}) and noted $|\Psi_{\text{BCS}}\rangle$; relation (A-23) must be used with $\mathbf{K} = 0$. As the exponential of a sum of commuting operators[4] is a product of operators, we get:

$$|\Psi_{\text{BCS}}\rangle = \exp\left\{\sum g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right\} |0\rangle = \prod_{\otimes \mathbf{k}} |\varphi_{\mathbf{k}}\rangle \tag{B-11}$$

with:

$$|\varphi_{\mathbf{k}}\rangle = \exp\left[g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle \tag{B-12}$$

As the square of any fermion creation operator is zero, the series expansion of the exponential is limited to its first two terms:

$$|\varphi_{\mathbf{k}}\rangle = \left[1 + g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle \qquad \text{(fermions only)} \tag{B-13}$$

B-3. Pairs of particles and pairs of individual states

Pairs of states is an important concept not to be confused with pairs of particles. In (B-7) as well as in (B-11), the individual states intervene as "pairs of states" $(\mathbf{k}, -\mathbf{k})$. The number of those pairs (which can be infinite if D is infinite) is not related to the particle number. For fermions in a singlet state, it is convenient to label the pair of states by the momentum $\mathbf{k}$ associated with the spin state $\uparrow$, while remembering that the momentum associated with the spin state $\downarrow$ is the opposite, $-\mathbf{k}$. We shall systematically use this simplification in what follows.

C. Properties of the kets characterizing the paired states

Let us examine a few properties of the states $|\varphi_{\mathbf{k}}\rangle$ that will be useful in what follows. To keep things simple, we continue limiting in this § C the generality of the cases under study, and assume the particles in the same spin states are bosons; as for the particles in different spin states, we shall continue using the example of fermions in a singlet state. The generalization to other paired cases does not introduce any particular difficulties.

C-1. Normalization

The normalization of the states $|\varphi_{\mathbf{k}}\rangle$ is actually simpler for fermions than for bosons; this is because, as we shall see below, the series expansion of the exponential (B-12) contains only two terms for fermions, instead of an infinity for bosons. This is why we do not keep in this § C the same order as in § A and start with the study of spin 1/2 fermions.

[4] The operators $a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}$ and $a^{\dagger}_{\mathbf{k}',\uparrow} a^{\dagger}_{-\mathbf{k}',\downarrow}$ associated with different pairs commute, since they are products of two fermionic operators.

C-1-a. Fermions in a singlet state

We choose to normalize separately each of the $|\varphi_{\mathbf{k}}\rangle$ by multiplying them by a number $u_{\mathbf{k}}$. This operation amounts to replacing $|\varphi_{\mathbf{k}}\rangle$ by:

$$|\overline{\varphi}_{\mathbf{k}}\rangle = \left[u_{\mathbf{k}} + v_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle \tag{C-1}$$

with:

$$v_{\mathbf{k}} = u_{\mathbf{k}}\ g_{\mathbf{k}} \tag{C-2}$$

The normalization condition becomes:

$$|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1 \tag{C-3}$$

It then becomes natural to set:

$$\begin{aligned} u_{\mathbf{k}} &= \cos\theta_{\mathbf{k}}\ e^{-i\zeta_{\mathbf{k}}} \\ v_{\mathbf{k}} &= \sin\theta_{\mathbf{k}}\ e^{i\zeta_{\mathbf{k}}} \end{aligned} \tag{C-4}$$

where $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$ are the two variables[5] the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$ depends on. One can choose $\theta_{\mathbf{k}}$ between 0 and $\pi/2$:

$$0 \leq \theta_{\mathbf{k}} \leq \frac{\pi}{2} \tag{C-5}$$

so that $\cos\theta_{\mathbf{k}}$ and $\sin\theta_{\mathbf{k}}$ are positive and represent the moduli of $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$. We saw in § A-2 that $g_{\mathbf{k}} = g_{-\mathbf{k}}$; the functions $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$ are therefore even with respect to $\mathbf{k}$.

The variational ket $|\Psi_{\mathrm{BCS}}\rangle$ now becomes the normalized ket $\left|\overline{\Psi}_{\mathrm{BCS}}\right\rangle$:

$$\begin{aligned} \left|\overline{\Psi}_{\mathrm{BCS}}\right\rangle &= \prod\nolimits_{\otimes \mathbf{k}} \left[u_{\mathbf{k}} + v_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle \\ &= \prod\nolimits_{\otimes \mathbf{k}} \left[\cos\theta_{\mathbf{k}}\ e^{-i\zeta_{\mathbf{k}}} + \sin\theta_{\mathbf{k}}\ e^{i\zeta_{\mathbf{k}}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle \end{aligned} \tag{C-6}$$

Comment:

A particular case occurs when all the $\theta_{\mathbf{k}}$ are either zero or equal to $\pi/2$. The ket $\left|\overline{\Psi}_{\mathrm{BCS}}\right\rangle$ then reduces to a simple Fock state, whose populations of individual states are either zero, or equal to one (for populations corresponding to states belonging to a pair for which $\theta_{\mathbf{k}} = \pi/2$). In that case, the phases $\zeta_{\mathbf{k}}$ no longer play any role: instead of fixing a relative phase, they only determine the global phase of the state vector.

If, furthermore, we choose $\theta_{\mathbf{k}} = \pi/2$ for all values of $\mathbf{k}$ whose modulus is less than a given value k_F, and zero otherwise, the paired state now describes an ensemble of fermions filling two Fermi spheres (one for each spin state), which is simply the ground state of an ideal gas of fermions. The ket $\left|\overline{\Psi}_{\mathrm{BCS}}\right\rangle$ then reduces to the trial ket of the Hartree-Fock method of Complement $\mathrm{B_{XV}}$; that method appears as a particular case of the more general pairing method used in this chapter.

[5] The variable $\zeta_{\mathbf{k}}$ determines the difference $2\zeta_{\mathbf{k}}$ between the phases of $v_{\mathbf{k}}$ and $u_{\mathbf{k}}$. We could also introduce a variable to determine their sum, but that would be pointless: such a variable would only change the total phase of the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$, without any physical consequences.

C-1-b. Bosons in the same spin state

For bosons, the results are slightly different. To maintain a certain analogy, we shall use the same parameters $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$ as for fermions, but it is now the hyperbolic sine and cosine of $\theta_{\mathbf{k}}$ that will come into play. Relation (B-9) leads to:

$$|\varphi_{\mathbf{k}}\rangle = \sum_{q=0}^{\infty} \frac{1}{q!} \left[\sqrt{2} g_{\mathbf{k}}\ a_{\mathbf{k}}^{\dagger} a_{-\mathbf{k}}^{\dagger}\right]^q |0\rangle = \sum_{q=0}^{\infty} \frac{1}{q!} \left[-x_{\mathbf{k}}\right]^q \left[a_{\mathbf{k}}^{\dagger} a_{-\mathbf{k}}^{\dagger}\right]^q |0\rangle \tag{C-7}$$

with[6]:

$$x_{\mathbf{k}} = -\sqrt{2}\ g_{\mathbf{k}} \tag{C-8}$$

As mentioned before, the spin index that comes in addition to the index $\mathbf{k}$ is not written explicitly as its value does not change.

Consequently:

$$\begin{aligned} \langle\varphi_{\mathbf{k}}\ |\varphi_{\mathbf{k}}\rangle &= \sum_{q=0}^{\infty} \left(\frac{1}{q!}\right)^2 |x_{\mathbf{k}}|^{2q} \left(\sqrt{q!}\right)^4 = \sum_{q=0}^{\infty} |x_{\mathbf{k}}|^{2q} \\ &= \frac{1}{1-|x_{\mathbf{k}}|^2} \end{aligned} \tag{C-9}$$

We assumed, to sum the series, that:

$$|x_{\mathbf{k}}|^2 < 1 \tag{C-10}$$

It is useful in what follows to characterize the complex variable $x_{\mathbf{k}}$ by two real variables: $\theta_{\mathbf{k}}$ to define its modulus, and an angle $\zeta_{\mathbf{k}}$ that characterizes its phase. We therefore set:

$$x_{\mathbf{k}} = \tanh\theta_{\mathbf{k}}\ e^{2i\zeta_{\mathbf{k}}} \qquad \text{with:} \quad \theta_{\mathbf{k}} \geq 0 \tag{C-11}$$

Inequality (C-10) is automatically satisfied since the modulus of a hyperbolic tangent is always less than 1; as the function $g_{\mathbf{k}}$ is even – see relation (A-4) – so are the variable $x_{\mathbf{k}}$ and the functions[7] $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$. We then get:

$$\langle\varphi_{\mathbf{k}}\ |\varphi_{\mathbf{k}}\rangle = \frac{1}{1-\tanh^2\theta_{\mathbf{k}}} = \cosh^2\theta_{\mathbf{k}} \tag{C-12}$$

The normalized kets $|\overline{\varphi}_{\mathbf{k}}\rangle$ can be written as:

$$|\overline{\varphi}_{\mathbf{k}}\rangle = \frac{1}{\cosh\theta_{\mathbf{k}}} |\varphi_{\mathbf{k}}\rangle = \frac{1}{\cosh\theta_{\mathbf{k}}} \exp\left[-x_{\mathbf{k}}\ a_{\mathbf{k}}^{\dagger} a_{-\mathbf{k}}^{\dagger}\right] |0\rangle \tag{C-13}$$

Replacing the $|\varphi_{\mathbf{k}}\rangle$ by the $|\overline{\varphi}_{\mathbf{k}}\rangle$, the ket $|\Psi_{\text{paired}}\rangle$ becomes normalized to 1.

Initially, the kets $|\varphi_{\mathbf{k}}\rangle$, as well as their normalized version $|\overline{\varphi}_{\mathbf{k}}\rangle$, have been defined in the tensor product (B-8) only when $\mathbf{k}$ belongs to the half-space D. They can, however, be defined by relations (C-7) and (C-13) for any $\mathbf{k}$; we then have simply $|\overline{\varphi}_{\mathbf{k}}\rangle = |\overline{\varphi}_{-\mathbf{k}}\rangle$, which was to be expected since $|\overline{\varphi}_{\mathbf{k}}\rangle$ involves the two individual states $\mathbf{k}$ and $-\mathbf{k}$ in the same way.

[6] The minus sign in this definition is arbitrary – a change of sign of the wave function $\chi(\mathbf{r})$ or of its Fourier transform $g_{\mathbf{k}}$ has no physical consequences – but it is convenient to introduce this sign to ensure coherence with the calculations in § E.

[7] Furthermore, rotational invariance generally requires those functions to depend only on the modulus k of $\mathbf{k}$.

C-2. Average value and root mean square deviation of particle number

The particle number in the individual state $\mathbf{k}$ corresponds to the operator:

$$\widehat{n}_{\mathbf{k}} = a^{\dagger}_{\mathbf{k}} a_{\mathbf{k}} \tag{C-14}$$

We are now going to compute the average value and the root mean square deviation of the particle number, first in a given pair of states, then for the system as a whole.

C-2-a. Fermions in a singlet state

Let us compute the average value of the particle number in the state $\left|\overline{\Psi}_{\mathrm{BCS}}\right\rangle$, which is the tensor product of the states $|\overline{\varphi}_{\mathbf{k}}\rangle$, each being associated with the pair of states $(\hbar\mathbf{k}, \nu = +; -\hbar\mathbf{k}, \nu = -)$; as defined above, each pair is labeled by the wave vector $\mathbf{k}$ of the spin $+$ particle. The particle number in each of these pairs of states corresponds to the operator:

$$\widehat{n}_{(\text{pair } \mathbf{k})} = \widehat{n}_{\mathbf{k},\uparrow} + \widehat{n}_{-\mathbf{k},\downarrow} = a^{\dagger}_{\mathbf{k},\uparrow} a_{\mathbf{k},\uparrow} + a^{\dagger}_{-\mathbf{k},\downarrow} a_{-\mathbf{k},\downarrow} \tag{C-15}$$

with eigenvalues 0, 1 and 2. Now $|\overline{\varphi}\rangle_{\mathbf{k}}$ is given by (C-1), the sum of two components, one with zero particles, and the other with two particles. This leads to:

$$\langle\overline{\varphi}_{\mathbf{k}}|\, \widehat{n}_{(\text{pair } \mathbf{k})}\, |\overline{\varphi}_{\mathbf{k}}\rangle = 2\,|v_{\mathbf{k}}|^2 = 2\sin^2\theta_{\mathbf{k}} \tag{C-16}$$

and:

$$\langle\overline{\varphi}_{\mathbf{k}}|\, \left[\widehat{n}_{(\text{pair } \mathbf{k})}\right]^2 |\overline{\varphi}_{\mathbf{k}}\rangle = 4\,|v_{\mathbf{k}}|^2 = 4\sin^2\theta_{\mathbf{k}} \tag{C-17}$$

The root mean square deviation $\Delta n_{(\text{pair } \mathbf{k})}$ of the particle number in a pair is thus:

$$\Delta n_{(\text{pair } \mathbf{k})} = \sqrt{4\,|v_{\mathbf{k}}|^2 \left[1 - |v_{\mathbf{k}}|^2\right]} = 2\sin\theta_{\mathbf{k}}\cos\theta_{\mathbf{k}} \tag{C-18}$$

Consequently, the fluctuations of the particle number in each pair of states can be large.

On the other hand, the fluctuations of the total particle number, obtained by summing over all the pairs, remain small. The average value of this total number is:

$$\left\langle \widehat{N} \right\rangle = 2\sum_{\mathbf{k}} |v_{\mathbf{k}}|^2 = 2\sum_{\mathbf{k}} \sin^2\theta_{\mathbf{k}} \tag{C-19}$$

As we will show just below, the square of the fluctuation ΔN of $\langle N\rangle$ is given by:

$$[\Delta N]^2 = 4\sum_{\mathbf{k}} |v_{\mathbf{k}}|^2 \left[1 - |v_{\mathbf{k}}|^2\right] \tag{C-20}$$

Since $\left[1 - |v_{\mathbf{k}}|^2\right] \leq 1$, we get:

$$[\Delta N]^2 \leq 4\sum_{\mathbf{k}} |v_{\mathbf{k}}|^2 = 2\left\langle \widehat{N} \right\rangle \tag{C-21}$$

so that:

$$\frac{\Delta N}{\langle \widehat{N} \rangle} \leq \sqrt{\frac{2}{\langle \widehat{N} \rangle}} \tag{C-22}$$

Hence, for large values of $\langle N \rangle$, the fluctuations of the particle number, in relative value, are very small, decreasing at least as fast as the inverse of the square root of the average value.

Demonstration:

The operator corresponding to the square of the particle number is:

$$\left[\widehat{N}\right]^2 = \sum_{\mathbf{k}} \left[\widehat{n}_{(\text{pair } \mathbf{k})}\right]^2 + \sum_{\mathbf{k} \neq \mathbf{k}'} \widehat{n}_{(\text{pair } \mathbf{k})}\ \widehat{n}_{(\text{pair } \mathbf{k}')} \tag{C-23}$$

As the state $\left|\overline{\Psi}_{\text{BCS}}\right\rangle$ is a product of states of pairs, the latter are not correlated and the average value of this operator is written:

$$\begin{aligned} &\left\langle \overline{\Psi}_{\text{BCS}} \right| \left[\widehat{N}\right]^2 \left| \overline{\Psi}_{\text{BCS}} \right\rangle \\ &= \sum_{\mathbf{k}} \langle \overline{\varphi_{\mathbf{k}}} | \left[\widehat{n}_{(\text{pair } \mathbf{k})}\right]^2 | \overline{\varphi}_{\mathbf{k}} \rangle + \sum_{\mathbf{k} \neq \mathbf{k}'} \langle \overline{\varphi}_{\mathbf{k}} | \widehat{n}_{(\text{pair } \mathbf{k})} | \overline{\varphi}_{\mathbf{k}} \rangle \langle \overline{\varphi}_{\mathbf{k}'} | \widehat{n}_{(\text{pair } \mathbf{k}')} | \overline{\varphi}_{\mathbf{k}'} \rangle \end{aligned} \tag{C-24}$$

Expression (C-1) for $|\overline{\varphi}_{\mathbf{k}}\rangle$ leads to:

$$\widehat{n}_{(\text{pair } \mathbf{k})} | \overline{\varphi}_{\mathbf{k}} \rangle = 2 v_{\mathbf{k}} | \mathbf{k}, \uparrow; -\mathbf{k}, \downarrow \rangle \tag{C-25}$$

so that:

$$\left\langle \left[\widehat{N}\right]^2 \right\rangle = 4 \sum_{\mathbf{k}} |v_{\mathbf{k}}|^2 + 4 \sum_{\mathbf{k} \neq \mathbf{k}'} |v_{\mathbf{k}}|^2 |v_{\mathbf{k}'}|^2 \tag{C-26}$$

Now the square of the average value $\langle N \rangle$ is equal to the last terms of this equality, but without the constraint $\mathbf{k} \neq \mathbf{k}'$ in the summation. It follows that the root mean square ΔN is written as:

$$[\Delta N]^2 = \left\langle \left[\widehat{N}\right]^2 \right\rangle - [\langle N \rangle]^2 = 4 \sum_{\mathbf{k}} \left[|v_{\mathbf{k}}|^2 - |v_{\mathbf{k}}|^4 \right] \tag{C-27}$$

which leads to (C-20).

C-2-b. Bosons in the same spin state

For bosons, each pair contains two individual states of opposite $\mathbf{k}$. We show below that for each of them, we have:

$$\langle \widehat{n}_{\mathbf{k}} \rangle = \sinh^2 \theta_{\mathbf{k}} \tag{C-28}$$

and that:

$$\left\langle [\widehat{n}_{\mathbf{k}}]^2 \right\rangle = 2 \langle \widehat{n}_{\mathbf{k}} \rangle^2 + \langle \widehat{n}_{\mathbf{k}} \rangle \tag{C-29}$$

The root mean square deviation of the distribution associated with the values of $n_{\mathbf{k}}$ is thus:

$$\begin{aligned}\Delta n_{\mathbf{k}} &= \sqrt{\left\langle \left[\widehat{n}_{\mathbf{k}}\right]^2 \right\rangle - \left\langle \widehat{n}_{\mathbf{k}} \right\rangle^2} = \sqrt{\left\langle \widehat{n}_{\mathbf{k}} \right\rangle^2 + \left\langle \widehat{n}_{\mathbf{k}} \right\rangle} \\ &= \sinh\theta_{\mathbf{k}} \cosh\theta_{\mathbf{k}} \end{aligned} \tag{C-30}$$

(the average value of the particle number in a pair of states is $2\left\langle \widehat{n}_{\mathbf{k}} \right\rangle$, and the root mean square deviation of that number is $2\Delta n_{\mathbf{k}}$).

Demonstration:

As $|\varphi_{\mathbf{k}}\rangle$ is symmetric with respect to the two individual states $\mathbf{k}$ and $-\mathbf{k}$, we have:

$$\left\langle \widehat{n}_{\mathbf{k}} \right\rangle = \left\langle \widehat{n}_{-\mathbf{k}} \right\rangle \tag{C-31}$$

with:

$$\begin{aligned}\left\langle \varphi_{\mathbf{k}} \right| \widehat{n}_{\mathbf{k}} \left| \varphi_{\mathbf{k}} \right\rangle &= \sum_{q=0}^{\infty} \left|x_{\mathbf{k}}\right|^{2q} \times q = \left|x_{\mathbf{k}}\right|^2 \frac{\partial}{\partial \left|x_{\mathbf{k}}\right|^2} \left\langle \varphi_{\mathbf{k}} \right| \left. \varphi_{\mathbf{k}} \right\rangle \\ &= \frac{\left|x_{\mathbf{k}}\right|^2}{\left[1 - \left|x_{\mathbf{k}}\right|^2\right]^2} \end{aligned} \tag{C-32}$$

so that:

$$\left\langle \widehat{n}_{\mathbf{k}} \right\rangle = \frac{\left\langle \varphi_{\mathbf{k}} \right| \widehat{n}_{\mathbf{k}} \left| \varphi_{\mathbf{k}} \right\rangle}{\left\langle \varphi_{\mathbf{k}} \right| \left. \varphi_{\mathbf{k}} \right\rangle} = \frac{\left|x_{\mathbf{k}}\right|^2}{1 - \left|x_{\mathbf{k}}\right|^2} \tag{C-33}$$

which leads to (C-28).

The average value of the particle number squared is computed in a similar way. Using the identity $q^2 = q(q-1) + q$ to bring up the second derivative with respect to $\left|x_{\mathbf{k}}\right|^2$, we can write:

$$\begin{aligned}\left\langle \varphi_{\mathbf{k}} \right| \left[\widehat{n}_{\mathbf{k}}\right]^2 \left| \varphi_{\mathbf{k}} \right\rangle &= \sum_{q=0}^{\infty} \left|x_{\mathbf{k}}\right|^{2q} \times q^2 \\ &= \left|x_{\mathbf{k}}\right|^4 \frac{\partial^2}{\left[\partial \left|x_{\mathbf{k}}\right|^2\right]^2} \left\langle \varphi_{\mathbf{k}} \right| \left. \varphi_{\mathbf{k}} \right\rangle + \left|x_{\mathbf{k}}\right|^2 \frac{\partial}{\partial \left|x_{\mathbf{k}}\right|^2} \left\langle \varphi_{\mathbf{k}} \right| \left. \varphi_{\mathbf{k}} \right\rangle \\ &= \frac{2\left|x_{\mathbf{k}}\right|^4}{\left[1 - \left|x_{\mathbf{k}}\right|^2\right]^3} + \frac{\left|x_{\mathbf{k}}\right|^2}{\left[1 - \left|x_{\mathbf{k}}\right|^2\right]^2} \end{aligned} \tag{C-34}$$

and hence:

$$\left\langle \left[\widehat{n}_{\mathbf{k}}\right]^2 \right\rangle = \frac{\left\langle \varphi_{\mathbf{k}} \right| \left[\widehat{n}_{\mathbf{k}}\right]^2 \left| \varphi_{\mathbf{k}} \right\rangle}{\left\langle \varphi_{\mathbf{k}} \right| \left. \varphi_{\mathbf{k}} \right\rangle} = 2\left\langle \widehat{n}_{\mathbf{k}} \right\rangle^2 + \left\langle \widehat{n}_{\mathbf{k}} \right\rangle \tag{C-35}$$

The total number of particles is written:

$$\left\langle \widehat{N} \right\rangle = \sum_{\mathbf{k}} \left|v_{\mathbf{k}}\right|^2 = \sum_{\mathbf{k}} \sinh^2\theta_{\mathbf{k}} \tag{C-36}$$

(as expected, each pair of states appears twice in this summation; one can also restrict the summation to the half-space D provided we add a factor 2). We also have:

$$\left\langle \left[\widehat{N}\right]^2 \right\rangle = \left\langle \sum_{\mathbf{k}} \widehat{n}_{\mathbf{k}} \sum_{\mathbf{k}'} \widehat{n}_{\mathbf{k}'} \right\rangle = \left\langle \sum_{\mathbf{k}} \widehat{n}_{\mathbf{k}} \left[\widehat{n}_{\mathbf{k}} + \widehat{n}_{-\mathbf{k}} + \sum_{\mathbf{k}' \neq \pm \mathbf{k}} \widehat{n}_{\mathbf{k}'} \right] \right\rangle$$
$$= \sum_{\mathbf{k}} \left[\langle [\widehat{n}_{\mathbf{k}}]^2 \rangle + \langle \widehat{n}_{\mathbf{k}} \widehat{n}_{-\mathbf{k}} \rangle \right] + \sum_{\mathbf{k}, \mathbf{k}' \neq \pm \mathbf{k}} \langle \widehat{n}_{\mathbf{k}} \rangle \langle \widehat{n}_{\mathbf{k}'} \rangle \tag{C-37a}$$

where, in the last term, we have used the fact that the paired state is the product of uncorrelated pairs. But this state is symmetric in $\mathbf{k}$ and $-\mathbf{k}$, and operators $\widehat{n}_{\mathbf{k}}$ and $\widehat{n}_{-\mathbf{k}}$ act on it in the same way. Therefore:

$$\left\langle \left[\widehat{N}\right]^2 \right\rangle = \sum_{\mathbf{k}} 2 \left[\left\langle [\widehat{n}_{\mathbf{k}}]^2 \right\rangle - \langle \widehat{n}_{\mathbf{k}} \rangle^2 \right] + \sum_{\mathbf{k}, \mathbf{k}'} \langle \widehat{n}_{\mathbf{k}} \rangle \langle \widehat{n}_{\mathbf{k}'} \rangle \tag{C-37b}$$

(the constraint in the second summmation has been eliminated by subtracting a term in the first summation). We then get:

$$\left\langle \left[\widehat{N}\right]^2 \right\rangle = 2 \sum_{\mathbf{k}} [\Delta n_{\mathbf{k}}]^2 + \langle N \rangle^2 \tag{C-37c}$$

The root mean square deviations $\Delta n_{\mathbf{k}}$ have been obtained in (C-30). Hence, the square of the root mean square deviation of the total number of particles is written as:

$$[\Delta N]^2 = 2 \sum_{\mathbf{k}} [\Delta n_{\mathbf{k}}]^2 = 2 \sum_{\mathbf{k}} \sinh^2 \theta_{\mathbf{k}} \cosh^2 \theta_{\mathbf{k}} \tag{C-38}$$

As for the fermion case, this square contains only a single summation on $\mathbf{k}$, whereas the square of the total particle number contains two. Now the number of non-zero terms in those summations is the number of Fourier components necessary to describe the pair of particles used, in § B, to build the paired state in a cube of edge length L (size of the momentum quantization box – see § A-1). This number is of the order of the cube of the ratio between L and the size of the pair, hence a very large number, as it is the ratio between a macroscopic and a microscopic volume. A double summation over $\mathbf{k}$ therefore contains many more terms than a simple summation, and since all the terms are positive and of comparable magnitude, we have:

$$\left\langle \widehat{N} \right\rangle^2 \gg [\Delta N]^2 \tag{C-39}$$

We again find, as for fermions, that $\Delta N \ll \left\langle \widehat{N} \right\rangle$.

C-3. "Anomalous" average values

For computing average values of the energy (in particular, in Complement $\mathrm{B_{XVII}}$), we will need the average values of products of two creation or annihilation operators. For example, for bosons we will need to calculate:

$$\langle \overline{\varphi}_{\mathbf{k}} |\, a_{\mathbf{k}}^\dagger a_{-\mathbf{k}}^\dagger \,| \overline{\varphi}_{\mathbf{k}} \rangle \qquad \text{and} \qquad \langle \overline{\varphi}_{\mathbf{k}} |\, a_{\mathbf{k}} a_{-\mathbf{k}} \,| \overline{\varphi}_{\mathbf{k}} \rangle \tag{C-40}$$

We note, right away, that they concern operators that do not conserve the particle number, and this is the reason they are often called "anomalous average values". One could be surprised that such average values come into play while studying physical processes that do not physically imply creation or destruction of particles. We will show that they actually occur in a very natural way in the calculation of the average value of a Hamiltonian that conserves the particle number. The reason is that $|\overline{\varphi}_{\mathbf{k}}\rangle$ is only a component of the total state vector (B-8), in which it is associated with many other $|\overline{\varphi}_{\mathbf{k}'}\rangle$; in the total state vector, the particle number in the state $|\overline{\varphi}_{\mathbf{k}}\rangle$ may, for example, decrease by 2 while the particle number in the state $|\overline{\varphi}_{\mathbf{k}'}\rangle$ simultaneously increases by the same quantity. We are, therefore, performing computations on the components of a state vector that has the same total particle number; the "anomalous" character is only apparent, and is due to the fact that we only consider part of the total state vector.

C-3-a. Fermions in a singlet state

Consider the action of the operator $a_{-\mathbf{k},\downarrow}a_{\mathbf{k},\uparrow}$ on the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$ written in (C-1). Only one of its component, in $v(\mathbf{k})$, remains and, after two anticommutations, we can write:

$$\begin{aligned} a_{-\mathbf{k},\downarrow}a_{\mathbf{k},\uparrow}|\overline{\varphi}_{\mathbf{k}}\rangle &= v_{\mathbf{k}}a_{-\mathbf{k},\downarrow}a_{\mathbf{k},\uparrow}a^{\dagger}_{\mathbf{k},\uparrow}a^{\dagger}_{-\mathbf{k},\downarrow}|0\rangle = v_{\mathbf{k}}\left[a_{-\mathbf{k},\downarrow}a^{\dagger}_{-\mathbf{k},\downarrow}\right]\left[a_{\mathbf{k},\uparrow}a^{\dagger}_{\mathbf{k},\uparrow}\right]|0\rangle \\ &= v_{\mathbf{k}}|0\rangle \end{aligned} \tag{C-41}$$

Taking the scalar product of this ket with the bra $\langle\overline{\varphi}_{\mathbf{k}}|$, only its component $u^{*}(\mathbf{k})\langle 0|$ remains; the average value is thus:

$$\langle\overline{\varphi}_{\mathbf{k}}|\, a_{-\mathbf{k},\downarrow}a_{\mathbf{k},\uparrow}|\overline{\varphi}_{\mathbf{k}}\rangle = u^{*}_{\mathbf{k}}v_{\mathbf{k}} = \sin\theta_{\mathbf{k}}\cos\theta_{\mathbf{k}}\; e^{2i\zeta_{\mathbf{k}}} \tag{C-42}$$

The anticommutation of these two operators then yields:

$$\langle\overline{\varphi}_{\mathbf{k}}|\, a_{\mathbf{k},\uparrow}a_{-\mathbf{k},\downarrow}|\overline{\varphi}_{\mathbf{k}}\rangle = -u^{*}_{\mathbf{k}}v_{\mathbf{k}} = -\sin\theta_{\mathbf{k}}\cos\theta_{\mathbf{k}}\; e^{2i\zeta_{\mathbf{k}}} \tag{C-43}$$

Taking the Hermitian conjugate of (C-42), we get:

$$\langle\overline{\varphi}_{\mathbf{k}}|\, a^{\dagger}_{\mathbf{k},\uparrow}a^{\dagger}_{-\mathbf{k},\downarrow}|\overline{\varphi}_{\mathbf{k}}\rangle = u_{\mathbf{k}}v^{*}_{\mathbf{k}} = \sin\theta_{\mathbf{k}}\cos\theta_{\mathbf{k}}\; e^{-2i\zeta_{\mathbf{k}}} \tag{C-44}$$

whereas the average value of $a^{\dagger}_{-\mathbf{k},\downarrow}a^{\dagger}_{\mathbf{k},\uparrow}$ is the opposite (anticommutation):

$$\langle\overline{\varphi}_{\mathbf{k}}|\, a^{\dagger}_{-\mathbf{k},\downarrow}a^{\dagger}_{\mathbf{k},\uparrow}|\overline{\varphi}_{\mathbf{k}}\rangle = -\sin\theta_{\mathbf{k}}\cos\theta_{\mathbf{k}}\; e^{-2i\zeta_{\mathbf{k}}} \tag{C-45}$$

We saw, in § C-1-a, that the functions $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$ are even; we can therefore change the sign of $\mathbf{k}$ on the left-hand side of the previous relations without changing the right-hand side.

C-3-b. Bosons in the same spin state

For bosons, it is easier to first compute the average value of a product of creation operators:

$$\langle\overline{\varphi}_{\mathbf{k}}|\, a^{\dagger}_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}|\overline{\varphi}_{\mathbf{k}}\rangle = \frac{1}{\cosh^{2}\theta_{\mathbf{k}}}\langle\varphi_{\mathbf{k}}|\, a^{\dagger}_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}|\varphi_{\mathbf{k}}\rangle \tag{C-46}$$

This expression contains the product of the ket:

$$a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} |\varphi_{\mathbf{k}}\rangle = a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} \sum_{q=0}^{\infty} [-x_{\mathbf{k}}]^q |n_{\mathbf{k}} = q; n_{-\mathbf{k}} = q\rangle$$
$$= \sum_{q=0}^{\infty} (q+1) [-x_{\mathbf{k}}]^q |n_{\mathbf{k}} = q+1; n_{-\mathbf{k}} = q+1\rangle \tag{C-47}$$

by the bra:

$$\sum_{q'=0}^{\infty} [-x^*_{\mathbf{k}}]^{q'} \langle n_{\mathbf{k}} = q'; n_{-\mathbf{k}} = q'| \tag{C-48}$$

To get a non-zero term, we must have $q' = q+1$, which leads to:

$$(q+1) [-x_{\mathbf{k}}]^q [-x^*_{\mathbf{k}}]^{q+1} \tag{C-49}$$

whose sum over q yields, taking (C-32) into account:

$$\sum_{q=0}^{\infty} (q+1) \quad [-x^*_{\mathbf{k}}] \quad |x_{\mathbf{k}}|^{2q} = [-x^*_{\mathbf{k}}] \quad [\langle n_{\mathbf{k}}\rangle + 1] \langle \varphi_{\mathbf{k}} | \varphi_{\mathbf{k}}\rangle \tag{C-50}$$

We finally divide by $\langle \varphi_{\mathbf{k}} | \varphi_{\mathbf{k}}\rangle$ as in (C-46), to obtain, inserting the value (C-11) of $x_{\mathbf{k}}$:

$$\langle \overline{\varphi}_{\mathbf{k}}| a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} |\overline{\varphi}_{\mathbf{k}}\rangle = -\tanh\theta_{\mathbf{k}}\, e^{-2i\zeta_{\mathbf{k}}} \left[1 + \sinh^2\theta_{\mathbf{k}}\right]$$
$$= -e^{-2i\zeta_{\mathbf{k}}} \sinh\theta_{\mathbf{k}} \cosh\theta_{\mathbf{k}} \tag{C-51}$$

As for the other "anomalous" average value $\langle \overline{\varphi}_{\mathbf{k}}| a_{\mathbf{k}} a_{-\mathbf{k}} |\overline{\varphi}_{\mathbf{k}}\rangle$, a simple Hermitian conjugation operation shows that it is the complex conjugate of the previous one:

$$\langle \overline{\varphi}_{\mathbf{k}}| a_{\mathbf{k}} a_{-\mathbf{k}} |\overline{\varphi}_{\mathbf{k}}\rangle = -e^{2i\zeta_{\mathbf{k}}} \sinh\theta_{\mathbf{k}} \cosh\theta_{\mathbf{k}} \tag{C-52}$$

As was the case for fermions, the functions $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$ are even, which allows changing the sign of $\mathbf{k}$ on the left-hand side of the previous equations without changing the result.

D. Correlations between particles, pair wave function

As already mentioned in this chapter's introduction, one of the major interest of the paired states is to allow varying the spatial correlation functions of a system of identical particles. In addition to the purely statistical correlations, coming from the indistinguishability of the particles and already present in an ideal gas, we now have a way to include dynamic correlations due to the interactions. Using paired states instead of simple Fock states allows, for example, a better optimization of the energy. We shall limit our study to the two-particle diagonal correlation function, as it is the one that fixes the average value of the interaction Hamiltonian. This will lead us to introduce a new wave function, that we shall name the "pair wave function". In the complements following

this chapter we shall also study non-diagonal correlation functions; it will concern the one-particle correlation function, whose long range behavior may signal the existence of Bose-Einstein condensation, as well as the two-particle correlation function.

In a general way, one may wonder about the physical significance of correlation functions computed in states $|\Psi_{\text{paired}}\rangle$ or $|\Psi_{\text{BCS}}\rangle$, since these states are coherent superpositions of kets containing different particle numbers N. However, correlation functions are average values of operators keeping the particle number constant, and hence independent of the coherence between kets of different N values. Furthermore, we saw in § C-2 that for large values of the average particle number $\left\langle \widehat{N} \right\rangle$, the relative fluctuations of that number were negligible. In the limit of large $\left\langle \widehat{N} \right\rangle$, one can thus expect the results obtained with $|\Psi_{\text{paired}}\rangle$ or $|\Psi_{\text{BCS}}\rangle$ to be very close to those obtained with the $|\Psi_P\rangle$, for which these fluctuations are strictly zero. This question will be discussed in more detail in § 1 of Complement B_{XVII}.

When studying correlation functions in the case where the paired particles are in the same spin state, the only relevant indices concern the orbital variables. We shall start with this simpler case, and study later the case of paired particles in a singlet state.

D-1. Particles in the same spin state

Relation (B-34) of Chapter XVI indicates that the two-particle diagonal correlation function $G_2(\mathbf{r}_1, \mathbf{r}_2)$ can be written:

$$G_2(\mathbf{r}_1, \mathbf{r}_2) = \left\langle \Psi^\dagger(\mathbf{r}_1)\, \Psi^\dagger(\mathbf{r}_2)\, \Psi(\mathbf{r}_2)\, \Psi(\mathbf{r}_1) \right\rangle \tag{D-1}$$

Replacing the field operators and their adjoints by expressions (A-3) and (A-6) of Chapter XVI, using as a basis the normalized plane waves, we get:

$$G_2(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{L^6} \sum_{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3, \mathbf{k}_4} e^{i[(\mathbf{k}_4 - \mathbf{k}_1)\cdot \mathbf{r}_1 + (\mathbf{k}_3 - \mathbf{k}_2)\cdot \mathbf{r}_2]} \left\langle a^\dagger_{\mathbf{k}_1} a^\dagger_{\mathbf{k}_2} a_{\mathbf{k}_3} a_{\mathbf{k}_4} \right\rangle \tag{D-2}$$

where the average value of the product of the 4 creation and annihilation operators must be taken in a paired state. Figure 1 symbolizes the different terms present in this correlation function.

D-1-a. Simplifications due to pairing

The computation is greatly simplified by noting that in a paired state, the populations of the two individual states having opposite wave numbers $\mathbf{k}$ and $-\mathbf{k}$ must always be equal. Consequently, only those combinations of the 4 operators that do not change this equality will lead to non-zero average values. Three cases are then possible:

– **Case I**: the two annihilation operators concern two individual states that do not belong to the same pair ($\mathbf{k}_3 \neq \pm \mathbf{k}_4$); the two creation operators must then restore to their initial values the populations of these two same states, or else their average value will be zero; these are the so called "forward scattering" terms. We then have either $\mathbf{k}_4 = \mathbf{k}_1$ and $\mathbf{k}_3 = \mathbf{k}_2$ (direct term), or $\mathbf{k}_4 = \mathbf{k}_2$ and $\mathbf{k}_3 = \mathbf{k}_1$ (exchange term).

– **Case II**: the two annihilation operators act on the two states of a first pair ($\mathbf{k}_4 = -\mathbf{k}_3$), and the creation operators on the two states of another pair ($\mathbf{k}_2 = -\mathbf{k}_1$). We then talk about a "pair annihilation-creation process" .

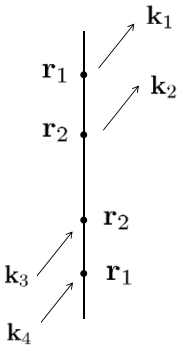


Figure 1: This diagram symbolizes the terms that come into play in the computation of the correlation function of two particles at points $\mathbf{r}_1$ *and* $\mathbf{r}_1$*. The two incoming arrows at the bottom left-hand side represent the two particles eliminated by the annihilation operators; they are associated with a positive imaginary exponent of the position. The two outgoing arrows on the top right-hand side represent the two particles resulting from the action of the creation operators, associated with a negative imaginary exponent. The correlation function is the sum of these terms over all the values of the* 4 $\mathbf{k}$ *vecteurs.*

– **Case III**: the two annihilation operators act on the two states of the same pair, and the creation operators replenish these same two states (this is a special case of the one we just discussed); another possibility is that the two annihilation operators act on the same individual state (all the wave numbers $\mathbf{k}$ must then be equal).

Using these conditions on the values of the wave numbers in (D-2), we note that the terms corresponding to cases I and II include two summations over the wave vectors, whereas there is only one summation in the terms corresponding to case III. Consequently, for a large (macroscopic) volume L^3, there are far fewer terms coming from case III than from cases I and II. We shall therefore only take into account terms arising from case I and II. For the same reason, we shall ignore in our computation of these terms the constraints $\mathbf{k}_3 \neq \pm\mathbf{k}_4$ or $\mathbf{k}_3 \neq \pm\mathbf{k}_1$, as this amounts to adding a negligible number of terms.

D-1-b. Expression of the correlation function

The direct term is obtained for $\mathbf{k}_4 = \mathbf{k}_1$ and $\mathbf{k}_3 = \mathbf{k}_2$; it no longer has any spatial dependence. Since $\mathbf{k}_1$ and $\mathbf{k}_2$ are different, the average value of the product of operators can also be written $\left\langle a^\dagger_{\mathbf{k}_1} a_{\mathbf{k}_1} a^\dagger_{\mathbf{k}_2} a_{\mathbf{k}_2} \right\rangle$ – for fermions, the two minus signs coming from the anticommutations cancel each other. Now relation (B-8) shows that the paired state is a tensor product of pairs of states. This means that the average value we wish to determine is simply the product of the average values of the first two operators and of the last two operators, i.e. the product of the average values of two occupation numbers. We thus

get a first contribution:

$$G_2^{\text{dir}}(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{L^6}\sum_{\mathbf{k}_1,\mathbf{k}_2}\langle\widehat{n}_{\mathbf{k}_1}\rangle\langle\widehat{n}_{\mathbf{k}_2}\rangle = \frac{\left\langle\widehat{N}\right\rangle^2}{L^6} \tag{D-3}$$

where the summation over $\mathbf{k}_1$ and $\mathbf{k}_2$ are considered as independent, since as we mentioned above, we can neglect the constraint linking these two indices.

The exchange term is obtained for $\mathbf{k}_4 = \mathbf{k}_2$ and $\mathbf{k}_3 = \mathbf{k}_1$; it exhibits a spatial dependence. As we did for the direct term, we regroup the creation and annihilation operators acting on the same individual states, but this operation now involves only one commutation between operators. We then introduce a factor η, equal to -1 for fermions, and we get:

$$G_2^{\text{ex}}(\mathbf{r}_1,\mathbf{r}_2) = \frac{\eta}{L^6}\sum_{\mathbf{k}_1,\mathbf{k}_2} e^{i(\mathbf{k}_2-\mathbf{k}_1)\cdot(\mathbf{r}_1-\mathbf{r}_2)}\ \langle\widehat{n}_{\mathbf{k}_1}\rangle\langle\widehat{n}_{\mathbf{k}_2}\rangle \tag{D-4}$$

The pair annihilation-creation term $\mathbf{k}_4 = -\mathbf{k}_3$ and $\mathbf{k}_2 = -\mathbf{k}_1$ also exhibits a spatial dependence, but no longer involves average values of occupation numbers. Its expression is:

$$G_2^{\text{pair-pair}}(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{L^6}\sum_{\mathbf{k}_1,\mathbf{k}_4} e^{i(\mathbf{k}_4-\mathbf{k}_1)\cdot(\mathbf{r}_1-\mathbf{r}_2)}\ \left\langle a^\dagger_{\mathbf{k}_1}a^\dagger_{-\mathbf{k}_1}\right\rangle\left\langle a_{-\mathbf{k}_4}a_{\mathbf{k}_4}\right\rangle \tag{D-5}$$

and its structure is schematized in Figure 2. Expression (D-5) contains average values of products of operators that do not conserve the particle number, but rather annihilate (or create) two of them. They are called "anomalous average values". As we explained in § C-3, these anomalous average values come into play quite naturally in the computation of the average value of an operator that does conserve the particle number. Defining the "pair wave function" ϕ_{pair} as:

$$\phi_{\text{pair}}(\mathbf{r}) = \frac{1}{L^3}\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{r}}\langle a_{\mathbf{k}}a_{-\mathbf{k}}\rangle \tag{D-6}$$

this correlation function can be written as:

$$G_2^{\text{pair-pair}}(\mathbf{r}_1,\mathbf{r}_2) = \left|\phi_{\text{pair}}(\mathbf{r}_1-\mathbf{r}_2)\right|^2 \tag{D-7}$$

The complete correlation function $G_2(\mathbf{r}_1,\mathbf{r}_2)$ is the sum of the three previous contributions:

$$G_2(\mathbf{r}_1,\mathbf{r}_2) = G_2^{\text{dir}}(\mathbf{r}_1,\mathbf{r}_2) + G_2^{\text{ex}}(\mathbf{r}_1,\mathbf{r}_2) + G_2^{\text{pair-pair}}(\mathbf{r}_1,\mathbf{r}_2) \tag{D-8}$$

For bosons in the same spin state, we can insert in this correlation function the average values given in (C-28), (C-51) and (C-52). We then get a binary correlation function that explicitly depends on the parameters $\theta_{\mathbf{k}}$, as well as on the phases $\zeta_{\mathbf{k}}$, which both define the paired state. This clearly verifies that these parameters introduce flexibility in the two-body correlation function. For example, we find:

$$\phi_{\text{pair}}(\mathbf{r}) = -\frac{1}{L^3}\sum_{\mathbf{k}}\sinh\theta_{\mathbf{k}}\ \cosh\theta_{\mathbf{k}}\ e^{-i(\mathbf{k}\cdot\mathbf{r}-2\zeta_{\mathbf{k}})} \tag{D-9}$$

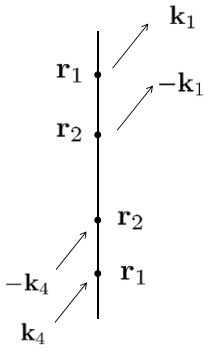


Figure 2: Diagram symbolizing the pair-pair term of the binary correlation function, with the same convention as for Figure 1.

The pair wave function thus directly depends on the phases $\zeta_{\mathbf{k}}$; as we shall see in the complements, these phases actually play a major role in the optimization of the energy. For bosons, this wave function is always even since, as we saw in § C-1-b, this is the case for the functions $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$ introduced in (C-11).

Comment:

To describe systems of interacting bosons undergoing Bose-Einstein condensation (see § 4 of Complement $\mathrm{B_{XVII}}$ and Complement $\mathrm{E_{XVII}}$, we shall add to the paired state $|\Psi_{\mathrm{paired}}\rangle$ another highly populated state with zero momentum ($\mathbf{k} = 0$). This will introduce new terms in the correlation functions, in addition to those computed in this chapter. When the population of that individual state with zero momentum is very high, these additional terms may become dominant.

D-2. Fermions in a singlet state

For fermions with spin 1/2, since each spin can point in two directions, there exists a larger number of correlation functions. Several among them will be studied in §2 of Complement $\mathrm{C_{XVII}}$. We shall only compute one of them here, involving opposite spins, as it plays the most significant role:

$$G_2\left(\mathbf{r}_1,\uparrow;\mathbf{r}_2,\downarrow\right) = \left\langle \Psi_\uparrow^\dagger\left(\mathbf{r}_1\right)\Psi_\downarrow^\dagger\left(\mathbf{r}_2\right)\Psi_\downarrow\left(\mathbf{r}_2\right)\Psi_\uparrow\left(\mathbf{r}_1\right)\right\rangle \tag{D-10}$$

Relation (D-2) now becomes:

$$G_2\left(\mathbf{r}_1,\uparrow;\mathbf{r}_2,\downarrow\right) = \frac{1}{L^6}\sum_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4} e^{i[(\mathbf{k}_4-\mathbf{k}_1)\cdot\mathbf{r}_1+(\mathbf{k}_3-\mathbf{k}_2)\cdot\mathbf{r}_2]}\left\langle a_{\mathbf{k}_1,\uparrow}^\dagger a_{\mathbf{k}_2,\downarrow}^\dagger a_{\mathbf{k}_3,\downarrow} a_{\mathbf{k}_4,\uparrow}\right\rangle \tag{D-11}$$

The diagram schematizing each term of this sum is obtained by adding spin indices to the positions in Figure 1 – as is done in Figure 4 of Complement $\mathrm{C_{XVII}}$.

The computation is then similar to that of § D-1. The direct term is written:

$$G_2^{\text{dir}}(\mathbf{r}_1,\uparrow;\mathbf{r}_2,\downarrow) = \frac{1}{L^6}\sum_{\mathbf{k}_1,\mathbf{k}_2}\langle\widehat{n}_{\mathbf{k}_1,\uparrow}\rangle\langle\widehat{n}_{\mathbf{k}_2,\downarrow}\rangle = \frac{\left\langle\widehat{N_\uparrow}\right\rangle\left\langle\widehat{N_\downarrow}\right\rangle}{L^6} \tag{D-12}$$

There is no exchange term where $\mathbf{k}_4 = \mathbf{k}_2$ and $\mathbf{k}_3 = \mathbf{k}_1$, as it would correspond to the average value of an operator changing the direction of one of the spins in two different pairs, hence destroying the equality between populations of opposite spins in each pair; this term does exist, however, in the special case where $\mathbf{k}_1 = -\mathbf{k}_2$, but its contribution is negligible. Finally, the pair annihilation-creation term corresponds to $\mathbf{k}_4 = -\mathbf{k}_3$ and $\mathbf{k}_2 = -\mathbf{k}_1$; it is written:

$$G_2^{\text{pair-pair}}(\mathbf{r}_1,\uparrow;\mathbf{r}_2,\downarrow) = \frac{1}{L^6}\sum_{\mathbf{k}_1,\mathbf{k}_4} e^{i(\mathbf{k}_4-\mathbf{k}_1)\cdot(\mathbf{r}_1-\mathbf{r}_2)}\ \left\langle a^\dagger_{\mathbf{k}_1,\uparrow}a^\dagger_{-\mathbf{k}_1,\downarrow}\right\rangle\left\langle a_{-\mathbf{k}_4,\downarrow}a_{\mathbf{k}_4,\uparrow}\right\rangle \tag{D-13}$$

Here again, the pair-pair term involves anomalous average values. As before, we can define a pair wave function ϕ_{pair} as:

$$\phi_{\text{pair}}(\mathbf{r}) = \frac{1}{L^3}\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{r}}\langle a_{\mathbf{k},\downarrow}a_{-\mathbf{k},\uparrow}\rangle = \frac{1}{L^3}\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}}\langle a_{-\mathbf{k},\downarrow}a_{\mathbf{k},\uparrow}\rangle \tag{D-14}$$

whose modulus squared appears in the correlation function:

$$G_2^{\text{pair-pair}}(\mathbf{r}_1,\uparrow;\mathbf{r}_2,\downarrow) = \left|\phi_{\text{pair}}(\mathbf{r}_1-\mathbf{r}_2)\right|^2 \tag{D-15}$$

Inserting relations (C-42) into (D-14) yields:

$$\phi_{\text{pair}}(\mathbf{r}) = \frac{1}{L^3}\sum_{\mathbf{k}} u^*_{\mathbf{k}}v_{\mathbf{k}}\ e^{i\mathbf{k}\cdot\mathbf{r}} = \frac{1}{L^3}\sum_{\mathbf{k}}\sin\theta_{\mathbf{k}}\ \cos\theta_{\mathbf{k}}\ e^{i(\mathbf{k}\cdot\mathbf{r}+2\zeta_{\mathbf{k}})} \tag{D-16}$$

The important role of this pair wave function in the BCS condensation phenomenon will be discussed in detail in Complement C_{XVII}. We will show in particular that this function not only plays a role in the diagonal binary correlation function; it also determines the long-range non-diagonal properties of the density operator, hence playing the role of an order parameter. We noted that the parameters $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$ are even functions of $\mathbf{k}$; consequently, the function $\phi_{\text{pair}}(\mathbf{r})$ is also an even function of $\mathbf{r}$.

The total correlation function is then:

$$G_2(\mathbf{r}_1,\uparrow;\mathbf{r}_2,\downarrow) = \frac{\left\langle\widehat{N_\uparrow}\right\rangle\left\langle\widehat{N_\downarrow}\right\rangle}{L^6} + \left|\phi_{\text{pair}}(\mathbf{r}_1-\mathbf{r}_2)\right|^2 \tag{D-17}$$

Inserting in this result expression (D-16) for $\phi_{\text{pair}}(\mathbf{r})$, we obtain the dependence of the correlation function on the parameters $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$. This illustrates how these parameters, which define the paired state, allow changing the correlation function.

Comment:

In the particular case where all the $\theta_{\mathbf{k}}$ are either zero or equal to $\pi/2$, we already mentioned (see end of § C-1-a) that the paired state becomes a Fock state in which the phases $\zeta_{\mathbf{k}}$ no longer play any role. It is easy to check that the anomalous average values are then all equal to zero, as is, obviously, the function $\phi_{\text{pair}}(\mathbf{r})$. On the other hand, for a different choice of the parameters $\theta_{\mathbf{k}}$, the phases $\zeta_{\mathbf{k}}$ play an especially important role, as will be shown for example in Complement C_{XVII}.

E. Paired states as a quasi-particle vacuum; Bogolubov-Valatin transformations

The Hamiltonian of a noninteracting particle system can be written as:

$$\widehat{H}_0 = \sum_i (\hbar\omega_i)\, a_i^\dagger a_i \tag{E-1}$$

where $\hbar\omega_i$ is the energy of an individual state labeled by the index i. The ground state $|\Phi_0\rangle$ of $\widehat{H}_0$ is an eigenvector of all the annihilation operators a_i, with a zero eigenvalue:

$$a_i\,|\Phi_0\rangle = 0 \tag{E-2}$$

The paired ket $|\Psi_{\text{paired}}\rangle$ is not an eigenvector of the usual annihilation operators a_i. We shall, however, introduce in § E-1 a linear transformation of the a_i and $a_i^\dagger$ into new annihilation and creation operators, and show in § E-2 that $|\Psi_{\text{paired}}\rangle$ is an eigenvector, with a zero eigenvalue, of all the new annihilation operators. The paired state will then appear as a "particle vacuum". Furthermore, in § E-3, we shall associate with $|\Psi_{\text{paired}}\rangle$ a family of operators having the same form as the Hamiltonian (E-1), but where the a_i and $a_i^\dagger$ are replaced by the new annihilation and creation operators. The interest of that association is the possibility, in certain cases (illustrated in the complements), to identify – with certain approximations if needed – an operator in this family with the Hamiltonian of a given physical situation. The problem of finding the ground state and the excited states is then solved, as if dealing with a system of independent particles. The state $|\Psi_{\text{paired}}\rangle$ can then be considered as the ground state of the Hamiltonian of independent "quasi-particles", while the new creation operators permit building a complete orthogonal basis of excited states.

E-1. Transformation of the creation and annihilation operators

For bosons in the same spin state, the state $|\varphi_{\mathbf{k}}\rangle$ belongs to the space $\mathcal{E}_{\mathbf{k}}$ associated with the pair $(\mathbf{k}, -\mathbf{k})$; this space is generated by the action of two creation operators $a_{\mathbf{k}}^\dagger$ and $a_{-\mathbf{k}}^\dagger$ on the vacuum. This is also the case for fermions in opposite spin states, if we simplify the notation $\mathbf{k},\uparrow$ to $\mathbf{k}$, as well as $-\mathbf{k},\downarrow$ to $-\mathbf{k}$ (we have labeled each pair of individual states by the value of $\mathbf{k}$ associated with the spin $\uparrow$). For both cases, we now define two new couples of creation and annihilation operators that act in $\mathcal{E}_{\mathbf{k}}$.

We introduce the two annihilation operators $b_{\mathbf{k}}$ and $b_{-\mathbf{k}}$, defined for $\mathbf{k} \neq 0$, as well as the Hermitian conjugate operators $b_{\mathbf{k}}^\dagger$ and $b_{-\mathbf{k}}^\dagger$, as:

$$\begin{aligned} b_{\mathbf{k}} &= u_{\mathbf{k}} a_{\mathbf{k}} + \eta\, v_{\mathbf{k}} a_{-\mathbf{k}}^\dagger & b_{\mathbf{k}}^\dagger &= u_{\mathbf{k}}^* a_{\mathbf{k}}^\dagger + \eta\, v_{\mathbf{k}}^* a_{-\mathbf{k}} \\ b_{-\mathbf{k}} &= u_{\mathbf{k}} a_{-\mathbf{k}} + v_{\mathbf{k}} a_{\mathbf{k}}^\dagger & b_{-\mathbf{k}}^\dagger &= u_{\mathbf{k}}^* a_{-\mathbf{k}}^\dagger + v_{\mathbf{k}}^* a_{\mathbf{k}} \end{aligned} \tag{E-3}$$

or:

$$\begin{pmatrix} b_{\mathbf{k}} \\ b_{-\mathbf{k}} \\ b_{\mathbf{k}}^\dagger \\ b_{-\mathbf{k}}^\dagger \end{pmatrix} = \begin{pmatrix} u_{\mathbf{k}} & & & \eta\, v_{\mathbf{k}} \\ & u_{\mathbf{k}} & v_{\mathbf{k}} & \\ & \eta\, v_{\mathbf{k}}^* & u_{\mathbf{k}}^* & \\ v_{\mathbf{k}}^* & & & u_{\mathbf{k}}^* \end{pmatrix} \begin{pmatrix} a_{\mathbf{k}} \\ a_{-\mathbf{k}} \\ a_{\mathbf{k}}^\dagger \\ a_{-\mathbf{k}}^\dagger \end{pmatrix} \tag{E-4}$$

As for now, $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are any two complex numbers.

As in Chapters XV and XVI, $[A, B]_{-\eta}$ denotes the commutator of A and B if $\eta = 1$ (bosons), and their anticommutator if $\eta = -1$ (fermions). We now compute $[b_{\mathbf{k}}, b_{-\mathbf{k}}]_{-\eta}$; as $a_{\mathbf{k}}$ (anti)commutes with $a_{-\mathbf{k}}$ and as $a^{\dagger}_{-\mathbf{k}}$ (anti)commutes with $a^{\dagger}_{\mathbf{k}}$, only the cross terms in $u_{\mathbf{k}} v_{\mathbf{k}}$ remain:

$$[b_{\mathbf{k}}, b_{-\mathbf{k}}]_{-\eta} = u_{\mathbf{k}} v_{\mathbf{k}} \left\{ \left[a_{\mathbf{k}}, a^{\dagger}_{\mathbf{k}} \right]_{-\eta} + \eta \left[a^{\dagger}_{-\mathbf{k}}, a_{-\mathbf{k}} \right]_{-\eta} \right\} \tag{E-5}$$

For bosons, the commutator of $a_{-\mathbf{k}}$ and $a^{\dagger}_{-\mathbf{k}}$ equals 1, and hence the commutator of $a^{\dagger}_{-\mathbf{k}}$ and $a_{-\mathbf{k}}$ equals -1; for fermions, the two anticommutators of those operators are equal to 1, so that we obtain, in both cases:

$$\begin{aligned} [b_{\mathbf{k}}, b_{-\mathbf{k}}]_{-\eta} &= u_{\mathbf{k}} v_{\mathbf{k}} \{1 - 1\} \\ &= 0 \end{aligned} \tag{E-6}$$

By Hermitian conjugation, we get:

$$\left[b^{\dagger}_{\mathbf{k}}, b^{\dagger}_{-\mathbf{k}} \right]_{-\eta} = 0 \tag{E-7}$$

We now compute $\left[b_{\mathbf{k}}, b^{\dagger}_{\mathbf{k}} \right]_{-\eta}$. This time, we get the two squared terms in $|u_{\mathbf{k}}|^2$ and $|v_{\mathbf{k}}|^2$. The one in $|u_{\mathbf{k}}|^2$ contains the (anti)commutator $\left[a_{\mathbf{k}}, a^{\dagger}_{\mathbf{k}} \right]$ equal to 1; the one in $|v_{\mathbf{k}}|^2$ contains, for bosons, the commutator of $a^{\dagger}_{-\mathbf{k}}$ and $a_{-\mathbf{k}}$ which equals -1, and for fermions, the anticommutator of those two operators which is equal to $+1$. We therefore get:

$$\left[b_{\mathbf{k}}, b^{\dagger}_{\mathbf{k}} \right]_{-\eta} = |u_{\mathbf{k}}|^2 - \eta \, |v_{\mathbf{k}}|^2 \tag{E-8}$$

In a similar way:

$$\left[b_{-\mathbf{k}}, b^{\dagger}_{-\mathbf{k}} \right]_{-\eta} = |u_{\mathbf{k}}|^2 - \eta \, |v_{\mathbf{k}}|^2 \tag{E-9}$$

Finally, we are left with the computation of $\left[b_{\mathbf{k}}, b^{\dagger}_{-\mathbf{k}} \right]_{-\eta}$ and $\left[b_{-\mathbf{k}}, b^{\dagger}_{\mathbf{k}} \right]_{-\eta}$. The first is zero since $a_{\mathbf{k}}$ (anti)commutes both with $a^{\dagger}_{-\mathbf{k}}$ and itself[8], and that $a^{\dagger}_{-\mathbf{k}}$ (anti)commutes with itself and with $a_{-\mathbf{k}}$; the reasoning is the same for the second, so that:

$$\begin{aligned} \left[b_{\mathbf{k}}, b^{\dagger}_{-\mathbf{k}} \right]_{-\eta} &= 0 \\ \left[b_{-\mathbf{k}}, b^{\dagger}_{\mathbf{k}} \right]_{-\eta} &= 0 \end{aligned} \tag{E-10}$$

To sum up, it suffices to impose, for all values of $\mathbf{k}$, the condition:

$$|u_{\mathbf{k}}|^2 - \eta \, |v_{\mathbf{k}}|^2 = 1 \tag{E-11}$$

[8] For fermions, its square is identically zero.

to get the relations:

$$\begin{aligned}\left[b_{\mathbf{k}}, b_{\mathbf{k}}^{\dagger}\right]_{-\eta} &= 1 \\ \left[b_{-\mathbf{k}}, b_{-\mathbf{k}}^{\dagger}\right]_{-\eta} &= 1\end{aligned} \tag{E-12}$$

so that the operators $b_{\mathbf{k}}$ and $b_{-\mathbf{k}}$, as well as their adjoints, obey the same relations of (anti)commutation as the usual annihilation and creation operators of identical particles.

For fermions, we find again condition (C-3), which allows us to simply set, as in (C-4):

$$\begin{aligned}u_{\mathbf{k}} &= \cos\theta_{\mathbf{k}}\ e^{-i\zeta_{\mathbf{k}}} \\ v_{\mathbf{k}} &= \sin\theta_{\mathbf{k}}\ e^{i\zeta_{\mathbf{k}}}\end{aligned} \tag{E-13}$$

We then see that the matrix in the right hand side of (E-4) is unitary. This unitary transformation of the creation and annihilation operators is called the "Bogolubov-Valatin transformation".

For bosons, we will set:

$$\begin{aligned}u_{\mathbf{k}} &= \cosh\theta_{\mathbf{k}}\ e^{-i\zeta_{\mathbf{k}}} \\ v_{\mathbf{k}} &= \sinh\theta_{\mathbf{k}}\ e^{i\zeta_{\mathbf{k}}}\end{aligned} \tag{E-14}$$

Comparison with relation (C-11) shows that:

$$x_{\mathbf{k}} = \frac{v_{\mathbf{k}}}{u_{\mathbf{k}}} \tag{E-15}$$

The transformation of the creation and annihilation operators for bosons is called the "Bogolubov transformation".

E-2. Effect on the kets $|\overline{\varphi}_{\mathbf{k}}\rangle$

We now show that the vectors $|\overline{\varphi}_{\mathbf{k}}\rangle$ are eigenkets of two annihilation operators $b_{\mathbf{k}}$ and $b_{-\mathbf{k}}$ with eigenvalues zero. This property makes them similar to a usual vacuum state, which yields zero under the action of all the annihilation operators $a_{\mathbf{k}}$.

E-2-a. Fermions in a singlet state

Let us compute the effect of those operators on the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$ defined by relation (C-1), that we write with the simplified notation already used above ($\mathbf{k}$ is associated with the spin index + and $-\mathbf{k}$ with the spin index −):

$$|\overline{\varphi}_{\mathbf{k}}\rangle = \left[u_{\mathbf{k}} + v_{\mathbf{k}}\ a_{\mathbf{k}}^{\dagger} a_{-\mathbf{k}}^{\dagger}\right]|0\rangle \tag{E-16}$$

We start with the operator $b_{\mathbf{k}}$ defined in (E-3). Its $u_{\mathbf{k}} a_{\mathbf{k}}$ term yields zero when acting on the term in $u_{\mathbf{k}}$ of $|\overline{\varphi}_{\mathbf{k}}\rangle$; only the term in $v_{\mathbf{k}}$ remains, for which the operator lowers from one to zero the occupation number of the state $\mathbf{k}, \uparrow$, since:

$$a_{\mathbf{k}} a_{\mathbf{k}}^{\dagger} a_{-\mathbf{k}}^{\dagger}|0\rangle = a_{-\mathbf{k}}^{\dagger}|0\rangle \tag{E-17}$$

As for the operator $v_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}$, it yields zero when acting on the $v_{\mathbf{k}}$ term of $|\overline{\varphi}_{\mathbf{k}}\rangle$ (for fermions, the square of a creation operator is zero), leaving only the term in $u_{\mathbf{k}}$. This leads to:

$$b_{\mathbf{k}}\,|\overline{\varphi}_{\mathbf{k}}\rangle = \left[u_{\mathbf{k}}v_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}} - v_{\mathbf{k}}u_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}\right]|0\rangle = 0 \tag{E-18}$$

The computation is the same for the operator $b_{-\mathbf{k}}$, except for the fact that the operator $a_{-\mathbf{k}}$ must first anticommute with $a^{\dagger}_{\mathbf{k}}$ before it can be regrouped with $a^{\dagger}_{-\mathbf{k}}$ and lower from one to zero the occupation number of the state $-\mathbf{k}$. The anticommutation therefore introduces a sign change, but as the definition of $b_{-\mathbf{k}}$ does not contain any $-$ sign, we again find:

$$b_{-\mathbf{k}}\,|\overline{\varphi}_{\mathbf{k}}\rangle = u_{\mathbf{k}}v_{\mathbf{k}}\left[-a^{\dagger}_{\mathbf{k}} + a^{\dagger}_{\mathbf{k}}\right]|0\rangle = 0 \tag{E-19}$$

We have shown that the two operators $b_{\mathbf{k}}$ and $b_{-\mathbf{k}}$ have the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$ as an eigenvector, with eigenvalue zero.

E-2-b. Bosons in the same spin state

Taking into account (E-15), relation (C-7) is written:

$$|\varphi_{\mathbf{k}}\rangle = \sum_{q=0}^{\infty}\frac{1}{q!}\left[-\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}}\right]^{q}\left[a^{\dagger}_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}\right]^{q}|0\rangle \tag{E-20}$$

Since:

$$a_{\mathbf{k}}\left[a^{\dagger}_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}\right]^{q}|0\rangle = \left(a_{\mathbf{k}}a^{\dagger}_{\mathbf{k}}\right)\left[a^{\dagger}_{\mathbf{k}}\right]^{q-1}\left[a^{\dagger}_{-\mathbf{k}}\right]^{q}|0\rangle = (q)\left[a^{\dagger}_{\mathbf{k}}\right]^{q-1}\left[a^{\dagger}_{-\mathbf{k}}\right]^{q}|0\rangle \tag{E-21}$$

we have:

$$u_{\mathbf{k}}a_{\mathbf{k}}\,|\varphi_{\mathbf{k}}\rangle = u_{\mathbf{k}}\sum_{q=0}^{\infty}\frac{q}{q!}\left[-\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}}\right]^{q}\left[a^{\dagger}_{\mathbf{k}}\right]^{q-1}\left[a^{\dagger}_{-\mathbf{k}}\right]^{q}|0\rangle \tag{E-22}$$

or else, since $a^{\dagger}_{-\mathbf{k}}$ commutes with all the operators in this expression:

$$\begin{aligned} u_{\mathbf{k}}a_{\mathbf{k}}\,|\varphi_{\mathbf{k}}\rangle &= -v_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}\sum_{q'=0}^{\infty}\frac{1}{q'!}\left[-\frac{v_{\mathbf{k}}}{u_{\mathbf{k}}}\right]^{q'}\left[a^{\dagger}_{\mathbf{k}}\right]^{q'}\left[a^{\dagger}_{-\mathbf{k}}\right]^{q'}|0\rangle \\ &= -v_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}\,|\varphi_{\mathbf{k}}\rangle \end{aligned} \tag{E-23}$$

where we have set $q' = q - 1$. This leads to:

$$\left(u_{\mathbf{k}}a_{\mathbf{k}} + v_{\mathbf{k}}a^{\dagger}_{-\mathbf{k}}\right)|\varphi_{\mathbf{k}}\rangle = 0 \tag{E-24}$$

which clearly shows that $|\varphi_{\mathbf{k}}\rangle$ is an eigenvector of the operator $b_{\mathbf{k}}$ defined in (E-3):

$$b_{\mathbf{k}}\,|\varphi_{\mathbf{k}}\rangle = 0 \tag{E-25}$$

The same computation leads to:

$$u_{\mathbf{k}}a_{-\mathbf{k}}\,|\varphi_{\mathbf{k}}\rangle = -v_{\mathbf{k}}a^{\dagger}_{\mathbf{k}}\,|\varphi_{\mathbf{k}}\rangle \tag{E-26}$$

and hence to:

$$b_{-\mathbf{k}} |\varphi_{\mathbf{k}}\rangle = 0 \tag{E-27}$$

As for fermions, the two operators $b_{\mathbf{k}}$ and $b_{-\mathbf{k}}$ have the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$ as an eigenvector with an eigenvalue zero.

E-3. Basis of excited states, quasi-particles

For bosons as for fermions, we just saw that the new creation and annihilation operators introduced in (E-3) and (E-4) have the same properties as the usual creation and annihilation operators. In particular, the two operators:

$$\widehat{n}\,(b_{\mathbf{k}}) = b_{\mathbf{k}}^{\dagger} b_{\mathbf{k}} \qquad\qquad \widehat{n}\,\left(b_{-\mathbf{k}}\right) = b_{-\mathbf{k}}^{\dagger} b_{-\mathbf{k}} \tag{E-28}$$

have as eigenvalues all the positive or zero integers, in perfect analogy with the operators corresponding to the population of individual states. By analogy with (E-1), it is therefore natural to introduce the operator:

$$\widehat{H}_B = \sum_{\mathbf{k}\in D} \hbar\omega_k \left[b_{\mathbf{k}}^{\dagger} b_{\mathbf{k}} + b_{-\mathbf{k}}^{\dagger} b_{-\mathbf{k}} \right] \tag{E-29}$$

where, for the moment, the ω_k are free parameters, as are the parameters which define the paired state (they will be fixed later on, depending on the physical problem we study). In relation (E-29), the summation is limited, as above, to a momentum half-space, which avoids taking opposite momenta into account twice. The eigenvalues of $\widehat{H}_B$ are all of the form:

$$E_B = \sum_{\mathbf{k}\in D} \left[n\,(b_{\mathbf{k}}) + n\,\left(b_{-\mathbf{k}}\right) \right] \hbar\omega_k \tag{E-30}$$

where $n\,(b_{\mathbf{k}})$ and $n\,(b_{-\mathbf{k}})$ are any positive or zero integers for bosons, and restricted to 0 or 1 for fermions.

The ground state $|\Phi_0(b)\rangle$ of $\widehat{H}_B$ is an eigenvector of all the annihilation operators $b_{\mathbf{k}}$ and $b_{-\mathbf{k}}$ with eigenvalues zero. Now we saw in (B-8) for bosons, and in (B-11) for fermions, that the paired state vector is a tensor product of states $|\varphi_{\mathbf{k}}\rangle$, which are precisely the eigenvectors, with zero eigenvalues, of these two operators. The paired state, $|\Psi_{\text{paired}}\rangle$ for bosons[9], or $|\Psi_{\text{BCS}}\rangle$ for fermions, is thus an eigenvector of $\widehat{H}_B$ with a zero eigenvalue (ground state).

One can then obtain the other eigenstates of $\widehat{H}_B$ (excited states) by the action of the creation operators $b_{\mathbf{k}}^{\dagger}$ and $b_{-\mathbf{k}}^{\dagger}$ on $|\Phi_0(b)\rangle$. For bosons, each of these two operators will be able to act any number of times. For fermions, on the other hand, we shall only get 3 excited states, by the action of either $b_{\mathbf{k}}^{\dagger}$, or $b_{-\mathbf{k}}^{\dagger}$, or their product; as these operators anticommute, any higher power of those operators' product will yield zero. We finally note that operator (E-29) shares many of the properties of the Hamiltonian of an ensemble of particles without mutual interactions. Just as the usual creation operators

[9]For bosons, in Complement B_{XVII}, we will associate to that paired state a coherent state $|\varphi_{\mathbf{0}}\rangle$ to obtain the state $|\Phi_B\rangle$. But, as none of the operators $b_{\mathbf{k}}$ or $b_{-\mathbf{k}}$ act in the Fock space associated with the individual state $\mathbf{k} = 0$, the conclusions will be unchanged for $|\Phi_B\rangle$.

can add particles in a system of free identical particles, the creation operators $b^\dagger_{\mathbf{k}}$ and $b^\dagger_{-\mathbf{k}}$ can be considered as the operators adding a supplementary "quasi-particle" into the physical system. These quasi-particles are not the same as particles in a system really without interactions, as illustrated by the expression of these creation operators. They yield, however, a basis of states in which we can reason as if there were no interactions, which is a very powerful framework for reasoning in many domains of physics.

For the previous considerations to be relevant from a physical point of view, we have yet to show that the Hamiltonian of the problem we study can be approximated by an operator $\widehat{H}_B$, provided we make a judicious choice of all the parameters $u_{\mathbf{k}}$, $v_{\mathbf{k}}$ and ω_k. This is not a priori easy: the Hamiltonian of an ensemble of particles includes, in general, two-body interaction terms, and those are expressed in terms of sums of products of two creation operators $a^\dagger_{\mathbf{k}}$ and two annihilation operators $a_{\mathbf{k}}$, hence of 4 operators. Now, if we insert definitions (E-3) and (E-4) into (E-29) to express $\widehat{H}_B$ as a function of the old creation and annihilation operators $a_{\mathbf{k}}$ and $a^\dagger_{\mathbf{k}}$, it is clear that we shall only obtain combinations of products of 2 operators. We shall need to make certain approximations to be able to consider $\widehat{H}_B$ as a physically pertinent approximate Hamiltonian. Examples of such situations will be given in the complements.

Conclusion

In conclusion, the paired states are a powerful tool for studying both fermions and bosons. They provide a systematic method allowing a certain flexibility in variational calculations in the presence of interactions. Furthermore, starting from a paired ground state, we were able to build a whole basis of excited states using creation and annihilation operators matching that ground state. In the complements of this chapter, we shall use the paired states to study different problems and compute the optimal parameters most relevant for each situation. The physical results will be quite different, depending on the cases, especially for fermions or for bosons; but the main point remains that the paired states offer a unified framework for obtaining all these different results.

COMPLEMENTS OF CHAPTER XVII, READER'S GUIDE

The first two complements provide more details about a number of results given in the chapter, concerning various properties of the pair operators and the paired states. The following three complements apply these concepts to physical phenomena involving fermions, and then bosons.

A_{XVII} : PAIR FIELD OPERATOR FOR IDENTICAL PARTICLES	The pair field operator is the analog, for a pair of particles, of the usual field operator for a single particle. It is a useful tool for computing average values in a paired state. The commutation relations of fermion pair operators are similar to those of bosons, except for an additional term due to the fermionic character of the pair constituents.
B_{XVII} : AVERAGE ENERGY IN A PAIRED STATE	This complement explains the computation of the average energy in a paired state. For bosons, we add to this paired state a condensate, described by a coherent state. The results of this complement are used in Complements C_{XVII} and E_{XVII} .
C_{XVII} : FERMION PAIRING, BCS THEORY	Even weak attractive interactions can greatly modify the ground state of a fermion system, via the BCS mechanism for pair formation. This complement discusses the theory of this phenomena, and its effect on the particle distribution and correlation functions, as well as its link to Bose-Einstein condensation of pairs of particles.
D_{XVII} : COOPER PAIRS	The simple Cooper model studies the bound states of two weakly attracted particles, in the presence of a Fermi sphere that prevents the particles from occupying states inside that sphere. Whereas, in general, a minimum depth of an attractive potential is required for two particles to form a bound state in 3-D, the presence of the Fermi sphere ensures the existence of a bound state, no matter how weak the attraction is. The Cooper model accounts in a somewhat intuitive way for a number of results of the BCS theory.
E_{XVII} : CONDENSED REPULSIVE BOSONS	For an ensemble of bosons, using paired states as variational states leads to the same results as the Bogolubov method based on operator transformations. We thus obtain the Bogolubov spectrum, compute the "quantum depletion" introduced by the interactions, etc.

Complement A_{XVII}

Pair field operator for identical particles

1 Pair creation and annihilation operators 1846

1-a Particles in the same spin state 1846

1-b Pairs in a singlet spin state 1849

2 Average values in a paired state 1851

2-a Average value of a field operator; pair wave function, and order parameter 1851

2-b Average value of a product of two field operators; factorization of the order parameter 1854

2-c Application to the computation of the correlation function (singlet pairs) 1858

3 Commutation relations of field operators 1861

3-a Particles in the same spin state 1861

3-b Singlet pairs 1866

In Chapter XVI we introduced a field operator $\Psi(\mathbf{r})$ acting in the state space of a system of identical particles. This operator was defined as a linear combination of annihilation operators associated with individual states having a given momentum. It proved to be a useful tool for various computations, and in particular for the determination of correlation functions. We then showed, in Chapter XVII, the relevance of paired states where, essentially, identical particles were grouped into pairs. We introduced creation and annihilation operators of pairs of particles in well defined momentum states, $A^{\dagger}_{\mathbf{K}}$ and $A_{\mathbf{K}}$. Consequently, it is natural to envisage the introduction of a field operator for pairs of particles, which will be the operator $\Phi_{\chi}(\mathbf{R})$ destroying a pair of particles whose center of mass is at point $\mathbf{R}$ and whose internal state is described by the wave function χ. Its adjoint, $\Phi^{\dagger}_{\chi}(\mathbf{R})$, creates a pair of particles in that same state. In this complement, we will define these operators and study some of their properties.

We start in § 1 by giving the expression of these field operators $\Phi_{\chi_n}(\mathbf{R})$ and $\Phi^{\dagger}_{\chi_n}(\mathbf{R})$ for pairs described by any orbital state χ_n. We consider the case where the particles are either in the same spin state, or in a spin singlet state. We then study, in § 2, the average values, in paired states, of pair field operators and of products of such operators. These average values have some very interesting properties leading us, in particular, to introduce a new wave function $\phi_{\text{pair}}(\mathbf{r})$, called the "pair wave function", which is not simply the two-particle wave function $\chi_{\text{pair}}(\mathbf{r})$ used to build the paired state. As we shall see in § 2-c, this new wave function explicitly appears in the binary correlation function of the particles' positions. Moreover, its norm is linked to the number of quanta present in the field of condensed pairs. The origin of this pair function is the fact that pairs can collectively contribute to the creation of a field whose average value is what we shall call an "order parameter". This non-zero order parameter indicates the existence of a macroscopic field associated with the pairs. We will show how it relates the "anomalous average values" (of operators that do not conserve particle number) to

the normal average values of a product of two field operators $\Phi_{\chi_n}(\mathbf{R})$ and $\Phi^\dagger_{\chi_n}(\mathbf{R})$, that does conserve particle number. In particular, we shall use, in § 2-c, the properties of the pair field operator to get the correlation functions in a paired BCS state, and to study the consequences of the existence of the macroscopic field associated with the pairs. We shall finally study, in § 3, the commutation properties of these operators; they will be found to be similar to those of bosons (whether the particles building the pair are bosons or fermions), but not completely identical as corrective terms must be added to the boson commutator. We shall see that, since the pairs are strongly bound and have a spatial extension much smaller than all the characteristic dimensions of the problem, the pairs can be assimilated to bosons; if, however, the pairs are weakly bound (as is the case, in particular, for the BCS mechanism we will discuss in Complement C_{XVII}, it is not possible to consider them as indivisible entities: the fermionic structure of their components plays an important role that cannot be ignored.

1. Pair creation and annihilation operators

By analogy with the field operator for the particles composing the pairs, we now introduce a field operator concerning the pairs themselves. The adjoint of this field operator allows the direct creation of a pair of particles at a given point, and with a given internal state; as for the operator itself, it annihilates that same pair.

1-a. Particles in the same spin state

We defined, in Chapter XVII, the operators $A^\dagger_{\mathbf{K}}$ and $A_{\mathbf{K}}$ for pairs of particles without spin (or in the same spin state), as:

$$A^\dagger_{\mathbf{K}} = \frac{1}{\sqrt{2}} \sum_{\mathbf{k}} g_{\mathbf{k}}\ a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}}$$
$$A_{\mathbf{K}} = \frac{1}{\sqrt{2}} \sum_{\mathbf{k}} g^*_{\mathbf{k}}\ a_{\frac{\mathbf{K}}{2}-\mathbf{k}} a_{\frac{\mathbf{K}}{2}+\mathbf{k}} \tag{1}$$

In this expression, $g_{\mathbf{k}}$ is the Fourier transform of the wave function $\chi(\mathbf{r})$ characterizing the pair:

$$g_{\mathbf{k}} = \frac{1}{L^{3/2}} \int_{L^3} d^3r\ e^{-i\mathbf{k}\cdot\mathbf{r}} \chi(\mathbf{r}) \tag{2}$$

and L is the edge length of a cube, of volume L^3 , which contains the physical system. We now generalize these definitions to the case where the pair is not necessarily in a given orbital state, but in any state belonging to an orthonormal basis of states $|\chi_n\rangle$, with the index n going from 1 to infinity; these states each have a wave function $\chi_n(\mathbf{r})$ whose Fourier transform is $g^n_{\mathbf{k}}$. We therefore simply add an index n to the previous definitions, as for example:

$$A^\dagger_{\mathbf{K},n} = \frac{1}{\sqrt{2}} \sum_{\mathbf{k}} g^n_{\mathbf{k}}\ a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k}} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k}} \tag{3}$$

We saw in Chapter XVII that the bosonic or fermionic character of the particles building the pair requires the functions $g^n_{\mathbf{k}}$ to have the parity η with respect to the

variable $\mathbf{k}$, which means $g^n_{-\mathbf{k}} = \eta\, g^n_{\mathbf{k}}$ with:

$$\eta = \begin{cases} +1 & \text{for bosons} \\ -1 & \text{for fermions} \end{cases} \tag{4}$$

Should $g^n_{\mathbf{k}}$ be of parity $-\eta$, it is easy to check by (anti)commutation of the operators that expression (3) yields zero. We can then only consider the case where the $g^n_{\mathbf{k}}$, and hence the corresponding wave functions $\chi_n(\mathbf{r})$, have the parity η. If, however, we need the basis of states $|\chi_n\rangle$ associated with these wave functions to be complete, we can include states of any parity, ± 1. We must then remember that the operators $A^\dagger_{\mathbf{K},n}$ are zero whenever the index n corresponds to a wave function of parity $-\eta$.

α. *Expression of $A^\dagger_{\mathbf{K},n}$ in terms of the particle field operator*

Relation (A-10) in Chapter XVI permits replacing the creation operators $a^\dagger_{\mathbf{k}}$ by:

$$a^\dagger_{\mathbf{k}} = \frac{1}{\sqrt{L^3}} \int \mathrm{d}^3 r\; e^{i\mathbf{k}\cdot\mathbf{r}}\; \Psi^\dagger(\mathbf{r}) \tag{5}$$

where $\Psi^\dagger(\mathbf{r})$ is the adjoint of the field operator associated with the elementary components of the pair (the "atoms" of each "molecule"). Using twice this relation in (3), we get:

$$A^\dagger_{\mathbf{K},n} = \frac{1}{\sqrt{2}L^3} \sum_{\mathbf{k}}\; g^n_{\mathbf{k}}\; \int \mathrm{d}^3 r \int \mathrm{d}^3 r'\; e^{i\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right)\cdot\mathbf{r}} e^{i\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right)\cdot\mathbf{r}'}\; \Psi^\dagger(\mathbf{r})\Psi^\dagger(\mathbf{r}') \tag{6}$$

or else, choosing as the integration variables $\mathbf{R} = (\mathbf{r}+\mathbf{r}')/2$ and $\mathbf{x} = \mathbf{r}-\mathbf{r}'$:

$$A^\dagger_{\mathbf{K},n} = \frac{1}{\sqrt{2}L^3} \int \mathrm{d}^3 R\; e^{i\mathbf{K}\cdot\mathbf{R}} \int \mathrm{d}^3 x \sum_{\mathbf{k}}\; g^n_{\mathbf{k}}\; e^{i\mathbf{k}\cdot\mathbf{x}}\; \Psi^\dagger(\mathbf{R}+\frac{\mathbf{x}}{2})\Psi^\dagger(\mathbf{R}-\frac{\mathbf{x}}{2}) \tag{7}$$

The summation over $\mathbf{k}$ on the right-hand side leads to expression (A-2) of Chapter XVII for the wave function χ_n, and we can write:

$$A^\dagger_{\mathbf{K},n} = \frac{1}{\sqrt{2L^3}} \int \mathrm{d}^3 R\; e^{i\mathbf{K}\cdot\mathbf{R}} \int \mathrm{d}^3 x\; \chi_n(\mathbf{x})\; \Psi^\dagger(\mathbf{R}+\frac{\mathbf{x}}{2})\Psi^\dagger(\mathbf{R}-\frac{\mathbf{x}}{2}) \tag{8}$$

This other form for the operator already introduced in (3) demonstrates the fact that it creates a pair of particles in a molecular state characterized, for its external variables, by a plane wave of wave vector $\mathbf{K}$, and for its internal variables, by the wave function χ_n .

β. *Pair field*

For each internal state $|\chi_n\rangle$ of the pair, we can introduce, using relation (A-3) of Chapter XVI, an operator $\Phi^\dagger_{\chi_n}(\mathbf{R})$ that creates a pair at point $\mathbf{R}$ and in the internal state χ_n:

$$\Phi^\dagger_{\chi_n}(\mathbf{R}) = \frac{1}{\sqrt{L^3}} \sum_{\mathbf{K}}\; e^{-i\mathbf{K}\cdot\mathbf{R}}\; A^\dagger_{\mathbf{K},n} \tag{9}$$

Replacing in (8) the integral variable $\mathbf{R}$ by $\mathbf{R}'$, and using the result in equality (9), we get:

$$\Phi^{\dagger}_{\chi_n}(\mathbf{R}) = \frac{1}{\sqrt{2}L^3}\sum_{\mathbf{K}}\int \mathrm{d}^3R'\ e^{i\mathbf{K}\cdot(\mathbf{R}'-\mathbf{R})}\int \mathrm{d}^3x\ \chi_n(\mathbf{x})\ \Psi^{\dagger}(\mathbf{R}'+\frac{\mathbf{x}}{2})\Psi^{\dagger}(\mathbf{R}'-\frac{\mathbf{x}}{2}) \quad (10)$$

The sum over $\mathbf{K}$ of $e^{i\mathbf{K}\cdot(\mathbf{R}'-\mathbf{R})}$ then yields $L^3\delta(\mathbf{R}-\mathbf{R}')$, which allows integrating over d^3R', and we obtain:

$$\Phi^{\dagger}_{\chi_n}(\mathbf{R}) = \frac{1}{\sqrt{2}}\int \mathrm{d}^3x\ \chi_n(\mathbf{x})\ \Psi^{\dagger}(\mathbf{R}+\frac{\mathbf{x}}{2})\Psi^{\dagger}(\mathbf{R}-\frac{\mathbf{x}}{2}) \quad (11)$$

This operator is therefore a product of field operators creating successively each of the two elements of the pair, which is easy to understand from a physical point of view. Note, however, that the two elements are not created at the same point, but symmetrically with respect to point $\mathbf{R}$, and with a spatial distribution whose amplitude is given by the wave function $\chi(\mathbf{x})$ of the "molecule". The spatial zone involved in the process thus extends over a distance of the order of the range of this wave function.

As for the pair field operator itself, which annihilates a pair, it is defined by Hermitian conjugation of the previous relation:

$$\Phi_{\chi_n}(\mathbf{R}) = \frac{1}{\sqrt{2}}\int \mathrm{d}^3x\ \chi^*_n(\mathbf{x})\ \Psi(\mathbf{R}-\frac{\mathbf{x}}{2})\Psi(\mathbf{R}+\frac{\mathbf{x}}{2}) \quad (12)$$

We now use relation (A-3) of Chapter XVI to come back to the annihilation operators in a basis of individual states with fixed momenta. Using (twice) this relation in (12), we get:

$$\Phi_{\chi_n}(\mathbf{R}) = \frac{1}{L^3\sqrt{2}}\int \mathrm{d}^3x\ \chi^*_n(\mathbf{x})\ \sum_{\mathbf{k}_1,\mathbf{k}_2} e^{i\mathbf{k}_1\cdot\left(\mathbf{R}-\frac{\mathbf{x}}{2}\right)}e^{i\mathbf{k}_2\cdot\left(\mathbf{R}+\frac{\mathbf{x}}{2}\right)}a_{\mathbf{k}_1}a_{\mathbf{k}_2} \quad (13)$$

This relation will be useful in what follows.

γ. *Inversion; expression for the interaction energy*

We call $|\chi_n\rangle$ the individual states corresponding to the wave functions $\chi_n(\mathbf{r})$ and assume they form a complete basis. The closure relation on these states is written:

$$\sum_n \langle\mathbf{k}\,|\chi_n\rangle\,\langle\chi_n\,|\mathbf{k}'\rangle = \sum_n (g^n_{\mathbf{k}})\,(g^n_{\mathbf{k}'})^* = \delta_{\mathbf{k},\mathbf{k}'} \quad (14)$$

We now multiply relation (3) by $(g^n_{\mathbf{k}'})^*$, and sum over n to get:

$$\sum_n (g^n_{\mathbf{k}'})^*\,A^{\dagger}_{\mathbf{K},n} = \frac{1}{\sqrt{2}}\ a^{\dagger}_{\frac{\mathbf{K}}{2}+\mathbf{k}'}a^{\dagger}_{\frac{\mathbf{K}}{2}-\mathbf{k}'} \quad (15)$$

It is thus possible to invert relations (3) and express any two creation operators as a sum of pair creation operators, according to:

$$a^{\dagger}_{\mathbf{k}_1}a^{\dagger}_{\mathbf{k}_2} = \sqrt{2}\sum_n \left(g^n_{(\mathbf{k}_1-\mathbf{k}_2)/2}\right)^* A^{\dagger}_{\mathbf{k}_1+\mathbf{k}_2,n} \quad (16)$$

where we have replaced $\mathbf{K}$ par $\mathbf{k}_1+\mathbf{k}_2$ and $\mathbf{k}'$ by $(\mathbf{k}_1-\mathbf{k}_2)/2$. By Hermitian conjugation, we get a similar relation for any product $a_{\mathbf{k}_3}a_{\mathbf{k}_4}$ of annihilation operators.

Interaction Hamiltonian:

Any operator $\widehat{W}_{\rm int}$ for the binary interactions between particles can therefore be written as:

$$\widehat{W}_{\rm int} = \sum_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4,n,n'} \langle 1:\mathbf{k}_1;2:\mathbf{k}_2|\, W_2\,(1,2)\,|1:\mathbf{k}_3;2:\mathbf{k}_4\rangle \left(g^n_{(\mathbf{k}_1-\mathbf{k}_2)/2}\right)^* \left(g^{n'}_{(\mathbf{k}_3-\mathbf{k}_4)/2}\right) A^\dagger_{\mathbf{K},n}\; A_{\mathbf{K},n'} \tag{17}$$

where $W_2\,(1,2)$ is the binary interaction between particles (as, for example, in Complement $\mathrm{E_{XV}}$; using momentum conservation, we have set:

$$\mathbf{K} = \mathbf{k}_1+\mathbf{k}_2 = \mathbf{k}_3+\mathbf{k}_4 \tag{18}$$

Written in terms of pair creation and annihilation operators, $\widehat{W}_{\rm int}$ is the sum of quadratic terms, and no longer of fourth degree terms as was the case with operators for individual particles. Note, however, that one must be careful when using relation (17) since, as we shall see in § 3, the pair creation and annihilation operators do not obey the usual commutation relations. The action of an operator $A_{\mathbf{K},n}$ on a paired state obtained by the action of $\left[A^\dagger_{\mathbf{K}',n}\right]^N$ on the vacuum, does not necessarily yield zero when $\mathbf{K}\neq\mathbf{K}'$. Pair creation operators are not as simple to handle as particle creation operators.

1-b. Pairs in a singlet spin state

For a pair of spin 1/2 particles in a singlet spin state, we use relation (A-23) of Chapter XVII and add an index n to represent the internal orbital state of the pair; this reads:

$$A^\dagger_{\mathbf{K},n} = \sum_{\mathbf{k}} g^n_{\mathbf{k}}\; a^\dagger_{\frac{\mathbf{K}}{2}+\mathbf{k},\uparrow} a^\dagger_{\frac{\mathbf{K}}{2}-\mathbf{k},\downarrow} \tag{19}$$

The following computations apply directly to fermions in a singlet state, for which the functions $g^n_{\mathbf{k}}$ must be even with respect to the variable $\mathbf{k}$. We noted however in Chapter XVII, in comment (ii) just before § B, that they can also apply to fermions in a triplet spin state, when the function $g^n_{\mathbf{k}}$ is odd; even though this case can be included in the following discussion, for the sake of simplicity we will continue to talk about singlet pairs.

α. Expression of $A^\dagger_{\mathbf{K},n}$ in terms of the particle field operator

Relation (A-9) of Chapter XVI becomes here, taking into account the spin indices:

$$a^\dagger_{\mathbf{k},\nu} = \frac{1}{\sqrt{L^3}} \int \mathrm{d}^3r\; e^{i\mathbf{k}\cdot\mathbf{r}}\; \Psi^\dagger_\nu(\mathbf{r}) \tag{20}$$

Inserting this equality in (19) yields:

$$A^\dagger_{\mathbf{K},n} = \frac{1}{L^3}\sum_{\mathbf{k}}\; g^n_{\mathbf{k}}\; \int \mathrm{d}^3r \int \mathrm{d}^3r'\; e^{i\left(\frac{\mathbf{K}}{2}+\mathbf{k}\right)\cdot\mathbf{r}} e^{i\left(\frac{\mathbf{K}}{2}-\mathbf{k}\right)\cdot\mathbf{r}'}\; \Psi^\dagger_\uparrow(\mathbf{r})\Psi^\dagger_\downarrow(\mathbf{r}') \tag{21}$$

As previously, the wave function $\chi_n(\mathbf{x})$ appears when we use as integral variables $\mathbf{R} = (\mathbf{r} + \mathbf{r}')/2$ and $\mathbf{x} = \mathbf{r} - \mathbf{r}'$, and we get:

$$A^\dagger_{\mathbf{K},n} = \frac{1}{\sqrt{L^3}} \int \mathrm{d}^3R\, e^{i\mathbf{K}\cdot\mathbf{R}} \int \mathrm{d}^3x\, \chi_n(\mathbf{x})\ \Psi^\dagger_\uparrow(\mathbf{R}+\frac{\mathbf{x}}{2})\Psi^\dagger_\downarrow(\mathbf{R}-\frac{\mathbf{x}}{2}) \tag{22}$$

This yields the form of the operator creating a pair of particles in a singlet molecular state, characterized by a plane wave of wave vector $\mathbf{K}$ for its external variables, and by the wave function χ_n for its internal variables.

β. Pair field

We now insert relation (22) in (9); we get:

$$\Phi^\dagger_{\chi_n}(\mathbf{R}) = \frac{1}{L^3} \sum_{\mathbf{K}} \int \mathrm{d}^3R'\, e^{i\mathbf{K}\cdot(\mathbf{R}'-\mathbf{R})} \int \mathrm{d}^3x\, \chi_n(\mathbf{x})\ \Psi^\dagger_\uparrow(\mathbf{R}'+\frac{\mathbf{x}}{2})\Psi^\dagger_\downarrow(\mathbf{R}'-\frac{\mathbf{x}}{2}) \tag{23}$$

As before, the sum over $\mathbf{K}$ of $e^{i\mathbf{K}\cdot(\mathbf{R}'-\mathbf{R})}$ yields $L^3\delta(\mathbf{R}-\mathbf{R}')$, and we get:

$$\Phi^\dagger_{\chi_n}(\mathbf{R}) = \int \mathrm{d}^3x\, \chi_n(\mathbf{x})\ \Psi^\dagger_\uparrow(\mathbf{R}+\frac{\mathbf{x}}{2})\Psi^\dagger_\downarrow(\mathbf{R}-\frac{\mathbf{x}}{2}) \tag{24}$$

The same comments as in § 1-a-β can be made: this operator successively creates the two elements of the pair at different points, with a probability amplitude given by the internal wave function $\chi_n(\mathbf{x})$ of the distance between these points. The field operator is obtained by Hermitian conjugation:

$$\Phi_{\chi_n}(\mathbf{R}) = \int \mathrm{d}^3x\, \chi^*_n(\mathbf{x})\ \Psi_\downarrow(\mathbf{R}-\frac{\mathbf{x}}{2})\Psi_\uparrow(\mathbf{R}+\frac{\mathbf{x}}{2}) \tag{25}$$

It will often be convenient to come back and use the annihilation operators in a basis of individual states of fixed momenta. Using (twice) relation (A-14) of Chapter XVI, we get:

$$\Phi_{\chi_n}(\mathbf{R}) = \frac{1}{L^3} \int \mathrm{d}^3x\, \chi^*_n(\mathbf{x}) \sum_{\mathbf{k}_1,\mathbf{k}_2} e^{i\mathbf{k}_1\cdot(\mathbf{R}-\frac{\mathbf{x}}{2})} e^{i\mathbf{k}_2\cdot(\mathbf{R}+\frac{\mathbf{x}}{2})} a_{\mathbf{k}_1,\downarrow} a_{\mathbf{k}_2,\uparrow} \tag{26}$$

Comment:

For singlet pairs, we could invert those relations, as we did before, and express the interaction energy in terms of the pair creation and annihilation operators. It is, however, a bit more complicated in this case than when the pairs were in the same spin state: as we shall see in § 2-c-α, it would be necessary to involve another pair creation operator (in a triplet state). This would lead to cumbersome notation, and the computation will not be presented here.

2. Average values in a paired state

We now compute the average values of pair field operators, or of products of such operators, in the paired state we defined Chapter XVII. We shall use relations (13) or (26), depending on whether the particles are in the same spin state, or in a singlet spin state. In both cases, the computation of the average value of those operators in a paired state involves the computation of average values of products of annihilation operators – i.e. of "anomalous" average values as defined in Chapter XVII.

2-a. Average value of a field operator; pair wave function, and order parameter

Expressions for the paired kets were obtained in § B-2 of Chapter XVII as tensor products of states of pairs that are not eigenstates of the occupation number operators. These pairs all have a zero total momentum; we therefore assume, from now on, that $\mathbf{K} = 0$. The average value computation of a pair field operator in these states will lead to a new wave function, that we will call the "pair wave function".

α. Particles in the same spin state

Relations (B-8) and (B-9) of Chapter XVII give the expression of the paired state vector $|\Psi_{\text{paired}}\rangle$ for an ensemble of a large number of particles:

$$|\Psi_{\text{paired}}\rangle = \prod_{\otimes \mathbf{k}\in\mathbf{D}} \exp\left\{\sqrt{2}\; g_{\mathbf{k}}\; a^{\dagger}_{\mathbf{k}} a^{\dagger}_{-\mathbf{k}}\right\} |0\rangle \tag{27}$$

The function $g_{\mathbf{k}}$ used to build this paired state is a priori totally independent of the functions $g^n_{\mathbf{k}}$ defining the pair field operators. In such paired states, the populations of the states of the same pair are always equal; consequently, the only non-zero average values $\langle a_{\mathbf{k}_1} a_{\mathbf{k}_2}\rangle$ are those in which the two annihilation operators act on the two states of the same pair, which have opposite momenta. As the total momenta of each pair is zero, we can set $\mathbf{k}_1 = -\mathbf{k}_2$ in (13) and obtain:

$$\begin{aligned}\left\langle \Phi_{\chi_n}(\mathbf{R})\right\rangle &= \frac{1}{L^3\sqrt{2}} \int \mathrm{d}^3x\; \chi^*_n(\mathbf{x}) \sum_{\mathbf{k}_1} e^{-i\mathbf{k}_1\cdot\mathbf{x}} \langle a_{\mathbf{k}_1} a_{-\mathbf{k}_1}\rangle \\ &= \int \mathrm{d}^3x\; \chi^*_n(\mathbf{x})\; \phi_{\text{pair}}(\mathbf{x})\end{aligned} \tag{28}$$

where the (non normalized) "pair wave function" has already been defined in (D-6) of Chapter XVII:

$$\boxed{\phi_{\text{pair}}(\mathbf{x}) = \langle \mathbf{x}\,|\phi_{\text{pair}}\rangle = \frac{1}{L^3\sqrt{2}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{x}} \langle a_{\mathbf{k}} a_{-\mathbf{k}}\rangle} \tag{29}$$

Changing the sign of the summation variable $\mathbf{k}$, allows writing the pair wave function $\bar{\phi}_{\text{pair}}(\mathbf{k})$ in the momentum representation as:

$$\bar{\phi}_{\text{pair}}(\mathbf{k}) = \langle \mathbf{k}\,|\phi_{\text{pair}}\rangle = \frac{1}{L^{3/2}\sqrt{2}} \langle a_{-\mathbf{k}} a_{\mathbf{k}}\rangle \tag{30}$$

Note that because of the condition $\mathbf{k}_1 = -\mathbf{k}_2$ (the total momentum of each pair is zero), the average value $\left\langle \Phi_{\chi_n}(\mathbf{R}) \right\rangle$ no longer depends on $\mathbf{R}$.

The average value of the pair field operator is thus:

$$\left\langle \Phi_{\chi_n}(\mathbf{R}) \right\rangle = \left\langle \Phi_{\chi_n} \right\rangle = \langle \chi_n \, | \phi_{\text{pair}} \rangle \tag{31}$$

As expected from the translation invariance of the system, it is independent of $\mathbf{R}$. On the other hand, it depends on the internal state $|\chi_n\rangle$, and reaches a maximum when $|\chi_n\rangle$ is equal to the normalized state $\left|\phi_{\text{pair}}^{\text{norm}}\right\rangle$ proportional to $|\phi_{\text{pair}}\rangle$:

$$\left|\phi_{\text{pair}}^{\text{norm}}\right\rangle = \frac{|\phi_{\text{pair}}\rangle}{\sqrt{\langle \phi_{\text{pair}} \, | \phi_{\text{pair}} \rangle}} \tag{32}$$

This computation therefore leads to a new state $\left|\phi_{\text{pair}}^{\text{norm}}\right\rangle$, different from the state $|\chi\rangle$ that was used in Chapter XVII to build the paired state $|\Psi_{\text{paired}}\rangle$. Choosing for the first vector of the basis $|\chi_1\rangle = \left|\phi_{\text{pair}}^{\text{norm}}\right\rangle$, the average value of the field is given by:

$$\boxed{\left\langle \Phi_{\chi_1}(\mathbf{R}) \right\rangle = \left\langle \Phi_{\chi_1} \right\rangle = \sqrt{\langle \phi_{\text{pair}} \, | \phi_{\text{pair}} \rangle}} \tag{33}$$

This average value $\left\langle \Phi_{\chi_1}(\mathbf{R}) \right\rangle$ is often called the "order parameter of the pairs"; its non-zero value is important as it indicates the existence of a field constructed collectively by the pairs. In the present case, the average value of this field is independent of $\mathbf{R}$, as the paired state was built from pairs having a total momentum $\mathbf{K} = 0$ and whose center of mass has a constant wave function.

The average values $\langle a_{\mathbf{k}} a_{-\mathbf{k}} \rangle$, which according to (29) determine the pair wave function, have been called, in § C-3 of Chapter XVII, "anomalous average values", as they involve operators that do not conserve particle number. For bosons in the same spin state, relation (C-52) of that chapter indicates that:

$$\langle a_{\mathbf{k}} a_{-\mathbf{k}} \rangle = -e^{2i\zeta_{\mathbf{k}}} \, \sinh\theta_{\mathbf{k}} \, \cosh\theta_{\mathbf{k}} \tag{34}$$

One may wonder, of course, what the purpose of computing an anomalous average value is, as it can only be zero in a state with a fixed total particle number. We shall see, however, in § 2-b that these anomalous average values are a useful tool for computing average values of operators that do conserve the total number of particles and hence have a direct physical interpretation.

For bosons, operators $a_{\mathbf{k}}$ and $a_{-\mathbf{k}}$ commute, and hence the definition (29) shows that the wave function $\phi_{\text{pair}}(\mathbf{x})$ is even:

$$\phi_{\text{pair}}(-\mathbf{x}) = \phi_{\text{pair}}(\mathbf{x}) \tag{35}$$

The field mean value (31) is thus zero for any state χ_n of the basis whose wave function is odd: the postulate of symmetrization with respect to the pair components requires that pair to be in an even orbital state[1].

[1] If the particles composing the pair were fermions in the same spin state, the conclusions would be opposite. The wave function would be odd (because of the anticommutation of the operators $a_{\mathbf{k}}$ and $a_{-\mathbf{k}}$); the average values for even internal states χ_n would be zero.

β. Singlet pairs

For fermions in a singlet state, the paired state vector $|\Psi_{\rm BCS}\rangle$ is given by relations (B-11) and (B-12) of Chapter XVII:

$$|\Psi_{\rm BCS}\rangle = \prod_{\otimes \mathbf{k}} \exp\left[g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle \tag{36}$$

Following (26), we must add the spin index $\downarrow$ to the state $\mathbf{k}$, and the spin index $\uparrow$ to the state $-\mathbf{k}$. Here again, the average value of the product of annihilation operators is different from zero only if their wave vectors are opposite, and relation (26) then leads to:

$$\begin{aligned} \left\langle \Phi_{\chi_n}(\mathbf{R})\right\rangle &= \frac{1}{L^3}\int \mathrm{d}^3x\ \chi_n^*(\mathbf{x})\ \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{x}} \langle a_{\mathbf{k},\downarrow} a_{-\mathbf{k},\uparrow}\rangle \\ &= \langle \chi_n\ |\phi_{\rm pair}\rangle \end{aligned} \tag{37}$$

with the definition (D-14) of Chapter XVII of the state $|\phi_{\rm pair}\rangle$, associated with the (non-normalized) wave function:

$$\phi_{\rm pair}(\mathbf{x}) = \langle \mathbf{x}\ |\phi_{\rm pair}\rangle = \frac{1}{L^3}\sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{x}} \langle a_{\mathbf{k},\downarrow} a_{-\mathbf{k},\uparrow}\rangle \tag{38}$$

In a similar way, the pair wave function in the momentum representation $\bar{\phi}_{\rm pair}(\mathbf{k})$ can be defined as:

$$\bar{\phi}_{\rm pair}(\mathbf{k}) = \frac{1}{L^{3/2}} \langle a_{-\mathbf{k},\downarrow} a_{\mathbf{k},\uparrow}\rangle \tag{39}$$

This wave function can be interpreted in the same way as the wave function defining the orbital variables of a pair in the singlet state. As in (32), we define the normalized ket $\left|\phi^{\rm norm}_{\rm pair}\right\rangle$. The field average value (37) is zero if $|\chi_n\rangle$ is orthogonal to $\left|\phi^{\rm norm}_{\rm pair}\right\rangle$ and reaches a maximum for $|\chi_n\rangle = |\chi_1\rangle = \left|\phi^{\rm norm}_{\rm pair}\right\rangle$; this maximum is equal to:

$$\boxed{\left\langle \Phi_{\chi_1}(\mathbf{R})\right\rangle = \left\langle \Phi_{\chi_1}\right\rangle = \sqrt{\langle \phi_{\rm pair}\ |\phi_{\rm pair}\rangle}} \tag{40}$$

and defines the order parameter of the physical system. It indicates the presence of a field created collectively by the pairs. As noted before, since the total momentum of each pair is zero, this average value does not depend on $\mathbf{R}$.

The average values that come into play in that definition are given by relation (C-42) of Chapter XVII:

$$\langle a_{\mathbf{k},\downarrow} a_{-\mathbf{k},\uparrow}\rangle = u^*_{\mathbf{k}} v_{\mathbf{k}} = \sin\theta_{\mathbf{k}} \cos\theta_{\mathbf{k}}\ e^{2i\zeta_{\mathbf{k}}} \tag{41}$$

(in the BCS state, the $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$ are even functions of $\mathbf{k}$). We noted, at the end of § C-1-a of that chapter that, in the specific case where the $\theta_{\mathbf{k}}$ are either zero or equal to $\pi/2$, the paired ket is simply a Fock state of individual particles, hence a ket without pairing. Since (41) is then equal to zero, we see that the pair wave function is zero in the absence of pairing.

2-b. Average value of a product of two field operators; factorization of the order parameter

The field operators do not conserve particle number, as opposed to the usual operators such as the Hamiltonian, the total momentum, the double density, etc. On the other hand, the product of operators $\Phi^{\dagger}_{\chi_n}(\mathbf{R})\,\Phi_{\chi_{n'}}(\mathbf{R}')$ does conserve that number, and may help characterizing the properties of the pairs while being easier to interpret from a physical point of view.

α. Particles in the same spin state

Using relation (13) we get:

$$\left\langle \Phi^{\dagger}_{\chi_n}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle = \frac{1}{2L^6}\int \mathrm{d}^3x\ \chi_n(\mathbf{x})\int \mathrm{d}^3x'\ \chi^*_{n'}(\mathbf{x}') \times \sum_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4} e^{-i\mathbf{k}_1\cdot\left(\mathbf{R}+\frac{\mathbf{x}}{2}\right)} e^{-i\mathbf{k}_2\cdot\left(\mathbf{R}-\frac{\mathbf{x}}{2}\right)} e^{i\mathbf{k}_3\cdot\left(\mathbf{R}'-\frac{\mathbf{x}'}{2}\right)} e^{i\mathbf{k}_4\cdot\left(\mathbf{R}'+\frac{\mathbf{x}'}{2}\right)} \left\langle a^{\dagger}_{\mathbf{k}_1} a^{\dagger}_{\mathbf{k}_2} a_{\mathbf{k}_3} a_{\mathbf{k}_4}\right\rangle \tag{42}$$

The integrals over d^3x and d^3x' yield Fourier transforms $g^n_{\mathbf{k}}$ of the wave functions $\chi_n(\mathbf{x})$:

$$g^n_{\mathbf{k}} = \frac{1}{L^{3/2}}\int \mathrm{d}^3x\ e^{-i\mathbf{k}\cdot\mathbf{x}}\ \chi_n(\mathbf{x}) \tag{43}$$

and we get:

$$\left\langle \Phi^{\dagger}_{\chi_n}(\mathbf{R})\,\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle = \frac{1}{2L^3}\sum_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4} g_n\left(\frac{\mathbf{k}_1-\mathbf{k}_2}{2}\right)\ g^*_{n'}\left(\frac{\mathbf{k}_4-\mathbf{k}_3}{2}\right) \times\ e^{i\left[(\mathbf{k}_3+\mathbf{k}_4)\mathbf{R}'-(\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{R}\right]\cdot} \left\langle a^{\dagger}_{\mathbf{k}_1} a^{\dagger}_{\mathbf{k}_2} a_{\mathbf{k}_3} a_{\mathbf{k}_4}\right\rangle \tag{44}$$

where, to simplify the notation, we have written $g_n(\mathbf{k})$ the Fourier transform of $g^n_{\mathbf{k}}$.

Computation of the average value $\left\langle a^{\dagger}_{\mathbf{k}_1} a^{\dagger}_{\mathbf{k}_2} a_{\mathbf{k}_3} a_{\mathbf{k}_4}\right\rangle$

This computation follows the same steps as the one in § D of Chapter XVII for the correlation between particles, as well as the one in § 3-a-β of Complement B_{XVII} for the interaction energy. Three cases must be distinguished:

– (**I**) The "forward scattering" terms are obtained either for $\mathbf{k}_4=\mathbf{k}_1$ and $\mathbf{k}_3=\mathbf{k}_2$ (direct terms), or $\mathbf{k}_3=\mathbf{k}_1$ and $\mathbf{k}_4=\mathbf{k}_2$ (exchange terms). We assume these forward scattering terms concern two different pairs, meaning $\mathbf{k}_1\neq\pm\mathbf{k}_2$. Since:

$$\left\langle a^{\dagger}_{\mathbf{k}_1} a^{\dagger}_{\mathbf{k}_2} a_{\mathbf{k}_2} a_{\mathbf{k}_1}\right\rangle = \eta\left\langle a^{\dagger}_{\mathbf{k}_1} a^{\dagger}_{\mathbf{k}_2} a_{\mathbf{k}_1} a_{\mathbf{k}_2}\right\rangle = \langle\widehat{n}_{\mathbf{k}_1}\rangle\,\langle\widehat{n}_{\mathbf{k}_2}\rangle \tag{45}$$

their sum yields the contribution:

$$\left\langle \Phi^{\dagger}_{\chi_n}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle_{\text{forward}} = \frac{1}{2L^3}\sum_{\mathbf{k}_1,\mathbf{k}_2} g_n(\mathbf{k})\left[g^*_{n'}(\mathbf{k})+\eta g^*_{n'}(-\mathbf{k})\right]\ e^{i\mathbf{K}\cdot\left(\mathbf{R}'-\mathbf{R}\right)}\ \langle\widehat{n}_{\mathbf{k}_1}\rangle\,\langle\widehat{n}_{\mathbf{k}_2}\rangle \tag{46}$$

(as in § D-1-a of Chapter XVII, we may consider the summations over $\mathbf{k}_1$ and $\mathbf{k}_2$ as independent, since, for a large volume L^3, ignoring the constraint $\mathbf{k}_1 \neq \pm\mathbf{k}_2$ leads to a negligible error); we have used the notation:

$$\begin{aligned} \mathbf{K} &= \mathbf{k}_1 + \mathbf{k}_2 \\ \mathbf{k} &= \frac{\mathbf{k}_1 - \mathbf{k}_2}{2} \end{aligned} \tag{47}$$

When the parity of the function $g_{n'}(\mathbf{k})$ is η, the two terms in the bracket of (46) are equal and we get the simpler relation:

$$\left\langle \Phi^\dagger_{\chi_n}(\mathbf{R}) \Phi_{\chi_{n'}}(\mathbf{R}') \right\rangle_{\text{forward}} = \frac{1}{L^3} \sum_{\mathbf{k}_1,\mathbf{k}_2} g_n(\mathbf{k})\, g^*_{n'}(\mathbf{k})\ e^{i\mathbf{K}\cdot(\mathbf{R}'-\mathbf{R})}\ \langle \widehat{n}_{\mathbf{k}_1} \rangle \langle \widehat{n}_{\mathbf{k}_2} \rangle \tag{48}$$

This result only depends on the difference $\mathbf{R}-\mathbf{R}'$ (translation invariance); it goes to zero when $|\mathbf{R}'-\mathbf{R}|$ becomes larger than the inverse of the momentum $\mathbf{K}$ distribution width of the function appearing on the right-hand side of (46), once it is summed over $\mathbf{k}$, the difference in momenta.

– (**II**) The terms corresponding to the annihilation-creation of different pairs are obtained for $\mathbf{k}_2 = -\mathbf{k}_1$ and $\mathbf{k}_4 = -\mathbf{k}_3$, with $\mathbf{k}_4 \neq \pm\mathbf{k}_2$. Their contribution is written:

$$\begin{aligned} &\left\langle \Phi^\dagger_{\chi_n}(\mathbf{R}) \Phi_{\chi_{n'}}(\mathbf{R}') \right\rangle_{\text{paire-paire}} \\ &\quad = \frac{1}{2L^3} \sum_{\mathbf{k}_1} \left\langle a^\dagger_{\mathbf{k}_1} a^\dagger_{-\mathbf{k}_1} \right\rangle\ g_n(\mathbf{k}_1) \sum_{\mathbf{k}_4} \left\langle a_{-\mathbf{k}_4} a_{\mathbf{k}_4} \right\rangle\ g^*_{n'}(\mathbf{k}_4) \end{aligned} \tag{49}$$

Now, using (30) and the definition (2) of the Fourier components of each pair state $|\chi_n\rangle$, we have:

$$\frac{1}{L^{3/2}\sqrt{2}} \sum_{\mathbf{k}_4} \langle a_{-\mathbf{k}_4} a_{\mathbf{k}_4} \rangle\ g^*_{n'}(\mathbf{k}_4) = \sum_{\mathbf{k}_4} \langle \mathbf{k}_4\ |\phi_{\text{pair}}\rangle \langle \chi_{n'}\ |\mathbf{k}_4\rangle = \langle \chi_{n'}\ |\phi_{\text{pair}}\rangle \tag{50}$$

The summation over $\mathbf{k}_1$ is computed in a similar way, via a simple complex conjugation. We then get on the right-hand side of (49) two scalar products, which finally yields:

$$\left\langle \Phi^\dagger_{\chi_n}(\mathbf{R}) \Phi_{\chi_{n'}}(\mathbf{R}') \right\rangle_{\text{pair-pair}} = \langle \phi_{\text{pair}}\ |\chi_n\rangle \langle \chi_{n'}\ |\phi_{\text{pair}}\rangle \tag{51}$$

Unlike the previous contribution, this one is independent of $\mathbf{R}-\mathbf{R}'$.

– (**III**) The terms corresponding to the annihilation-creation of the same pair are obtained for $\mathbf{k}_1 = -\mathbf{k}_2 = \pm\mathbf{k}_3 = \mp\mathbf{k}_4$, and yield the average values $\langle \widehat{n}_{\mathbf{k}_1} \rangle \langle \widehat{n}_{-\mathbf{k}_1} \rangle$ and $\eta \langle \widehat{n}_{\mathbf{k}_1} \rangle \langle \widehat{n}_{-\mathbf{k}_1} \rangle$ respectively. Those terms are just a particular case of the terms appearing in the summation (46) when $\mathbf{k}_1 = -\mathbf{k}_2$, and do not require a specific calculation. Finally, the terms $\mathbf{k}_1 = \mathbf{k}_2$ that we ignored in (**I**), and for which all the $\mathbf{k}$'s must be equal, contain only one summation over the wave vectors; consequently, they are negligible compared to (46), and will be omitted in this computation.

We are then left with the total (**I**) + (**II**), which yields:

$$\left\langle \Phi^\dagger_{\chi_n}(\mathbf{R})\, \Phi_{\chi_{n'}}(\mathbf{R}') \right\rangle = \left\langle \Phi^\dagger_{\chi_n}(\mathbf{R}) \Phi_{\chi_{n'}}(\mathbf{R}') \right\rangle_{\text{forward}} + \left\langle \Phi^\dagger_{\chi_n}(\mathbf{R}) \Phi_{\chi_{n'}}(\mathbf{R}') \right\rangle_{\text{pair-pair}} \tag{52}$$

where only the second term on the right-hand side does not go to zero when $|\mathbf{R}' - \mathbf{R}|$ becomes large, which indicates a long-range non-diagonal order. According to (51), this second term reaches a maximum when the two internal states $|\chi_n\rangle$ and $|\chi_{n'}\rangle$ are equal to the state $\left|\phi_{\text{pair}}^{\text{norm}}\right\rangle$ defined in (32). It indicates the existence of a cooperative field of pairs that have a total momentum $\mathbf{K} = 0$ as their external state, and $\left|\phi_{\text{pair}}^{\text{norm}}\right\rangle$ as their internal state.

Comparing (31) and (51) shows that:

$$\left\langle \Phi_{\chi_n}^{\dagger}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle_{\text{pair-pair}} = \left\langle \Phi_{\chi_n}(\mathbf{R})\right\rangle^{*} \times \left\langle \Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle \tag{53}$$

The pair-pair term of the two-point correlation function can thus be factored into a product of two one-point correlation functions; for $n = 1$, we get the same function we previously called the "order parameter". As already pointed out in § 2-a-α, it is because the pairs have a zero total momentum that any $\mathbf{R}$ and $\mathbf{R}'$ dependence has disappeared from both sides of (53), but this point is not essential. It is more important to note that introducing such an order parameter, a priori difficult to understand from a physical point of view as it is an average value that does not conserve particle number, is actually quite useful for computing other more physical parameters. We will make the connection between the factorization relation (53) and the Penrose-Onsager criterion for Bose-Einstein condensation in § 2-b-γ.

β. Singlet pairs

Using (26) instead of (13) now leads to a relation very similar to (42); the factor 1/2 is, however, missing, and we must make the substitution:

$$\left\langle a_{\mathbf{k}_1}^{\dagger} a_{\mathbf{k}_2}^{\dagger} a_{\mathbf{k}_3} a_{\mathbf{k}_4}\right\rangle \Rightarrow \left\langle a_{\mathbf{k}_1,\uparrow}^{\dagger} a_{\mathbf{k}_2,\downarrow}^{\dagger} a_{\mathbf{k}_3,\downarrow} a_{\mathbf{k}_4,\uparrow}\right\rangle \tag{54}$$

Relation (44) then becomes:

$$\begin{aligned}\left\langle \Phi_{\chi_n}^{\dagger}(\mathbf{R})\,\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle = \frac{1}{L^3}\sum_{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4} & g_n\left(\frac{\mathbf{k}_1-\mathbf{k}_2}{2}\right)\; g_{n'}^{*}\left(\frac{\mathbf{k}_4-\mathbf{k}_3}{2}\right) \\ \times\; e^{i\left[(\mathbf{k}_3+\mathbf{k}_4)\cdot\mathbf{R}'-(\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{R}\right]} & \left\langle a_{\mathbf{k}_1,\uparrow}^{\dagger} a_{\mathbf{k}_2,\downarrow}^{\dagger} a_{\mathbf{k}_3,\downarrow} a_{\mathbf{k}_4,\uparrow}\right\rangle\end{aligned} \tag{55}$$

The rest of the calculation is very similar to the one we just did, and involves the sum of several terms:

– (**I**) The forward scattering terms are obtained for $\mathbf{k}_4 = \mathbf{k}_1$ and $\mathbf{k}_3 = \mathbf{k}_2$. In two different pairs, a particle is destroyed and then created again in the same individual state (as we now have spin indices, there is no exchange term in this case). The computation is the same as the one that yielded (48) for spinless particles; with the notation (47) for the wave vectors, we get here:

$$\left\langle \Phi_{\chi_n}^{\dagger}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle_{\text{forward}} = \frac{1}{L^3}\sum_{\mathbf{k}_1,\mathbf{k}_2} g_n(\mathbf{k})\, g_{n'}^{*}(\mathbf{k})\;\; e^{i\mathbf{K}\cdot(\mathbf{R}'-\mathbf{R})}\;\; \langle \widehat{n}_{\mathbf{k}_1,\uparrow}\rangle\, \langle \widehat{n}_{\mathbf{k}_2,\downarrow}\rangle \tag{56}$$

– (**II**) The terms corresponding to the annihilation-creation of different pairs are obtained for $\mathbf{k}_2 = -\mathbf{k}_1$ and $\mathbf{k}_4 = -\mathbf{k}_3$, with $\mathbf{k}_4 \neq \mathbf{k}_1$. The computation is now the same

as the one that yielded (51). The right-hand side of (55) becomes:

$$\frac{1}{L^3}\sum_{\mathbf{k}_1,\mathbf{k}_4} g_n(\mathbf{k}_1)\, g^*_{n'}(\mathbf{k}_4)\left\langle a^\dagger_{\mathbf{k}_1\uparrow} a^\dagger_{-\mathbf{k}_1\downarrow} a_{-\mathbf{k}_4\downarrow} a_{\mathbf{k}_4\uparrow}\right\rangle \tag{57}$$

Using the definition (38) for the pair wave function in the singlet case, we again obtain:

$$\left\langle \Phi^\dagger_{\chi_n}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle_{\text{pair-pair}} = \langle \phi_{\text{pair}}\,|\chi_n\rangle\,\langle\chi_{n'}\,|\phi_{\text{pair}}\rangle \tag{58}$$

As mentioned above, any $\mathbf{R}$ dependence has disappeared from this average value since the paired state was built from pairs having a zero total momentum.

– (**III**) The terms corresponding to the annihilation-creation of the same pair are obtained for $\mathbf{k}_2 = \mathbf{k}_3 = -\mathbf{k}_1 = -\mathbf{k}_4$; they are proportional to $\langle \widehat{n}_{\mathbf{k}_1,\uparrow}\rangle \left\langle \widehat{n}_{-\mathbf{k}_1,\downarrow}\right\rangle$ and already included in the terms (**I**). The terms where all the $\mathbf{k}$'s are equal are neglected for the same reason as above.

To sum up, we find as before:

$$\begin{aligned}\left\langle \Phi^\dagger_{\chi_n}(\mathbf{R})\,\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle &= \left\langle \Phi^\dagger_{\chi_n}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle_{\text{forward}} + \left\langle \Phi^\dagger_{\chi_n}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle_{\text{pair-pair}} \\ &= \left\langle \Phi^\dagger_{\chi_n}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle_{\text{forward}} + \langle \Phi_{\chi_n}(\mathbf{R})\rangle^* \left\langle \Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle \end{aligned} \tag{59}$$

We arrive, finally, at the same results as for spinless bosons, with the same long-range non-diagonal order of the pairs, as well as the factorization (53) of the order parameters. We shall see in Complement C_{XVII} that this long-range order parameter is intimately linked to the nature of the BCS transition. Here again, the anomalous average values turn out to be useful tools for computing normal average values that conserve the particle number.

γ. *Link with Bose-Einstein condensation of pairs*

There is a close link between the order parameter of the pairs and the existence of Bose-Einstein condensation of those pairs. To show this, it is convenient to introduce the density operator of pairs, limiting ourselves, for the sake of simplicity, to the case of spinless particles. In Chapter XVII, the one-particle density operator ρ_I for identical particles was given, in terms of the field operator, by its matrix elements (B-26):

$$\langle \mathbf{r}', \nu'|\,\rho_I\,|\mathbf{r},\nu\rangle = \left\langle \Psi^\dagger_\nu(\mathbf{r})\Psi_{\nu'}(\mathbf{r}')\right\rangle \tag{60}$$

where $\mathbf{r}$ is the particle's position and ν its spin. For pairs, the corresponding relation is written[2]:

$$\langle \mathbf{R}', \chi_{n'}|\,\rho_I^{\text{pair}}\,|\mathbf{R},\chi_n\rangle = \left\langle \Phi^\dagger_{\chi_n}(\mathbf{R})\Phi_{\chi_{n'}}(\mathbf{R}')\right\rangle \tag{61}$$

where $\mathbf{R}$ is the position of the center of mass, and χ_n and $\chi_{n'}$ define the internal state of the pair; the index n plays a role similar to that of a spin index for a single particle (even though it corresponds to an internal orbital state).

[2] As we shall see in § 3, the pair field operators do not exactly satisfy the boson commutation relations. Consequently, operator (61) is not, strictly speaking, a density operator; to underline this difference, ρ_I^{pair} is sometimes called a "density quasi-operator".

In the momentum representation, the diagonal matrix elements of this density operator are:

$$\langle \mathbf{K}', \chi_n | \, \rho_I^{\text{pair}} \, | \mathbf{K}', \chi_n \rangle = \frac{1}{L^3} \int \mathrm{d}^3 R \int \mathrm{d}^3 R' \; e^{i\mathbf{K}'\cdot(\mathbf{R}-\mathbf{R}')} \left\langle \Phi_{\chi_n}^\dagger(\mathbf{R}) \Phi_{\chi_n}(\mathbf{R}') \right\rangle \tag{62}$$

Since $\left\langle \Phi_{\chi_n}^\dagger(\mathbf{R}) \Phi_{\chi_{n'}}(\mathbf{R}') \right\rangle$ only depends on $\mathbf{R} - \mathbf{R}'$, we perform the change of variables $\mathbf{X} = \mathbf{R} - \mathbf{R}'$; the integral over $\mathrm{d}^3 R$ is then trivial and cancels the factor $1/L^3$; we therefore get:

$$\langle \mathbf{K}', \chi_n | \, \rho_I^{\text{pair}} \, | \mathbf{K}', \chi_n \rangle = \int \mathrm{d}^3 X \; e^{i\mathbf{K}'\cdot\mathbf{X}} \; \left\langle \Phi_{\chi_n}^\dagger(\mathbf{R}) \Phi_{\chi_n}(\mathbf{R} - \mathbf{X}) \right\rangle \tag{63}$$

Inserting relation (52) in this result, we get the sum of a contribution from the forward scattering term and from the pair-pair term.

(i) The first contribution comes from inserting (48) in (63). The integral over $\mathrm{d}^3 X$ yields a delta function $\delta_{\mathbf{K},\mathbf{K}'}$ and a factor L^3 that cancels the same factor in the denominator. As the sum $\mathbf{k}_1 + \mathbf{k}_2$ must now be equal to $\mathbf{K}'$, the double summation over $\mathbf{k}_1$ and $\mathbf{k}_2$ reduces to a summation over $\mathbf{k}$. We then get:

$$\sum_{\mathbf{k}} |g_n(\mathbf{k})|^2 \; \left\langle \widehat{n}_{\frac{\mathbf{K}'}{2}+\mathbf{k}} \right\rangle \left\langle \widehat{n}_{\frac{\mathbf{K}'}{2}-\mathbf{k}} \right\rangle \tag{64}$$

This result is a regular function of $\mathbf{K}'$, related to the wave number dependence of the occupation numbers.

(ii) The pair-pair term contains the integral of the function (53), which is a product of two constant order parameters; it therefore leads to:

$$\delta_{\mathbf{K},\mathbf{0}} \; L^3 \left| \left\langle \Phi_{\chi_n} \right\rangle \right|^2 \tag{65}$$

The presence of the delta function $\delta_{\mathbf{K},\mathbf{0}}$ shows that the $\mathbf{K} = 0$ level has an additional population (number of quanta of the pair field) that does not exist for any other value of the momentum $\mathbf{K}$; this population is simply the square of the order parameter, multiplied by the system's volume; it is thus an extensive quantity. It indicates that the pairs of the system undergo Bose-Einstein condensation. As the corresponding population is proportional to the square of the order parameter, this clearly shows the close link between the long-range non-diagonal order, the order parameter and the existence of condensation. The factorization appearing in (53) is often called the "Penrose-Onsager condensation criterion".

2-c. Application to the computation of the correlation function (singlet pairs)

The average values of products of pair field operators can also be used to get the correlation functions between particles. We are going to show, in particular, that the correlation function G_2 is the sum of an "incoherent" term, independent of the positions, and of a coherent term that involves the pair wave function defined previously. In order to keep the demonstration short, we shall limit the discussion to the case of fermions described by a paired state built from singlet pairs[3], but the transposition to spinless particles is fairly straightforward.

[3]Condensed bosons will be studied in Complement E_{XVII}. We will then show that the properties of the paired state, built from the $\mathbf{k} \neq 0$ states, are not determined by the interactions within this paired state, but rather by the interactions with a condensate $\mathbf{k} = 0$, external to the paired state. It is therefore a completely different case.

Relation (24) expresses the conjugate of the pair field operator as a function of products of creation operators for its constituent particles; we shall start by inverting this relation.

α. Inversion of the relation between fields

The closure relation for the orthonormal basis of the wave functions $\chi_p(\mathbf{r})$ with $p = 1, 2, ...$ is written:

$$\sum_n \chi_n^*(\mathbf{x}')\ \chi_n(\mathbf{x}) = \delta(\mathbf{x} - \mathbf{x}') \tag{66}$$

This summation over n must include even orbital functions $\chi_n(\mathbf{x})$ (associated with a pair field Φ_{χ_n} describing paired fermions in a singlet state) as well as odd functions (associated with a pair field describing fermions paired in a triplet state). We then multiply (24) by $\chi_n^*(\mathbf{x}')$ and perform the summation over n. We recognize in the integral on the right-hand side the closure relation (66), which yields:

$$\sum_n \chi_n^*(\mathbf{x}')\ \Phi_{\chi_n}^\dagger(\mathbf{R}) = \Psi_\uparrow^\dagger(\mathbf{R} + \frac{\mathbf{x}'}{2})\Psi_\downarrow^\dagger(\mathbf{R} - \frac{\mathbf{x}'}{2}) \tag{67}$$

This leads to:

$$\Psi_\uparrow^\dagger(\mathbf{r}_1)\Psi_\downarrow^\dagger(\mathbf{r}_2) = \sum_n \chi_n^*(\mathbf{r}_1 - \mathbf{r}_2)\ \Phi_{\chi_n}^\dagger\left(\frac{\mathbf{r}_1 + \mathbf{r}_2}{2}\right) \tag{68}$$

Creating two particles of opposite spins at points $\mathbf{r}_1$ and $\mathbf{r}_2$ thus amounts to creating a coherent superposition of pairs with a center of mass at $(\mathbf{r}_1 + \mathbf{r}_2)/2$, in a singlet or triplet spin state, and with coefficients equal to the wave functions χ_n^* taken at the position $\mathbf{r}_1 - \mathbf{r}_2$.

The average value of this expression can be computed in a paired state, using relation (37). This leads to:

$$\left\langle \Psi_\uparrow^\dagger(\mathbf{r}_1)\Psi_\downarrow^\dagger(\mathbf{r}_2) \right\rangle = \sum_n \chi_n^*(\mathbf{r}_1 - \mathbf{r}_2)\ \ \langle \chi_n\, |\phi_{\text{pair}}\rangle \tag{69}$$

As ϕ_{pair} is an even function, it easily follows that only the even χ_n will contribute to this average value; the triplet pair fields have a zero average value in a singlet pair state.

As in § 2-a-β, we can choose for the $|\chi_n\rangle$ a basis whose first ket $|\chi_1\rangle$ coincides with the normalized pair ket $\left|\phi_{\text{pair}}^{\text{norm}}\right\rangle$. We then get:

$$\left\langle \Psi_\uparrow^\dagger(\mathbf{r}_1)\Psi_\downarrow^\dagger(\mathbf{r}_2) \right\rangle = \frac{\phi_{\text{pair}}^*(\mathbf{r}_1 - \mathbf{r}_2)}{\sqrt{\langle\phi_{\text{pair}}\,|\phi_{\text{pair}}\rangle}} \frac{\langle\phi_{\text{pair}}\,|\phi_{\text{pair}}\rangle}{\sqrt{\langle\phi_{\text{pair}}\,|\phi_{\text{pair}}\rangle}} = \phi_{\text{pair}}^*(\mathbf{r}_1 - \mathbf{r}_2) \tag{70}$$

β. 4-point correlation function

According to (68), the 4-point correlation function (for opposite spins) is written as:

$$\begin{aligned}&\left\langle \Psi_\uparrow^\dagger(\mathbf{r}_1)\Psi_\downarrow^\dagger(\mathbf{r}_2)\Psi_\downarrow(\mathbf{r}_2')\Psi_\uparrow(\mathbf{r}_1') \right\rangle \\ &= \sum_{n,n'} \chi_n^*(\mathbf{r}_1 - \mathbf{r}_2)\,\chi_{n'}(\mathbf{r}_1' - \mathbf{r}_2') \left\langle \Phi_{\chi_n}^\dagger\left(\frac{\mathbf{r}_1 + \mathbf{r}_2}{2}\right)\Phi_{\chi_{n'}}\left(\frac{\mathbf{r}_1' + \mathbf{r}_2'}{2}\right)\right\rangle\end{aligned} \tag{71}$$

It is expressed in terms of the average values of products of pair creation and annihilation operators, hence in terms of the average values of products of fields for which the index n plays the role of an internal state of the molecule. We will show below that it can be expressed as:

$$\left\langle \Psi_{\uparrow}^{\dagger}(\mathbf{r}_1)\Psi_{\downarrow}^{\dagger}(\mathbf{r}_2)\Psi_{\downarrow}(\mathbf{r}_2')\Psi_{\uparrow}(\mathbf{r}_1')\right\rangle = G_1\left(\mathbf{r}_1,\uparrow;\mathbf{r}_1',\uparrow\right)\times G_1\left(\mathbf{r}_2,\downarrow;\mathbf{r}_2',\downarrow\right) + \phi_{\text{pair}}^{*}\left(\mathbf{r}_1-\mathbf{r}_2\right)\times\phi_{\text{pair}}\left(\mathbf{r}_1'-\mathbf{r}_2'\right) \tag{72}$$

where $G_1\left(\mathbf{r},\uparrow;\mathbf{r}',\uparrow\right)$ is the non-diagonal one-particle correlation function, the Fourier transform of $\langle n_{\mathbf{k},\uparrow}\rangle$:

$$G_1\left(\mathbf{r},\uparrow;\mathbf{r}',\uparrow\right)=\frac{1}{L^3}\sum_{\mathbf{k}}e^{i\mathbf{k}\cdot\left(\mathbf{r}'-\mathbf{r}\right)}\ \langle n_{\mathbf{k},\uparrow}\rangle \tag{73}$$

and with a similar definition for $G_1\left(\mathbf{r},\downarrow;\mathbf{r}',\downarrow\right)$, the occupation number $\langle n_{\mathbf{k},\uparrow}\rangle$ being simply replaced by $\langle n_{\mathbf{k},\downarrow}\rangle$; the pair wave function ϕ_{pair} has already been defined in (38).

The function $G_1\left(\mathbf{r},\uparrow;\mathbf{r}',\uparrow\right)$, being the Fourier transform of a regular function $\langle n_{\mathbf{k},\uparrow}\rangle$, tends toward zero when the difference $|\mathbf{r}-\mathbf{r}'|$ is larger than a certain (microscopic) limit; the only terms left are those on the second line of (72). Imagine then that positions $\mathbf{r}_1$ and $\mathbf{r}_2$ are close to each other, forming a first group, and that the same is true for positions $\mathbf{r}_1'$ and $\mathbf{r}_2'$, forming a second group, while these two groups are far from each other. The non-diagonal correlation function can then be factored into a product of functions ϕ_{pair}. This situation is reminiscent of the Penrose-Onsager criterion for Bose-Einstein condensation of bosons (Complement A_{XVI}, § 3-a), but it now concerns the 4-point (instead of 2-point) non-diagonal correlation function. As the norm of ϕ_{pair} is the order parameter, it again underlines the important role of this parameter.

An important particular case of the 4-point correlation function is the two-body (diagonal) correlation function for opposite spins:

$$G_2\left(\mathbf{r}_1,\uparrow;\mathbf{r}_2,\downarrow\right)=\left\langle \Psi_{\uparrow}^{\dagger}(\mathbf{r}_1)\Psi_{\downarrow}^{\dagger}(\mathbf{r}_2)\Psi_{\downarrow}(\mathbf{r}_2)\Psi_{\uparrow}(\mathbf{r}_1)\right\rangle \tag{74}$$

The intensity of the pair field is therefore written:

$$\begin{aligned} G_2\left(\mathbf{r}_1,\uparrow;\mathbf{r}_2,\downarrow\right) &= \frac{1}{L^6}\sum_{\mathbf{k}_1,\mathbf{k}_2}\langle\widehat{n}_{\mathbf{k}_1,\uparrow}\rangle\langle\widehat{n}_{\mathbf{k}_2,\downarrow}\rangle+\left|\phi_{\text{pair}}\left(\mathbf{r}_1-\mathbf{r}_2\right)\right|^2 \\ &= \frac{\left\langle\widehat{N_{\uparrow}}\right\rangle\left\langle\widehat{N_{\downarrow}}\right\rangle}{L^6}+\left|\phi_{\text{pair}}\left(\mathbf{r}_1-\mathbf{r}_2\right)\right|^2 \end{aligned} \tag{75}$$

We find again relation (D-17) of Chapter XVII, but via another method. The two-body correlation function is the sum of a contribution independent of the positions (hence, with no correlations) and of the modulus squared of the pair wave function. This latter contribution comes from the term that, for pairs, indicates the existence of a long-range non-diagonal order (Bose-Einstein condensation). This is an important property, which is at the heart of the BCS mechanism, and which will be discussed in more detail in Complement B_{XVII}.

Demonstration

We insert in (71) relations (56) and (58). In the forward scattering term, we get the following expression:

$$\begin{aligned}
&\sum_{n,n'} \chi_n^*(\mathbf{r}_1-\mathbf{r}_2)\,\chi_{n'}\left(\mathbf{r}_1'-\mathbf{r}_2'\right) g_n(\mathbf{k})\, g_{n'}^*(\mathbf{k})\, e^{i\mathbf{K}\cdot\left(\mathbf{r}_1'+\mathbf{r}_2'-\mathbf{r}_1-\mathbf{r}_2\right)/2} \\
&= \sum_{n,n'} \langle\chi_n\,|\mathbf{r}_1-\mathbf{r}_2\rangle\ \left\langle\mathbf{r}_1'-\mathbf{r}_2'\,|\chi_{n'}\right\rangle\ \langle\mathbf{k}\,|\chi_n\rangle\ \langle\chi_{n'}\,|\mathbf{k}\rangle\ e^{i\mathbf{K}\cdot\left(\mathbf{r}_1'+\mathbf{r}_2'-\mathbf{r}_1-\mathbf{r}_2\right)/2} \\
&= \langle\mathbf{k}\,|\mathbf{r}_1-\mathbf{r}_2\rangle\left\langle\mathbf{r}_1'-\mathbf{r}_2'\,|\mathbf{k}\right\rangle\ e^{i\mathbf{K}\cdot\left(\mathbf{r}_1'+\mathbf{r}_2'-\mathbf{r}_1-\mathbf{r}_2\right)/2} \\
&= \frac{1}{L^3}\ e^{i(\mathbf{k}_2-\mathbf{k}_1)\cdot\left(\mathbf{r}_1'-\mathbf{r}_2'-\mathbf{r}_1+\mathbf{r}_2\right)/2}\ e^{i(\mathbf{k}_1+\mathbf{k}_2)\cdot\left(\mathbf{r}_1'+\mathbf{r}_2'-\mathbf{r}_1-\mathbf{r}_2\right)/2} \\
&= \frac{1}{L^3}\ e^{i\mathbf{k}_2\cdot\left(\mathbf{r}_1'-\mathbf{r}_1\right)}\ e^{i\mathbf{k}_1\cdot\left(\mathbf{r}_2'-\mathbf{r}_2\right)}
\end{aligned} \tag{76}$$

where $\mathbf{k}$ and $\mathbf{K}$ were defined in (47). Inserting this result in (71), we get the first term of the right-hand side of (72)

As for the pair annihilation-creation term (58), it yields:

$$\begin{aligned}
&\left\langle\Psi_\uparrow^\dagger(\mathbf{r}_1)\Psi_\downarrow^\dagger(\mathbf{r}_2)\Psi_\downarrow(\mathbf{r}_2')\Psi_\uparrow(\mathbf{r}_1')\right\rangle_{\text{pair-pair}} \\
&= \sum_{n,n'} \chi_n^*(\mathbf{r}_1-\mathbf{r}_2)\,\chi_{n'}\left(\mathbf{r}_1'-\mathbf{r}_2'\right)\langle\phi_{\text{pair}}\,|\chi_n\rangle\,\langle\chi_{n'}\,|\phi_{\text{pair}}\rangle \\
&= \sum_{n,n'} \left\langle\mathbf{r}_1'-\mathbf{r}_2'\,|\chi_{n'}\right\rangle\langle\chi_{n'}\,|\phi_{\text{pair}}\rangle\times\langle\phi_{\text{pair}}\,|\chi_n\rangle\,\langle\chi_n\,|\mathbf{r}_1-\mathbf{r}_2\rangle \\
&= \phi_{\text{pair}}^*(\mathbf{r}_1-\mathbf{r}_2)\times\phi_{\text{pair}}\left(\mathbf{r}_1'-\mathbf{r}_2'\right)
\end{aligned} \tag{77}$$

and we obtain the second term of the right-hand side of (72). Note that only the singlet pair fields (associated with the even function χ_n) contribute to this term.

3. Commutation relations of field operators

We now study the commutation relations between the pair field operators just defined. The "spin-statistics theorem" (Chapter XIV, § C-1) states that particles with integer spin are bosons, and particles with half-integer spin are fermions. If we consider two paired fermions, the rules for adding angular momenta (Chapter X) indicate that this composite system necessarily has an integer spin. Intuitively, one could thus expect two bound fermions to behave like a boson; this is the question we now discuss by examining the commutation relations between the operators $A_{\mathbf{K}}$ and $A_{\mathbf{K}}^\dagger$, and establishing the correction factors introduced by the underlying fermionic structure.

3-a. Particles in the same spin state

Starting with spinless particles, we shall explain in this simple case the main commutation properties of the pair operators. If the pairs created and annihilated by the operators $A_{\mathbf{K}}$ and $A_{\mathbf{K}'}$ and their Hermitian conjugates were really bosons, the commutator of these two operators should be equal to $\delta_{\mathbf{K}\mathbf{K}'}$. We are going to show that the commutator does contain such a term, but with several additional corrections.

α. Commutation relations of the $A_{\mathbf{K},n}$

Any product of two creation operators commute with any product of two creation operators (for fermions, two minus signs cancel each other when products of two operators cross each other); the same is true for two products of annihilation operators. We therefore have:

$$\begin{aligned}\left[A^\dagger_{\mathbf{K},n}, A^\dagger_{\mathbf{K}',n'}\right] &= 0\\ \left[A_{\mathbf{K},n}, A_{\mathbf{K}',n'}\right] &= 0\end{aligned} \tag{78}$$

The commutator of $A_{\mathbf{K},n}$ and $A^\dagger_{\mathbf{K}',n'}$ has yet to be computed:

$$\left[A_{\mathbf{K},n}, A^\dagger_{\mathbf{K}',n'}\right] = \frac{1}{2}\sum_{\mathbf{k}} \left(g^n_{\mathbf{k}}\right)^* \sum_{\mathbf{k}'} g^{n'}_{\mathbf{k}'} \left[a_{\frac{\mathbf{K}}{2}-\mathbf{k}} a_{\frac{\mathbf{K}}{2}+\mathbf{k}} \ , \ a^\dagger_{\frac{\mathbf{K}'}{2}+\mathbf{k}'} a^\dagger_{\frac{\mathbf{K}'}{2}-\mathbf{k}'}\right] \tag{79}$$

We will show below (§ 3-a-β) that:

$$\begin{aligned}\left[A_{\mathbf{K},n}, A^\dagger_{\mathbf{K}',n'}\right] &= \delta_{\mathbf{K},\mathbf{K}'}\ \delta_{n,n'} + 2\eta \sum_{\mathbf{k}} \left(g^n_{\mathbf{k}}\right)^*\ g^{n'}_{\frac{\mathbf{K}-\mathbf{K}'}{2}+\mathbf{k}}\ a^\dagger_{\mathbf{K}'-(\mathbf{K}/2)-\mathbf{k}}\ a_{(\mathbf{K}/2)-\mathbf{k}}\\ &= \delta_{\mathbf{K},\mathbf{K}'}\ \delta_{n,n'} + 2\eta \sum_{\boldsymbol{\kappa}} \left(g^n_{\boldsymbol{\kappa}-\frac{\mathbf{K}}{2}}\right)^*\ g^{n'}_{\boldsymbol{\kappa}-\frac{\mathbf{K}'}{2}}\ a^\dagger_{\mathbf{K}'-\boldsymbol{\kappa}}\ a_{\mathbf{K}-\boldsymbol{\kappa}}\end{aligned} \tag{80}$$

(in the second line, we have set $\boldsymbol{\kappa} = \mathbf{k} + \mathbf{K}/2$ and used the parity η of the function g); if needed, we can get rid of the η coefficient on the right-hand side provided we change the sign of the subscript of g (or of g^*).

The first term $\delta_{\mathbf{K},\mathbf{K}'}\ \delta_{n,n'}$, on the right-hand side of (80) , is exactly the commutator of two bosons with internal states n and n' (spin states for example): this term is different from zero only if both the external and internal variables are the same (in the present case, these internal states are actually orbital states). This first term is, however, followed by an additional term that shows that the fermionic structure of the pairs still plays a role. This latter term is a one-particle operator in the sense defined in § B of Chapter XV; relation (B-12) of that chapter permits computing the matrix elements of the corresponding operator $\widehat{f}$. This additional term contains creation and annihilation operators in normal order, which means that it will go to zero when the populations of the individual states tend to zero; in this limit, the pairs can be assimilated to bosons.

When $\mathbf{K}' = \mathbf{K}$ and $n = n'$ (pairs in the same internal and external states), we get the simpler relation:

$$\left[A_{\mathbf{K},n}, A^\dagger_{\mathbf{K},n}\right] = 1 + 2\eta \sum_{\boldsymbol{\kappa}} \left|g^n_{\boldsymbol{\kappa}-\frac{\mathbf{K}}{2}}\right|^2\ \widehat{n}_{\mathbf{K}-\boldsymbol{\kappa}} \tag{81}$$

with the usual definition of the population operator $\widehat{n}_{\mathbf{k}}$:

$$\widehat{n}_{\mathbf{k}} = a^\dagger_{\mathbf{k}} a_{\mathbf{k}} \tag{82}$$

The corrections to a purely bosonic commutation are then proportional to the populations of the individual states of the particles forming the pair, hence confirming the fact that they become negligible when the sum of all these populations is small enough.

β. Demonstration

We start by computing the commutator:

$$\left[a_1 a_2\ ,\ a_3^\dagger a_4^\dagger\right] = a_1 a_2 a_3^\dagger a_4^\dagger - a_3^\dagger a_4^\dagger a_1 a_2 \tag{83}$$

where $1, 2, 3, 4$ are indices labeling any individual states. We can write:

$$\begin{aligned} a_1 a_2 a_3^\dagger a_4^\dagger &= \delta_{23}\ a_1 a_4^\dagger + \eta\ a_1 a_3^\dagger a_2 a_4^\dagger \\ &= \delta_{23}\ a_1 a_4^\dagger + \eta\delta_{13}\ a_2 a_4^\dagger + a_3^\dagger a_1 a_2 a_4^\dagger \\ &= \delta_{23}\ a_1 a_4^\dagger + \eta\delta_{13}\ a_2 a_4^\dagger + \delta_{24}\ a_3^\dagger a_1 + \eta\ a_3^\dagger a_1 a_4^\dagger a_2 \\ &= \delta_{23}\ a_1 a_4^\dagger + \eta\delta_{13}\ a_2 a_4^\dagger + \delta_{24}\ a_3^\dagger a_1 + \eta\ \delta_{14}\ a_3^\dagger a_2 + a_3^\dagger a_4^\dagger a_1 a_2 \end{aligned} \tag{84}$$

so that:

$$\left[a_1 a_2\ ,\ a_3^\dagger a_4^\dagger\right] = \delta_{23}\ a_1 a_4^\dagger + \delta_{24}\ a_3^\dagger a_1 + \eta\left[\delta_{13}\ a_2 a_4^\dagger + \delta_{14}\ a_3^\dagger a_2\right] \tag{85}$$

Putting all the operators in normal order, we get[4]:

$$\left[a_1 a_2\ ,\ a_3^\dagger a_4^\dagger\right] = \delta_{23}\delta_{14} + \eta\delta_{13}\delta_{24} + \delta_{24}\ a_3^\dagger a_1 + \delta_{13}\ a_4^\dagger a_2 + \eta\left[\delta_{23}\ a_4^\dagger a_1 + \delta_{14}\ a_3^\dagger a_2\right] \tag{86}$$

The commutator appearing on the right-hand side of (79) is therefore equal to:

$$\begin{aligned} &\delta_{\mathbf{K},\mathbf{K}'}\ \delta_{\mathbf{k},\mathbf{k}'} + \eta\delta_{\mathbf{K},\mathbf{K}'}\ \delta_{\mathbf{k},-\mathbf{k}'} + \Big[\delta_{(\mathbf{K}-\mathbf{K}')/2,-\mathbf{k}-\mathbf{k}'}\ a^\dagger_{(\mathbf{K}'/2)+\mathbf{k}'}\ a_{(\mathbf{K}/2)-\mathbf{k}} \\ &+ \delta_{(\mathbf{K}-\mathbf{K}')/2,\mathbf{k}+\mathbf{k}'}\ a^\dagger_{(\mathbf{K}'/2)-\mathbf{k}'}\ a_{(\mathbf{K}/2)+\mathbf{k}} \\ &+\eta\delta_{(\mathbf{K}-\mathbf{K}')/2,\mathbf{k}'-\mathbf{k}}\ a^\dagger_{(\mathbf{K}'/2)-\mathbf{k}'}\ a_{(\mathbf{K}/2)-\mathbf{k}} + \eta\delta_{(\mathbf{K}-\mathbf{K}')/2,\mathbf{k}-\mathbf{k}'}\ a^\dagger_{(\mathbf{K}'/2)+\mathbf{k}'}\ a_{(\mathbf{K}/2)+\mathbf{k}}\Big] \end{aligned} \tag{87}$$

Inserting the first two terms back into (79), we get the following contribution to the commutator $\left[A_{\mathbf{K},n}, A^\dagger_{\mathbf{K}',n'}\right]$:

$$\frac{1}{2}\delta_{\mathbf{K},\mathbf{K}'}\sum_{\mathbf{k}}\ (g^n_{\mathbf{k}})^*\ \left[g^{n'}_{\mathbf{k}} + \eta g^{n'}_{-\mathbf{k}}\right] = \delta_{\mathbf{K},\mathbf{K}'}\sum_{\mathbf{k}}\ (g^n_{\mathbf{k}})^*\, g^{n'}_{\mathbf{k}} = \delta_{\mathbf{K},\mathbf{K}'}\ \delta_{n,n'} \tag{88}$$

where we have taken into account the parity with respect to $\mathbf{k}$ of the the functions $g^n_{\mathbf{k}}$ – see relation (A-4) of Chapter XVII – and used the fact that the internal states are orthonormal. As mentioned above, this $\delta_{\mathbf{K},\mathbf{K}'}\ \delta_{n,n'}$ is precisely what is expected for a boson commutation relation.

It is, however, followed in (87) by four other terms, which are written:

$$\begin{aligned} &\frac{1}{2}\sum_{\mathbf{k}}\ (g^n_{\mathbf{k}})^*\ g^{n'}_{-\frac{\mathbf{K}-\mathbf{K}'}{2}-\mathbf{k}}\ a^\dagger_{\mathbf{K}'-(\mathbf{K}/2)-\mathbf{k}} a_{(\mathbf{K}/2)-\mathbf{k}} \\ &\frac{1}{2}\sum_{\mathbf{k}}\ (g^n_{\mathbf{k}})^*\ g^{n'}_{\frac{\mathbf{K}-\mathbf{K}'}{2}-\mathbf{k}}\ a^\dagger_{\mathbf{K}'-(\mathbf{K}/2)+\mathbf{k}} a_{(\mathbf{K}/2)+\mathbf{k}} \\ &\frac{\eta}{2}\sum_{\mathbf{k}}\ (g^n_{\mathbf{k}})^*\ g^{n'}_{\frac{\mathbf{K}-\mathbf{K}'}{2}+\mathbf{k}}\ a^\dagger_{\mathbf{K}'-(\mathbf{K}/2)-\mathbf{k}} a_{(\mathbf{K}/2)-\mathbf{k}} \\ &\frac{\eta}{2}\sum_{\mathbf{k}}\ (g^n_{\mathbf{k}})^*\ g^{n'}_{-\frac{\mathbf{K}-\mathbf{K}'}{2}+\mathbf{k}}\ a^\dagger_{\mathbf{K}'-(\mathbf{K}/2)+\mathbf{k}} a_{(\mathbf{K}/2)+\mathbf{k}} \end{aligned} \tag{89}$$

[4]Since $\left[a_i, a_j^\dagger\right]_{-\eta} = \delta_{ij}$, we have $a_i a_j^\dagger = \eta a_j^\dagger a_i + \delta_{ij}$.

In each of them, and without modifying the result, we can change the sign of the summation dummy index $\mathbf{k}$, or change the sign of the subscript of the functions g or g^* (provided we introduce a factor η). For example, in the second term, we can change the sign of the subscripts of the two functions g and g^* (two factors η then cancel each other), then change the summation index $\mathbf{k}$ into $-\mathbf{k}$: this second term then doubles the first one. As for the third term, we simply change the sign of the subscript $(\mathbf{K}-\mathbf{K}')/2+\mathbf{k}$ of the function g (which introduces a factor η canceling that same factor already present) and reproduce the first term. Finally, for the fourth term, a parity operation on the function g^* followed by a change of the summation index from $\mathbf{k}$ to $-\mathbf{k}$ makes it equal to the first. The four terms are therefore equal; choosing for example the expression of the third one, and replacing the summation index $\mathbf{k}$ by $\boldsymbol{\kappa}=\mathbf{k}+\mathbf{K}/2$, we get relation (80).

γ. Commutation relations of pair field operators

For the same reasons as explained above (commutation of any products of two annihilation operators), the operators $\Phi_{\chi_n}(\mathbf{R})$ all commute with each others; the same is true for the adjoint operators $\Phi^\dagger_{\chi_n}(\mathbf{R})$. We have yet to examine the relations between the $\Phi_{\chi_n}(\mathbf{R})$ and the $\Phi^\dagger_{\chi_{n'}}(\mathbf{R}')$. Relations (11) and (12) show that:

$$\begin{aligned}\left[\Phi_{\chi_n}(\mathbf{R}),\Phi^\dagger_{\chi_{n'}}(\mathbf{R}')\right] = \\ \frac{1}{2}\int \mathrm{d}^3x\ \chi^*_n(\mathbf{x})\int \mathrm{d}^3x'\ \chi_{n'}(\mathbf{x}') \\ \left[\Psi(\mathbf{R}-\frac{\mathbf{x}}{2})\Psi(\mathbf{R}+\frac{\mathbf{x}}{2}),\Psi^\dagger(\mathbf{R}'+\frac{\mathbf{x}'}{2})\Psi^\dagger(\mathbf{R}'-\frac{\mathbf{x}'}{2})\right]\end{aligned} \tag{90}$$

The computation will not be carried out explicitly (though it does not present any particular difficulty); it leads to:

$$\begin{aligned}&\left[\Phi_{\chi_n}(\mathbf{R}),\Phi^\dagger_{\chi_{n'}}(\mathbf{R}')\right] \\ &\quad=\delta(\mathbf{R}-\mathbf{R}')\ \delta_{n,n'} \\ &\qquad+16\int \mathrm{d}^3x\ \chi^*_n[\mathbf{x}]\,\chi_{n'}[2(\mathbf{R}'-\mathbf{R})-\mathbf{x}]\,\Psi^\dagger(2\mathbf{R}'-\mathbf{R}-\frac{\mathbf{x}}{2})\Psi(\mathbf{R}-\frac{\mathbf{x}}{2}) \\ &\quad=\delta(\mathbf{R}-\mathbf{R}')\ \delta_{n,n'} \\ &\qquad+16\int \mathrm{d}^3z\ \chi^*_n(\mathbf{R}'-\mathbf{R}-\mathbf{z})\,\chi_{n'}(\mathbf{R}'-\mathbf{R}+\mathbf{z}) \\ &\qquad\times\Psi^\dagger(\frac{\mathbf{R}+\mathbf{R}'+\mathbf{z}}{2}+\mathbf{R}'-\mathbf{R})\ \Psi(\frac{\mathbf{R}+\mathbf{R}'+\mathbf{z}}{2}+\mathbf{R}-\mathbf{R}')\end{aligned} \tag{91}$$

In the second more symmetrical form of this commutator, we have used the notation $\mathbf{z}=\mathbf{R}'-\mathbf{R}-\mathbf{x}$. These relations are the equivalent, in the position representation, of the commutation relations (80) in the momentum representation (as already mentioned, it is possible to make the factor η appear or disappear in front of the integral by changing the sign of the variable of one of the two functions χ^*_n or $\chi_{n'}$).

The commutator thus includes several terms. The first term in $\delta(\mathbf{R}-\mathbf{R}')\,\delta_{n,n'}$ corresponds to the commutation relation of a usual bosonic field (whether the pair constituents are bosons or fermions); the $\delta_{n,n'}$ reflects the commutation of field components

corresponding to different orbital internal states of the pairs. To this term must be added a correction that depends on the structure of the pair, characterized by the functions $\chi_n(\mathbf{r})$ and $\chi_{n'}(\mathbf{r})$. We again find a result similar to that obtained before: a first simple bosonic term, which only takes into account the simultaneous exchanges of the two constituents of a "molecule" with the two constituents of another one. This term is followed by a correction that comes from the possibility of exchanges other than those involving complete pairs. Note that this correction is expressed in terms of field operators of the elementary constituents themselves and not of the pairs, as was to be expected since it is the constituents themselves that are involved. This correction term is a one-particle operator, non-diagonal in the position representation, since it destroys a particle at point $\mathbf{r}$ and recreates another one at point $\mathbf{r} + \mathbf{2}(\mathbf{R}' - \mathbf{R})$, always at the same distance.

To keep things simple, let us assume the dimensions of the "molecules" that define the pair field for the two internal states n and n' we are concerned with, are both of the order of the same dimension a_0; this means that the wave functions $\chi_n(\mathbf{r})$ and $\chi_{n'}(\mathbf{r})$ go to zero when $r \gg a_0$. In relation (91), the values of $\mathbf{x}$ that contribute to the integral are those for which none of the two functions $\chi_n^*[\mathbf{x}]$ and $\chi_{n'}[2(\mathbf{R} - \mathbf{R}') - \mathbf{x}]$ takes a negligible value; this requires that neither $|\mathbf{x}|$, nor $|2(\mathbf{R} - \mathbf{R}') - \mathbf{x}|$, be large compared to a_0. This double condition imposes $|\mathbf{R} - \mathbf{R}'| \simeq a_0$, in which case there are values of $\mathbf{x}$ for which both functions take simultaneously large values and the correction to the commutator cannot be neglected. On the other hand, if $|\mathbf{R} - \mathbf{R}'| \gg a_0$, there is no common domain where both functions χ_n^* and $\chi_{n'}$ take on significant values and the integral over d^3z is practically negligible. In other words, the molecular wave function range a_0 also plays the role of the commutator correction range.

The limit $a_0 \to 0$ can be obtained by choosing functions χ proportional to a function δ^ε (Appendix II, § 1-b), whose width goes to zero as $\varepsilon \to 0$ and whose integral equals one (it takes values of the order of $1/\varepsilon^3$ in a domain of volume of the order of ε^3). For the sake of simplicity, we assume that $n' = n$; as it is the square of the function χ_n that is normalized to 1 (and not the function itself), we must choose:

$$\chi_n(\mathbf{r}) \simeq \varepsilon^{3/2}\ \delta^\varepsilon(\mathbf{r}) \tag{92}$$

Using this form for the functions χ, the integral over d^3z in (91) leads to the convolution of two delta functions, which yields a function $\delta(\mathbf{R} - \mathbf{R}')$ multiplied by the operator $\Psi^\dagger(\mathbf{R})\Psi(\mathbf{R})$; nevertheless, the coefficient ε^3 of this term yields zero in the limit where $\varepsilon \to 0$. Consequently, if the molecules' size is very small compared to all the characteristic lengths of the system (such as the distance between molecules), the commutation relations of the field operator are exactly the same as for fields associated with bosons.

In conclusion, when the "molecules" have no spatial overlap[5], the only relevant exchanges concern exchanges between both of their constituents. On the other hand, when the two molecules do overlap, individual exchanges between their constituents become possible. If the molecules are loosely bound, as in the example of the BCS fermion pairing mechanism (Complement $\mathrm{C_{XVII}}$), they cannot be treated as bosons without structure, and one must use the complete formula (91) for the commutator.

[5]This does not exclude the case where the distance between molecules is small or comparable to the de Broglie wavelength of their centers of mass: the gas of molecules may be degenerate.

3-b. Singlet pairs

We now study the case of particles in a singlet pair, as in § 1-b.

α. Commutation relations of the $A_{\mathbf{K},n}$

As before, any products of two creation operators commute with any products of two creation operators; the same is true for any products of two annihilation operators. Relations (78) are thus still valid. We now have to compute the commutator:

$$\left[A_{\mathbf{K},n}, A^{\dagger}_{\mathbf{K}',n'}\right] = \sum_{\mathbf{k}} \; (g^n_{\mathbf{k}})^* \sum_{\mathbf{k}'} \; g^{n'}_{\mathbf{k}'} \; \left[a_{\frac{\mathbf{K}}{2}-\mathbf{k},\downarrow} a_{\frac{\mathbf{K}}{2}+\mathbf{k},\uparrow} \; , \; a^{\dagger}_{\frac{\mathbf{K}'}{2}+\mathbf{k}',\uparrow} a^{\dagger}_{\frac{\mathbf{K}'}{2}-\mathbf{k}',\downarrow}\right] \tag{93}$$

We are going to show that:

$$\begin{aligned} &\left[A_{\mathbf{K},n}, A^{\dagger}_{\mathbf{K}',n'}\right] = \delta_{\mathbf{K}\mathbf{K}'} \; \delta_{n,n'} \\ &+\eta \textstyle\sum_{\mathbf{k}} (g^n_{\mathbf{k}})^* \; g^{n'}_{\frac{\mathbf{K}-\mathbf{K}'}{2}+\mathbf{k}} \left[a^{\dagger}_{\mathbf{K}'-(\mathbf{K}/2)-\mathbf{k},\downarrow} a_{\frac{\mathbf{K}}{2}-\mathbf{k},\downarrow} + a^{\dagger}_{\mathbf{K}'-(\mathbf{K}/2)-\mathbf{k},\uparrow} \; a_{\frac{\mathbf{K}}{2}-\mathbf{k},\uparrow}\right] \\ &= \delta_{\mathbf{K}\mathbf{K}'} \; \delta_{n,n'} + \eta \textstyle\sum_{\kappa} \left(g^n_{\boldsymbol{\kappa}-\frac{\mathbf{K}}{2}}\right)^* \; g^{n'}_{\boldsymbol{\kappa}-\frac{\mathbf{K}'}{2}} \left[a^{\dagger}_{\mathbf{K}'-\boldsymbol{\kappa},\downarrow} a_{\mathbf{K}-\boldsymbol{\kappa},\downarrow} + a^{\dagger}_{\mathbf{K}'-\boldsymbol{\kappa},\uparrow} \; a_{\mathbf{K}-\boldsymbol{\kappa},\uparrow}\right] \end{aligned} \tag{94}$$

(with, in the last line, the notation $\boldsymbol{\kappa} = \mathbf{k}+\mathbf{K}/2$). Here again we find that the commutator of the two operators $A_{\mathbf{K},n}$ and $A^{\dagger}_{\mathbf{K}',n'}$ includes, to begin, a purely bosonic term, followed by corrections containing operators in normal order (which go to zero in the limit of low occupation numbers); a correction must be added for each of the two spin states.

Demonstration

To prove (94), we again use relation (86). As the indices 1, 2, 3 and 4 represent all the quantum numbers associated with an individual state, they must now contain the spin indices; these are added to the momentum indices, which play the same role as in the previous calculation. It then follows that the states 1 and 3 are always orthogonal, as are the states 2 and 4; the only terms remaining on the right-hand side of (86) are the terms in δ_{23} and δ_{14}, so that:

$$\begin{aligned} &\left[A_{\mathbf{K},n}, A^{\dagger}_{\mathbf{K}',n'}\right] = \textstyle\sum_{\mathbf{k}} (g^n_{\mathbf{k}})^* \sum_{\mathbf{k}'} g^{n'}_{\mathbf{k}'} \left[\delta_{\mathbf{K}\mathbf{K}'} \; \delta_{\mathbf{k}\mathbf{k}'}\right. \\ &\quad \left.+\eta \left(\delta_{(\mathbf{K}-\mathbf{K}')/2,\mathbf{k}'-\mathbf{k}} \; a^{\dagger}_{\frac{\mathbf{K}'}{2}-\mathbf{k}',\downarrow} a_{\frac{\mathbf{K}}{2}-\mathbf{k},\downarrow} + \delta_{(\mathbf{K}-\mathbf{K}')/2,\mathbf{k}-\mathbf{k}'} \; a^{\dagger}_{\frac{\mathbf{K}'}{2}+\mathbf{k}',\uparrow} \; a_{\frac{\mathbf{K}}{2}+\mathbf{k},\uparrow}\right)\right] \end{aligned} \tag{95}$$

or else (since the basis of the functions $g^n_{\mathbf{k}}$ is orthonormal):

$$\begin{aligned} &\delta_{\mathbf{K}\mathbf{K}'} \; \delta_{n,n'} \\ &+ \eta \sum_{\mathbf{k}} (g^n_{\mathbf{k}})^* \left[\; g^{n'}_{\frac{\mathbf{K}-\mathbf{K}'}{2}+\mathbf{k}} a^{\dagger}_{\mathbf{K}'-(\mathbf{K}/2)-\mathbf{k},\downarrow} a_{\frac{\mathbf{K}}{2}-\mathbf{k},\downarrow}\right. \\ &\qquad\qquad \left. + g^{n'}_{-\frac{\mathbf{K}-\mathbf{K}'}{2}+\mathbf{k}} a^{\dagger}_{\mathbf{K}'-(\mathbf{K}/2)+\mathbf{k},\uparrow} \; a_{\frac{\mathbf{K}}{2}+\mathbf{k},\uparrow}\right] \end{aligned} \tag{96}$$

We now modify the second term in the bracket of the summation, to make it similar to the first one: as the functions $g^n_{\mathbf{k}}$ have a definite parity $(-\eta)$ with respect to $\mathbf{k}$, we can change the index signs of g and g^*, and the sign of the dummy index $\mathbf{k}$ of the summation. The only difference between the two terms is then the spin directions, and we therefore obtain (94).

β. Commutation relations of pair field operators

A similar calculation to the one that led to (91) permits obtaining:

$$\begin{aligned}\left[\Phi_{\chi}\left(\mathbf{R}\right),\Phi_{\chi}^{\dagger}\left(\mathbf{R}'\right)\right] &= \delta\left(\mathbf{R}-\mathbf{R}'\right)\ \delta_{n,n'} \\ &+ 8\eta\int \mathrm{d}^3x\ \chi_n^*\left[\mathbf{x}\right]\chi_{n'}\left[2\left(\mathbf{R}'-\mathbf{R}\right)-\mathbf{x}\right] \\ &\times\left[\Psi_{\downarrow}^{\dagger}(2\mathbf{R}'-\mathbf{R}-\frac{\mathbf{x}}{2})\Psi_{\downarrow}(\mathbf{R}-\frac{\mathbf{x}}{2})+\Psi_{\uparrow}^{\dagger}(2\mathbf{R}'-\mathbf{R}-\frac{\mathbf{x}}{2})\Psi_{\uparrow}(\mathbf{R}-\frac{\mathbf{x}}{2})\right] \qquad (97)\end{aligned}$$

(a more symmetrical form of the right-hand side can be obtained by again using the notation $\mathbf{z}=\mathbf{R}'-\mathbf{R}-\mathbf{x}$). The commutator is thus equal to that of elementary bosons plus a correction term. This latter term plays an important role over a distance $\mathbf{R}'-\mathbf{R}$, of the order of the range of the wave functions χ, and is the sum of contributions independent of the two spin states.

Conclusion

In conclusion, note that the pair field operator provides interesting insights concerning the physical properties of paired states in a many-body system. For a N-particle state, built from a two-particle wave function χ, it leads to a new wave function ϕ_{pair} when the particle indistinguishability is taken into account. In the framework of the BCS theory, we will see how this pair wave function allows characterizing the cooperative effects of pair interactions. Introducing an order parameter is also useful for showing the link between anomalous average values (which do not conserve particle number) and the normal average values. The results take on different forms for paired states of bosons or fermions. There is, however, a strong analogy between the two cases, which provides a unified framework for the study of different phenomena, such as Bose-Einstein condensation of particles or pairs.

Complement B_{XVII}

Average energy in a paired state

1 Using states that are not eigenstates of the total particle number 1869

1-a Computation of the average values 1870

1-b A good approximation 1870

2 Hamiltonian 1871

2-a Operator expression 1871

2-b Simplifications due to pairing 1874

3 Spin 1/2 fermions in a singlet state 1874

3-a Different contributions to the energy 1874

3-b Total energy 1880

4 Spinless bosons 1881

4-a Choice of the variational state 1881

4-b Different contributions to the energy 1882

4-c Total energy 1887

In Chapter XVII, the paired states were introduced in a general way, without specifying any particular form of the Hamiltonian. In order to use the paired states $|\Psi_{\text{paired}}\rangle$ in the framework of a variational method, i.e. to be able to minimize the average value of the energy of an N-particle system, we must compute the average value of the energy in these paired states; this is the purpose of this complement. We start (§ 1) by examining the consequences of the fact that these states are not eigenvectors of the total particle number operator $\widehat{N}$. In § 2, we clarify the notation and give the expression of the Hamiltonian $\widehat{H}$. We then deal successively with the fermion case (§ 3) and the boson case (§ 4). This second case is slightly more complicated since it requires the adjunction of a specific state to describe the condensate.

1. Using states that are not eigenstates of the total particle number

The paired states $|\Psi_{\text{paired}}\rangle$ are coherent superpositions of states containing different numbers of particles. One may wonder how the average values computed in such states can be relevant for a physical system where N has a fixed value. As we already mentioned in § D of Chapter XVII, this approach is correct for large values of the average particle number, provided the operators, whose average values we are computing, conserve the particle number (i.e. commute with the total particle number $\widehat{N}$, as is the case for the Hamiltonian operator $\widehat{H}$). We are going to show in more detail that when these conditions are met, the average values do not depend on the state vector's coherences between different N values; they can thus be obtained using the paired states.

1-a. Computation of the average values

The state $|\Psi_{\text{paired}}\rangle$ defined in (B-5) of Chapter XVII is a superposition of states $|\Psi_P\rangle$ where the particle number is exactly $N = 2P$:

$$|\Psi_{\text{paired}}\rangle = \sum_{P=0}^{\infty} \frac{1}{P!} |\Psi_P\rangle \tag{1}$$

As the matrix elements of the operator $\widehat{H}$ between eigenkets of $\widehat{N}$ corresponding to different eigenvalues are zero, we have:

$$\begin{aligned} \langle\Psi_{\text{paired}}| \widehat{H} |\Psi_{\text{paired}}\rangle &= \sum_{P=0}^{\infty} \left[\frac{1}{P!}\right]^2 \langle\Psi_P| \widehat{H} |\Psi_P\rangle \\ &= \sum_{P=0}^{\infty} \left[\frac{1}{P!}\right]^2 \langle\Psi_P |\Psi_P\rangle \ E_P \end{aligned} \tag{2}$$

where E_P is the energy average value in the state $|\Psi_P\rangle$:

$$E_P = \frac{\langle\Psi_P| \widehat{H} |\Psi_P\rangle}{\langle\Psi_P |\Psi_P\rangle} \tag{3}$$

Consequently, if we define the weight distribution $\mathcal{D}(P)$ as:

$$\mathcal{D}(P) = \left[\frac{1}{P!}\right]^2 \langle\Psi_P |\Psi_P\rangle \tag{4}$$

the diagonal element of $\widehat{H}$ in $|\Psi_{\text{paired}}\rangle$ is given by:

$$\langle\Psi_{\text{paired}}| \widehat{H} |\Psi_{\text{paired}}\rangle = \sum_{P=0}^{\infty} \mathcal{D}(P) \ E_P \tag{5}$$

The average value $\left\langle \widehat{H} \right\rangle$ is then obtained by dividing this expression by the square of the norm $\langle\Psi_{\text{paired}} |\Psi_{\text{paired}}\rangle$.

In a general way, the diagonal element in $|\Psi_{\text{paired}}\rangle$ of any operator $\widehat{A}$ that commutes with $\widehat{N}$ is given by a linear combination of the average values of this operator in the states $|\Psi_P\rangle$ with the weight distribution $\mathcal{D}(P)$. As an example, for any function $F\left(\widehat{N}\right)$ of the operator $\widehat{N}$, we can write:

$$\langle\Psi_{\text{app}}| F\left(\widehat{N}\right) |\Psi_{\text{app}}\rangle = \sum_{P=0}^{\infty} \mathcal{D}(P) \ F(2P) \tag{6}$$

1-b. A good approximation

For a system with a fixed $N = 2P$ particle number, we are trying to determine the eigenvalues E_P and the kets $|\Psi_P\rangle$; the most direct method would be to vary separately

each ket $|\Psi_P\rangle$ to optimize E_P. This would lead, however, to complicated calculations. It turns out to be much more practical to vary $|\Psi_{\text{paired}}\rangle$ and optimize the corresponding energy; this leads to nearly the same results for large particle numbers, as we now explain.

We saw in § C-2 of Chapter XVII that the particle number fluctuations in a state $|\Psi_{\text{paired}}\rangle$ are very small in relative value when N is large. This means that the distribution $\mathcal{D}(P)$ has a sharp peak around a certain value P_0 of P, which determines half the average value of the particle number. Now if the energies E_P are practically constant over the width of that distribution, the Hamiltonian diagonal matrix element (2) can be written:

$$\begin{aligned}\langle\Psi_{\text{paired}}|\,\widehat{H}\,|\Psi_{\text{paired}}\rangle &\simeq E_{P_0}\sum_{P=0}^{\infty}\left[\frac{1}{P!}\right]^2\langle\Psi_P\,|\Psi_P\rangle \\ &= E_{P_0}\,\langle\Psi_{\text{paired}}\,|\Psi_{\text{paired}}\rangle \end{aligned} \tag{7}$$

Making this diagonal matrix element stationary (keeping constant the norm of $|\Psi_{\text{paired}}\rangle$) is equivalent to making E_{P_0} stationary. The optimal value obtained for this matrix element, divided by the squared norm of $|\Psi_{\text{paired}}\rangle$, yields a good approximation of the energy E_{P_0} we are looking for. Once $|\Psi_{\text{paired}}\rangle$ has been optimized in this way, it can be projected onto the various subspaces with fixed particle numbers, and therefore obtain the $|\Psi_P\rangle$, corresponding to stationary states with fixed particle numbers. In the following complements, we shall use the paired states rather than the states $|\Psi_P\rangle$ with fixed particle number.

Comment:

In the following complements, rather than optimizing the average energy, it is the difference between this average energy and the average particle number multiplied by the chemical potential μ that we shall optimize. As the two operators $\widehat{H}$ and $\widehat{N}$ commute with the total particle number, the line of reasoning we just followed also applies to that case.

2. Hamiltonian

Consider a physical system composed of fermions or bosons, placed in a cubic box of edge length L.

2-a. Operator expression

The Hamiltonian $\widehat{H}$ is the same as the one used on several occasions, for example in Complement E_{XV} (but we assume here that there is no external potential):

$$\widehat{H} = \widehat{H}_0 + \widehat{W}_{\text{int}} \tag{8}$$

The operator $\widehat{H}_0$ is the sum of the kinetic energy operators $K_0(q)$ associated with each particle q:

$$\widehat{H}_0 = \sum_q K_0(q) = \sum_q \frac{\mathbf{P}^2(q)}{2m} \tag{9}$$

and $\widehat{W}_{\rm int}$ is the sum of the interaction energies between particles:

$$\widehat{W}_{\rm int} = \frac{1}{2}\sum_{q\neq q'=1}^{N} W_2(\mathbf{R}_q, \mathbf{R}_{q'}) \tag{10}$$

where $W_2(\mathbf{R}_q, \mathbf{R}_{q'})$ only depends on the difference $\mathbf{R}_q - \mathbf{R}_{q'}$ (translation invariance) and does not act on the spins.

We now express $\widehat{H}$ in terms of creation and annihilation operators, according to formulas established in Chapter XV. We use the basis of individual states $|u_{\mathbf{k},\nu}\rangle$, where $\mathbf{k}$ labels the momentum $\hbar\mathbf{k}$ of a plane wave that satisfies the periodic conditions in the box; the index ν labels the spin state of the particles, but if they are all in the same spin state, it can be omitted in what follows. We get:

$$\begin{aligned}\widehat{H} = &\sum_{\mathbf{k},\nu} e_k\ a^\dagger_{\mathbf{k},\nu} a_{\mathbf{k},\nu} + \\ &+\frac{1}{2}\sum_{\mathbf{k},\nu;\mathbf{k}',\nu';\mathbf{k}'';\mathbf{k}'''} \langle 1:\mathbf{k}'',\nu;\ 2:\mathbf{k}''',\nu'|\,\widehat{W}_2(\mathbf{R}_1-\mathbf{R}_2)\,|1:\ \mathbf{k},\nu;\ 2:\mathbf{k}',\nu'\rangle \\ &\qquad\qquad \times\ a^\dagger_{\mathbf{k}'',\nu} a^\dagger_{\mathbf{k}''',\nu'} a_{\mathbf{k}',\nu'} a_{\mathbf{k},\nu}\end{aligned} \tag{11}$$

with:

$$e_k = \frac{\hbar^2 k^2}{2m} \tag{12}$$

(since the interaction potential does not act on the spins, we were able to replace the spin index ν'' associated with $\mathbf{k}''$ by the index ν, as well as the index ν''' associated with $\mathbf{k}'''$ by the index ν'). The matrix elements of $\widehat{W}_2$ appearing in (11) can be written:

$$\int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2\ W_2(\mathbf{r}_1-\mathbf{r}_2)\frac{1}{L^6}\ e^{i(\mathbf{k}-\mathbf{k}'')\cdot\mathbf{r}_1} e^{i(\mathbf{k}'-\mathbf{k}''')\cdot\mathbf{r}_2} \tag{13}$$

We make the following change of variables: $\mathbf{R} = (\mathbf{r}_1+\mathbf{r}_2)/2$ and $\mathbf{r} = \mathbf{r}_1-\mathbf{r}_2$. The integral over d^3R of the exponential yields the Kronecker delta function:

$$\frac{1}{L^3}\int \mathrm{d}^3 R\ e^{i\left(\mathbf{k}+\mathbf{k}'-\mathbf{k}''-\mathbf{k}'''\right)\cdot\mathbf{R}} = \delta_{\mathbf{k}+\mathbf{k}',\mathbf{k}''+\mathbf{k}'''} \tag{14}$$

which enforces the conservation of the total momentum:

$$\mathbf{k}+\mathbf{k}' = \mathbf{k}''+\mathbf{k}''' \tag{15}$$

The integral over $\mathrm{d}^3 r$ introduces the Fourier transform $V_{\mathbf{q}}$ of the potential[1]:

$$V_{\mathbf{q}} = \frac{1}{L^3}\int \mathrm{d}^3 r\ e^{-i\mathbf{q}\cdot\mathbf{r}}\ W_2(\mathbf{r}) \tag{16}$$

with:

$$\mathbf{q} = \frac{(\mathbf{k}''-\mathbf{k}) - (\mathbf{k}'''-\mathbf{k}')}{2} \tag{17}$$

[1]The factor $1/L^3$ in (16) comes from the normalization of the plane waves $e^{i\mathbf{k}\cdot\mathbf{r}}/L^{3/2}$ in a cube of edge length L; it ensures the potential V has the dimension of an energy.

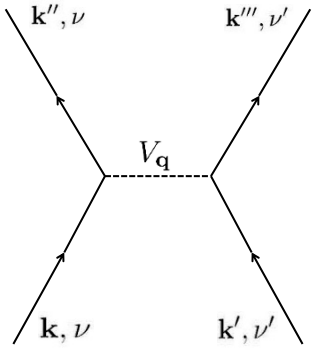


Figure 1: Symbolic plot of a general interaction process where two particles of momenta $\hbar\mathbf{k}$ and $\hbar\mathbf{k}'$ are replaced, as the result of their mutual interaction, by particles of momenta $\hbar\mathbf{k}''$ and $\hbar\mathbf{k}'''$. The indices ν and ν' label the spins, which are not modified by the interaction. The horizontal line represents the momentum transfer $\hbar\mathbf{q}$ whose value is given by (17) and (18).

or else, taking (15) into account:

$$\mathbf{q} = \mathbf{k}'' - \mathbf{k} = \mathbf{k}' - \mathbf{k}''' \tag{18}$$

The momentum transfer $\mathbf{q}$ gives the momentum variation of particle 1, as well as the opposite of the momentum variation of particle 2. Since $W_2(\mathbf{r}_1 - \mathbf{r}_2)$ is symmetric with respect to the exchange of the variables $\mathbf{r}_1$ and $\mathbf{r}_2$, the functions $W_2(\mathbf{r})$ and $V_{\mathbf{q}}$ are both even and real.

The matrix element of the interaction potential is then:

$$\langle 1 : \mathbf{k}'', \nu;\ 2 : \mathbf{k}''', \nu' | \, \widehat{W}_2(\mathbf{R}_1, \mathbf{R}_2) \, | 1 :\ \mathbf{k}, \nu;\ 2 : \mathbf{k}', \nu' \rangle = \delta_{\mathbf{k}+\mathbf{k}', \mathbf{k}''+\mathbf{k}'''}\ V_{\mathbf{q}} \tag{19}$$

and is schematized in Figure 1, where the horizontal line represents the momentum transfer $\hbar\mathbf{q}$ resulting from the interaction between the ingoing and outgoing particles. The interaction potential operator can thus be written:

$$\widehat{W}_{\text{int}} = \frac{1}{2} \sum_{\mathbf{k}, \mathbf{k}', \mathbf{k}'', \mathbf{k}'''}{}' V_{\mathbf{q}} \sum_{\nu, \nu'} a^{\dagger}_{\mathbf{k}'', \nu} a^{\dagger}_{\mathbf{k}''', \nu'} a_{\mathbf{k}', \nu'} a_{\mathbf{k}, \nu} \tag{20}$$

where the summation over the $\mathbf{k}$ actually concerns only three wave vectors, since $\mathbf{k}''' = \mathbf{k} + \mathbf{k}' - \mathbf{k}''$.

In a frequently used approximation, one assumes the interaction potential range to be very small compared with the de Broglie wavelengths of all the particles involved

(contact potential). The variations with $\mathbf{k}$ of $\overline{W}_2(\mathbf{k})$ can then be neglected, and all the matrix elements of the interaction potential are equal to a given constant $V_\mathbf{0}$ (provided they conserve the total momentum; otherwise, they are obviously zero):

$$V_\mathbf{0} = \frac{1}{L^3}\int \mathrm{d}^3 r\; W_2(\mathbf{r}) \tag{21}$$

2-b. Simplifications due to pairing

In general, the computation of the average value of the operator (11) is very complex, due to the large number of possible interaction terms. However, as we already saw in § D-1-a of Chapter XVII, some simplifications occur for a paired state. The main reason is that in the various components of a paired state on Fock states, all the paired individual states have the same population. If the population of an individual state $\mathbf{k}$ changes, the population of the individual state $-\mathbf{k}$ must change by the same quantity, otherwise the average value of the operator is zero. To get a non-zero average value in a paired state, the combination of creation and annihilation operators in the considered interaction term must respect this parity condition.

Now the interaction operator (20) is a sum of terms containing two annihilation operators on the right, and two creation operators on the left. Only two possibilities exist for the population balance of all the pairs to be conserved upon the action of these four operators: either the two creation operators re-establish the initial populations of the two states that were depopulated by the annihilation operators (in which case none of the populations are changed); or else, the two annihilation operators destroy particles in the same pair of states, and the creation operators produce another pair (in which case the population of the first pair[2] is lowered by 2, and the population of the second increased by 2). The two possibilities are combined in the particular case where the creation operators restore precisely the pair of particles destroyed by the annihilation operators. We are then led to the different cases examined in detail in § D-1-a of Chapter XVII: Case **I** (direct and exchange forward scattering terms), Case **II** (pair annihilation-creation terms) and Case **III** (combination of the two previous terms, yielding a negligible contribution).

3. Spin 1/2 fermions in a singlet state

We now compute the average value $\langle H\rangle$ of the operator $\widehat{H}$, written in (11), in the state $|\Psi_{\mathrm{BCS}}\rangle$ defined in § B-2-b of Chapter XVII. As far as the interaction energy is concerned, we will show that the terms associated with Case **I** only yield the usual mean field contributions, already discussed in the previous chapters. On the other hand, the terms associated with Case **II** are a direct consequence of the pairing, and are therefore totally new; they play a leading role in the BCS theory. The terms associated with Case **III**, being a particular case of the other two cases, generally play a negligible role.

3-a. Different contributions to the energy

The different contributions to the energy will be computed successively, starting with the kinetic energy.

[2] We defined in § C-2 of Chapter XVII, the pair population operator $\hat{n}_{\mathrm{pair}}$ as the sum of the population operators of each of the two individual states forming the pair.

α. Kinetic energy

The first term (kinetic energy) is, as for the particle number, the sum of the contributions of the pairs of states, labeled by $\mathbf{k}$ (each of the two states having the same kinetic energy):

$$\begin{aligned}\langle H_0\rangle &= \sum_{\mathbf{k}} e_k \ \langle\overline{\varphi}_{\mathbf{k}}|\,\widehat{n}_{(\text{pair }\mathbf{k})}\,|\overline{\varphi}_{\mathbf{k}}\rangle = 2\sum_{\mathbf{k}} e_k \ \left|v\left(\mathbf{k}\right)\right|^2 \\ &= 2\sum_{\mathbf{k}} e_k \ \sin^2\theta_{\mathbf{k}} \end{aligned} \tag{22}$$

β. Interaction energy

The average of the interaction potential energy is the sum of the averages of the terms on the right-hand side of (20), i.e. of terms that belong to one of the three possibilities **I**, **II** and **III** cited above; we study them successively.

– **Case I (the creation operators repopulate the states depopulated by the annihilation operators)**

For such terms, the occupation numbers of each individual state remain unchanged in the course of the interaction process. They are "diagonal terms" (sometimes called "mean field terms"). Two cases may arise, depending on whether the spin index ν is the same as, or different from ν'; we examine each of these possibilities in turn.

(i) If $\nu' = -\nu$, as the interaction potential does not act on the spin, we can trace each particle using its spin direction; it is as if the particles were distinguishable. If the creation operators repopulate exactly the same individual states depopulated by the annihilation operators, the only possible interaction is schematized in Figure 2, and corresponds to a *forward scattering*. As the momentum transfer $\mathbf{q}$ is zero, the potential term includes the constant $V_{\mathbf{0}}$, and we get the following contribution to the average energy:

$$\frac{V_{\mathbf{0}}}{2}\sum_{\mathbf{k}\neq-\mathbf{k}',\nu}\left\langle\overline{\Psi}_{\text{BCS}}\right|\left(a^{\dagger}_{\mathbf{k},\nu}a^{\dagger}_{\mathbf{k}',-\nu}a_{\mathbf{k}',-\nu}a_{\mathbf{k},\nu}\right)\left|\overline{\Psi}_{\text{BCS}}\right\rangle \tag{23}$$

(the condition $\mathbf{k}\neq-\mathbf{k}'$ comes from the fact that the pairs are different, each pair being labeled by the value of $\mathbf{k}$ associated with the spin $+$). Two anticommutations permit bringing the last operator $a_{\mathbf{k},\nu}$ right after the first one $a^{\dagger}_{\mathbf{k},\nu}$ (with two sign changes that cancel each other). If we now sum all the contributions from $\nu=+$ and from $\nu=-$, we get:

$$\begin{aligned}\frac{V_{\mathbf{0}}}{2}\sum_{\mathbf{k}\neq-\mathbf{k}'}&\left[\langle\overline{\varphi}_{\mathbf{k}}|\,a^{\dagger}_{\mathbf{k},\uparrow}a_{\mathbf{k},\uparrow}\,|\overline{\varphi}_{\mathbf{k}}\rangle\left\langle\overline{\varphi}_{-\mathbf{k}'}\right|a^{\dagger}_{\mathbf{k}',\downarrow}a_{\mathbf{k}',\downarrow}\left|\overline{\varphi}_{-\mathbf{k}'}\right\rangle\right.\\ &\left.+\left\langle\overline{\varphi}_{-\mathbf{k}}\right|a^{\dagger}_{\mathbf{k},\downarrow}a_{\mathbf{k},\downarrow}\left|\overline{\varphi}_{-\mathbf{k}}\right\rangle\langle\overline{\varphi}_{\mathbf{k}'}|\,a^{\dagger}_{\mathbf{k}',\uparrow}a_{\mathbf{k}',\uparrow}\,|\overline{\varphi}_{\mathbf{k}'}\rangle\right]\end{aligned} \tag{24}$$

We can show that the two terms inside the brackets yield the same contribution by interchanging the two dummy indices $\mathbf{k}$ and $\mathbf{k}'$ in the summation. We thus double the first term, and after changing the sign of $\mathbf{k}'$, we get:

$$\begin{aligned}V_{\mathbf{0}}\sum_{\mathbf{k}\neq\mathbf{k}'}\langle\overline{\varphi}_{\mathbf{k}}|\,a^{\dagger}_{\mathbf{k},\uparrow}a_{\mathbf{k},\uparrow}\,|\overline{\varphi}_{\mathbf{k}}\rangle\,\langle\overline{\varphi}_{\mathbf{k}'}|\,a^{\dagger}_{-\mathbf{k}',\downarrow}a_{-\mathbf{k}',\downarrow}\,|\overline{\varphi}_{\mathbf{k}'}\rangle &= V_{\mathbf{0}}\sum_{\mathbf{k}\neq\mathbf{k}'}\left|v\left(\mathbf{k}\right)\right|^2\left|v\left(\mathbf{k}'\right)\right|^2\\ &= V_{\mathbf{0}}\sum_{\mathbf{k}\neq\mathbf{k}'}\sin^2\theta_{\mathbf{k}}\ \sin^2\theta_{\mathbf{k}'}\end{aligned} \tag{25}$$

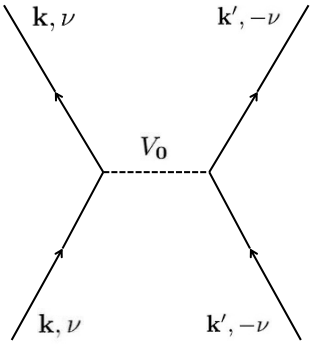


Figure 2: Schematic plot of the interaction between particles of opposite spins, which do not belong to the same pair (forward scattering). This diagram contributes to the particles' mean field.

When the particles are distributed in a large number of individual states, the value of the summation in the above expression is barely changed if we ignore the constraint $\mathbf{k} \neq \mathbf{k}'$. If we now use for $\langle N \rangle$ the expression (C-19) of Chapter XVII, we can write this contribution as:

$$\frac{V_0}{4} \langle N \rangle^2 \tag{26}$$

According to relation (21), the constant V_0 is proportional to the inverse of the volume L^3. This term can be interpreted as a mean field term, where $\langle N \rangle /2$ particles with a spin + interact with $\langle N \rangle /2$ particles having a spin −; a particle with a given spin direction feels the mean field exerted by all the particles with opposite spin, whose numerical density is $\langle N \rangle /2L^3$.

(ii) if $\nu' = \nu$, it is no longer possible to distinguish the particles by the direction of their spin, and the indistinguishability effects play their full role. Two cases must be distinguished for these "diagonal terms": either $\mathbf{k}'' = \mathbf{k}$ and $\mathbf{k}''' = \mathbf{k}'$, which yields a *direct term*; or $\mathbf{k}''' = \mathbf{k}$ and $\mathbf{k}'' = \mathbf{k}'$, which yields an *exchange term.* In both cases, the individual states populated in the bra and the ket are the same, and we are dealing with "diagonal processes" that can be called "mean field terms".

For the direct term, no particle changes its momentum, which again corresponds to a "forward scattering" (left-hand side of Figure 3), and the potential term again includes the constant V_0 written in (21). The average value of this direct term is:

$$\frac{V_0}{2} \sum_{\mathbf{k} \neq \mathbf{k}', \nu} \left\langle \overline{\Psi}_{\mathrm{BCS}} \right| a^\dagger_{\mathbf{k},\nu} a^\dagger_{\mathbf{k}',\nu} a_{\mathbf{k}',\nu} a_{\mathbf{k},\nu} \left| \overline{\Psi}_{\mathrm{BCS}} \right\rangle \tag{27}$$

Here again, since $\mathbf{k} \neq \mathbf{k}'$ (otherwise we would have the square of an annihilation operator, which is zero), two anticommutations let us bring the operator $a_{\mathbf{k},\nu}$ to the second position,

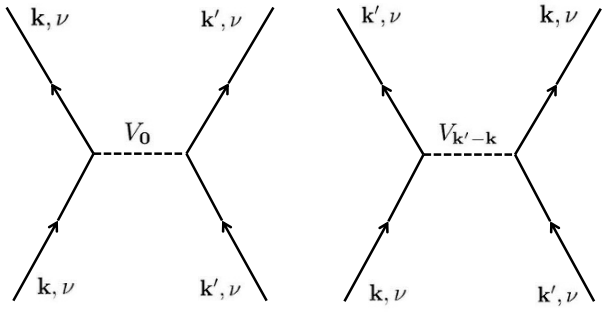


Figure 3: Interaction between particles having the same spin; the direct term (forward scattering) is schematized on the left, and the exchange term on the right. These two diagrams add their contributions to the diagram of Figure 2 to build the particles' mean field.

and we get:

$$\frac{V_0}{2}\sum_{\mathbf{k}\neq\mathbf{k}',\nu}\langle\overline{\varphi}_{\mathbf{k}}|\,a^{\dagger}_{\mathbf{k},\nu}a_{\mathbf{k},\nu}\,|\overline{\varphi}_{\mathbf{k}}\rangle\,\langle\overline{\varphi}_{\mathbf{k}'}|\,a^{\dagger}_{\mathbf{k}',\nu}a_{\mathbf{k}',\nu}\,|\overline{\varphi}_{\mathbf{k}'}\rangle = V_0\sum_{\mathbf{k}\neq\mathbf{k}'}|v_{\mathbf{k}}|^2\,|v_{\mathbf{k}'}|^2$$
$$= V_0\sum_{\mathbf{k}\neq\mathbf{k}'}\sin^2\theta_{\mathbf{k}}\;\sin^2\theta_{\mathbf{k}'} \tag{28}$$

(the two values of ν yield the same contribution, hence the disappearance of the factor $1/2$ on the right-hand side). As for the exchange term, we have $\mathbf{k}''' = \mathbf{k}$ and $\mathbf{k}'' = \mathbf{k}'$ (right-hand side of Figure 3); for such a momentum exchange, the transfer $\mathbf{q}$ is no longer zero, but equal to:

$$\mathbf{q} = \mathbf{k}'-\mathbf{k} \tag{29}$$

and the potential term now includes $V_{\mathbf{k}'-\mathbf{k}}$ obtained by inserting $\mathbf{q} = \mathbf{k}'-\mathbf{k}$ in (16). Furthermore, when $\mathbf{k} \neq \mathbf{k}'$:

$$a^{\dagger}_{\mathbf{k},\nu}a^{\dagger}_{\mathbf{k}',\nu}a_{\mathbf{k},\nu}a_{\mathbf{k}',\nu} = -a^{\dagger}_{\mathbf{k},\nu}a^{\dagger}_{\mathbf{k}',\nu}a_{\mathbf{k}',\nu}a_{\mathbf{k},\nu} \tag{30}$$

Apart from this sign change, the computation is the same as for the direct term. The sum of the two direct plus the exchange contributions finally yields:

$$\sum_{\mathbf{k}\neq\mathbf{k}'}\left[V_0 - V_{\mathbf{k}'-\mathbf{k}}\right]|v_{\mathbf{k}}|^2\,|v_{\mathbf{k}'}|^2 = \sum_{\mathbf{k}\neq\mathbf{k}'}\left[V_0 - V_{\mathbf{k}'-\mathbf{k}}\right]\;\sin^2\theta_{\mathbf{k}}\;\sin^2\theta_{\mathbf{k}'} \tag{31}$$

In the short-range potential approximation where $V_{\mathbf{k}'-\mathbf{k}} = V_0$, this sum is zero: the Pauli exclusion principle prevents particles having the same spin components from interacting via a contact potential.

– **Case II (particles annihilated in a pair of states and restored in another pair)**

Considering the nature of the creation and annihilation operators it contains, this process may be called "pair annihilation-creation". It plays an essential role in the BCS pairing, as we shall see in Complement C_{XVII}; the corresponding term in the Hamiltonian is thus often called the "pairing term".

We then have, on one side $\mathbf{k}' = -\mathbf{k}$ and $\nu' = -\nu$, and on the other, $\mathbf{k}''' = -\mathbf{k}''$, so that, according to its definition (17), the momentum transfer is $\mathbf{q} = \mathbf{k}'' - \mathbf{k}$; the corresponding diagram is shown in Figure 4. We are going to show below that its contribution to the energy can be written as:

$$\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}} \ \sin\theta_{\mathbf{k}}\cos\theta_{\mathbf{k}} \ \sin\theta_{\mathbf{k}''}\cos\theta_{\mathbf{k}''} \ e^{2i(\zeta_{\mathbf{k}}-\zeta_{\mathbf{k}''})} \tag{32}$$

This term is new, in the sense that it is not a mean field term, like the previous ones, but that its existence is due to the pairing process. We will show in Complement C_{XVII} that its contribution to the average energy plays an essential role in the BCS theory.

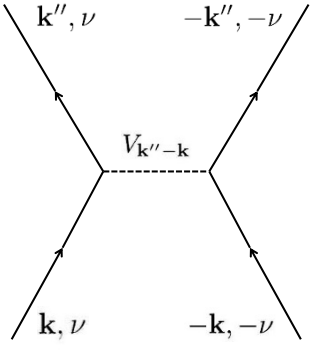


Figure 4: Interaction process between two particles in the same pair, which, in their final states, end up in another pair. In terms of creation and annihilation operators, this process is a "pair annihilation-creation" (two particles of the same pair are annihilated, while two particles are created in another pair). As opposed to the terms introduced by the other interaction processes, this term's contribution to the energy depends on the pairing; it is sometimes called the "pairing term", and is responsible for the energy gain in the BCS theory (Complement C_{XVII}).

Demonstration:

If $\nu = +$, the contribution contains a product of "anomalous" average values:

$$\frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}}\ \left\langle \overline{\Psi}_{\rm BCS}\right| a^{\dagger}_{\mathbf{k}'',\uparrow}a^{\dagger}_{-\mathbf{k}'',\downarrow}a_{-\mathbf{k},\downarrow}a_{\mathbf{k},\uparrow}\left|\overline{\Psi}_{\rm BCS}\right\rangle$$
$$= \frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}}\ \langle\overline{\varphi}_{\mathbf{k}''}|\, a^{\dagger}_{\mathbf{k}'',\uparrow}a^{\dagger}_{-\mathbf{k}'',\downarrow}\,|\overline{\varphi}_{\mathbf{k}''}\rangle\,\langle\overline{\varphi}_{\mathbf{k}}|\, a_{-\mathbf{k},\downarrow}a_{\mathbf{k},\uparrow}\,|\overline{\varphi}_{\mathbf{k}}\rangle \tag{33}$$

that is, using (C-42) and (C-44) of Chapter XVII:

$$\frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}}\ u_{\mathbf{k}''}v^{*}_{\mathbf{k}''}u^{*}_{\mathbf{k}}v_{\mathbf{k}}$$
$$= \frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}}\sin\theta_{\mathbf{k}}\cos\theta_{\mathbf{k}}\ \ \sin\theta_{\mathbf{k}''}\cos\theta_{\mathbf{k}''}\ e^{2i(\zeta_{\mathbf{k}}-\zeta_{\mathbf{k}''})} \tag{34}$$

If $\nu = -$, it is now the kets $\left|\overline{\varphi}_{-\mathbf{k}}\right\rangle$ and $\left|\overline{\varphi}_{-\mathbf{k}''}\right\rangle$ that come into play, and we obtain another product of anomalous average values for which we must use (C-43) and (C-45) of Chapter XVII (as well as the fact that the functions of $\mathbf{k}$ are even, as indicated in that chapter):

$$\frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}}\ \left\langle \overline{\Psi}_{\rm BCS}\right| a^{\dagger}_{\mathbf{k}'',\downarrow}a^{\dagger}_{-\mathbf{k}'',\uparrow}a_{-\mathbf{k},\uparrow}a_{\mathbf{k},\downarrow}\left|\overline{\Psi}_{\rm BCS}\right\rangle$$
$$= \frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}}\ v^{*}_{-\mathbf{k}''}u_{-\mathbf{k}''}v_{-\mathbf{k}}u^{*}_{-\mathbf{k}} \tag{35}$$

This expression is the same as the previous one, since it only differs by the sign of the summation dummy indices $\mathbf{k}$ and $\mathbf{k}''$ (remember that $V_{\mathbf{q}}$ is even). We therefore remove the factor 1/2 in (34) and get (32).

– **Case III (particles annihilated in a pair of states, then restored in the same pair)**

We then have again $\mathbf{k}' = -\,\mathbf{k}$ and $\nu' = -\nu$, but in addition $\mathbf{k}'' = \mathbf{k}$ (and hence necessarily $\mathbf{k}''' = \mathbf{k}'$), as shown in Figure 5; this is another case of forward scattering.

We now check that this term can be neglected. Its contribution to the energy is:

$$\frac{V_0}{2}\sum_{\mathbf{k},\nu}\ \left\langle \overline{\Psi}_{\rm BCS}\right| a^{\dagger}_{\mathbf{k},\nu}a^{\dagger}_{-\mathbf{k},-\nu}a_{-\mathbf{k},-\nu}a_{\mathbf{k},\nu}\left|\overline{\Psi}_{\rm BCS}\right\rangle \tag{36}$$

If $\nu = +$, we get (after two operator anticommutations):

$$\frac{V_0}{2}\sum_{\mathbf{k}}\langle\overline{\varphi}_{\mathbf{k}}|\left(a^{\dagger}_{\mathbf{k},\uparrow}a_{\mathbf{k},\uparrow}\right)\left(a^{\dagger}_{-\mathbf{k},\downarrow}a_{-\mathbf{k},\downarrow}\right)|\overline{\varphi}_{\mathbf{k}}\rangle = \frac{V_0}{2}\sum_{\mathbf{k}}|v_{\mathbf{k}}|^2 \tag{37}$$

and if $\nu = -$:

$$\frac{V_0}{2}\sum_{\mathbf{k}}\left\langle\overline{\varphi}_{-\mathbf{k}}\right|\left(a^{\dagger}_{\mathbf{k},\downarrow}a_{\mathbf{k},\downarrow}\right)\left(a^{\dagger}_{-\mathbf{k},\uparrow}a_{-\mathbf{k},\uparrow}\right)\left|\overline{\varphi}_{-\mathbf{k}}\right\rangle = \frac{V_0}{2}\sum_{\mathbf{k}}|v_{-\mathbf{k}}|^2 \tag{38}$$

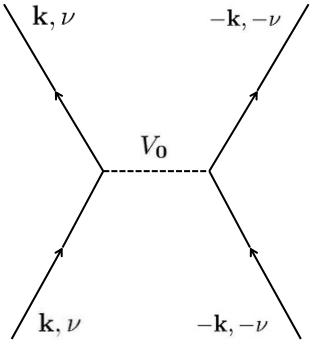


Figure 5: Interaction process where two particles of the same pair are scattered in the forward direction.

This term is the same as the previous one, as it only differs by the sign of the summation dummy index $\mathbf{k}$. Taking into account expression (C-19) of Chapter XVII for $\left\langle \hat{N} \right\rangle$, we can write the total contribution as:

$$V_0 \sum_{\mathbf{k}} |v_{\mathbf{k}}|^2 = \frac{V_0}{2} \left\langle \hat{N} \right\rangle \tag{39}$$

This contribution is interpreted as the average attraction energy in an ensemble of $\left\langle \hat{N} \right\rangle /2$ pairs. When the average particle number is large, we can neglect (39) compared to (26). Consequently, the pairing effects we are going to discuss cannot be simply interpreted as an attraction among an ensemble of $N/2$ pairs.

3-b. Total energy

Finally, adding the terms (22), (31), (26) and the double of (34), we get the average energy[3]:

$$\langle H \rangle = 2 \sum_{\mathbf{k}} e_k \ \sin^2 \theta_{\mathbf{k}} + \frac{V_0}{4} \langle N \rangle^2 + \sum_{\mathbf{k} \neq \mathbf{k}'} [V_0 - V_{\mathbf{k}'-\mathbf{k}}] \sin^2 \theta_{\mathbf{k}} \sin^2 \theta_{\mathbf{k}'} + \sum_{\mathbf{k},\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}} \ \sin \theta_{\mathbf{k}} \sin \theta_{\mathbf{k}''} \cos \theta_{\mathbf{k}} \cos \theta_{\mathbf{k}''} e^{2i(\zeta_{\mathbf{k}} - \zeta_{\mathbf{k}''})} \tag{40}$$

[3]The summations over $\mathbf{k}$ have no restrictions, contrary to the tensor product appearing in relation (B-8) of Chapter XVII, where the summation is limited to a half-space to avoid redundancy.

The first term on the right-hand side corresponds to the kinetic energy, the second to the mean field for particles of opposite spins, the third one is the analogous term for particles having the same spin (it goes to zero for a short-range potential); these three terms were already present in the Hartree-Fock theory. The fourth term, however, is new: it corresponds to the pair annihilation-creation (pairing term) whose average value is non-zero only in a paired state. It is the only one that depends on the phases $\zeta_{\mathbf{k}}$, which will prove to be essential in the BCS theory (Complement C_{XVII}).

4. Spinless bosons

For bosons, we must take into account the Bose-Einstein condensation phenomenon (Complements B_{XV}, C_{XV} and F_{XV}): in the ground state, a large fraction of the particles can occupy a single quantum state, the state $\mathbf{k} = 0$. This is not the case for a paired state; we must therefore choose a variational state permitting such a condensation.

We assume the interactions to be repulsive, in order to avoid the instabilities occurring for a system of attractive bosons (Complement H_{XV}, § 4-b).

4-a. Choice of the variational state

In Complement C_{XV}, we used the Gross-Pitaevskii approximation to treat, in the simplest way, Bose-Einstein condensation: the system of N bosons is supposed to be, at a given instant, in a state that is the product of identical individual states, generally chosen as the zero momentum state, $\mathbf{k} = 0$; the system state is thus written as:

$$|\Phi\rangle = \left[a_0^\dagger\right]^N |0\rangle \tag{41}$$

($a_0^\dagger$ is the creation operator in the individual state $\mathbf{k} = 0$). However, whereas such a state is suitable for an ideal gas ground state, it can only be an approximation for a gas of interacting particles: it is an eigenvector of the kinetic energy, but not of the operator associated with the interaction energy. The interaction potential actually couples this state to all the states where two particles are transferred from the individual state $\mathbf{k} = 0$ toward any two individual states of opposite momenta $\hbar\mathbf{k}$ and $-\hbar\mathbf{k}$ (because of momentum conservation), such as, for example, the state:

$$|\Phi'\rangle = \left[a_{\mathbf{k}}^\dagger\right]\left[a_{-\mathbf{k}}^\dagger\right]\left[a_0^\dagger\right]^{N-2} |0\rangle \tag{42}$$

where two states of a pair are occupied. This suggests using a state $|\Psi_{\text{paired}}\rangle$ as a variational ket[4] for describing the components of the system state vector associated with all the individual states $\mathbf{k} \neq 0$. We must also include the components corresponding to the individual state $\mathbf{k} = 0$; those will be described[5] by a "coherent state" (Complement G_V).

[4] The interaction potential also couples a state such as (42) to numerous states of the form $\left[a_{\mathbf{k}+\mathbf{q}}^\dagger\right]\left[a_{-\mathbf{q}}^\dagger\right]\left[a_{-\mathbf{k}}^\dagger\right]\left[a_0^\dagger\right]^{N-3}|0\rangle$, where $\mathbf{q}$ can take on any value. An exact theory would require taking those "unpaired" states into account, but leads to complex calculations. This is why we limit ourselves to a variational method in the framework of an approximation where the states $\mathbf{k} \neq 0$ are only accessible to pairs (we assume $N - N_0 \ll N$).

[5] This individual state must be treated separately, as applying the general formula (B-9) of Chapter XVII, used when $\mathbf{k} \neq \mathbf{0}$, to obtain $|\varphi_{\mathbf{k}=\mathbf{0}}\rangle$ would involve the exponential of the square of the operator

We therefore choose a variational state vector of the form:

$$|\Phi_B\rangle = |\varphi_{\mathbf{0}}\rangle \otimes |\Psi_{\text{paired}}\rangle = |\varphi_{\mathbf{0}}\rangle \otimes \prod_{\otimes \mathbf{k}\in D} |\overline{\varphi}_{\mathbf{k}}\rangle \tag{43}$$

(the notation B refers to the name Bogolubov).

In this expression, $|\Psi_{\text{paired}}\rangle$ is the paired state for spinless particles (B-8) of Chapter XVII, a tensor product of the normalized states $|\overline{\varphi}_{\mathbf{k}}\rangle$ defined in (B-9) and (C-13). The domain D of the tensor product in (43) is half the $\mathbf{k}$-space to avoid (as seen previously) a double appearance of each state $|\overline{\varphi}_{\mathbf{k}}\rangle$; the origin $\mathbf{k} = 0$ is excluded from D. This domain could eventually have an upper bound for $\mathbf{k}$.

As for $|\varphi_{\mathbf{0}}\rangle$, it is the coherent state whose expression can be found, for example, in Complement G_V, whose relations (65) and (66) provide[6]:

$$|\varphi_{\mathbf{0}}\rangle = e^{-N_0/2}\ e^{\alpha_0 a_0^\dagger}\,|n_0 = 0\rangle \tag{44}$$

This state depends on a complex parameter α_0, characterized by its modulus $\sqrt{N_0}$ and its phase ζ_0:

$$\alpha_0 = \sqrt{N_0}e^{i\zeta_0} \tag{45}$$

It is a normalized eigenvector of the operator a_0 with the eigenvalue α_0:

$$a_0\,|\varphi_{\mathbf{0}}\rangle = \alpha_0\,|\varphi_{\mathbf{0}}\rangle \tag{46}$$

The average particle number in the state $\mathbf{k} = 0$ is thus:

$$\langle\varphi_{\mathbf{0}}|\,a_0^\dagger a_0\,|\varphi_0\rangle = \alpha_0^*\alpha_0\,\langle\varphi_{\mathbf{0}}\,|\varphi_{\mathbf{0}}\rangle = N_0 \tag{47}$$

The width of the corresponding distribution is $\sqrt{N_0}$ (Complement G_V), hence negligible compared to N_0 (supposed to be a large number).

The variational variables contained in the trial ket (43) are thus the set of $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$, as well as N_0 and ζ_0.

4-b. Different contributions to the energy

We now compute the average energy, in the variational state $|\Phi_B\rangle$ written in (43), of the Hamiltonian operator given by (11).

$a^\dagger_{\mathbf{k}=\mathbf{0}}$, leading to large fluctuations of the particle number in the state $\mathbf{k} = 0$ (condensed particles). This would necessarily yield large fluctuations of the total particle number, as well as of the average repulsive energy, whereas, as we saw in § 3-b-β of Complement G_{XV}, those fluctuations are not possible precisely because of this repulsion.

[6] One shold be careful about the change of notation: in Chapter V and its complements, $|\varphi_0\rangle$ denotes the ground state of the one particle harmonic oscillator, which here corresponds to the vacuum $|n_0 = 0\rangle$. In the present complement, $|\varphi_{\mathbf{0}}\rangle$ is the ket associated with a large number of particles occupying the same individual state, as is also the case of the wave function $\varphi(\mathbf{r})$ of the Gross-Pitaevskii equation (Complement C_{XV}); with the notation of Complement G_V, this ket would rather correspond to a $|\alpha_0\rangle$ state.

α. Kinetic energy

The kinetic energy term is the sum of the contributions from the different individual states $\mathbf{k}$, with no contribution from the $\mathbf{k} = 0$ state (since $e_{k=0} = 0$). Each term of the summation contains the operator $\widehat{n}_{\mathbf{k}}$, whose average value in the factored (over the $\mathbf{k}$) state (43) is given by the average value $\langle\overline{\varphi}_{\mathbf{k}}|\,\widehat{n}_{\mathbf{k}}\,|\overline{\varphi}_{\mathbf{k}}\rangle$ in the state $|\overline{\varphi}_{\mathbf{k}}\rangle$. This average value is given by relation (C-33) of Chapter XVII as $\sinh^2\theta_{\mathbf{k}}$, which yields, for the average value E_c of the kinetic energy in the state $|\Phi_B\rangle$ the expression:

$$E_c = \sum_{\mathbf{k}\neq 0} e_k \;\langle\overline{\varphi}_{\mathbf{k}}|\,\widehat{n}_{\mathbf{k}}\,|\overline{\varphi}_{\mathbf{k}}\rangle = \sum_{\mathbf{k}\neq 0} e_k \,\sinh^2\theta_{\mathbf{k}} \tag{48}$$

with:

$$e_k = \frac{\hbar^2 k^2}{2m} \tag{49}$$

where m is the particle mass.

β. Interaction energy

The average value of the interaction potential energy is a sum over four indices $\mathbf{k}, \mathbf{k}', \mathbf{k}'', \mathbf{k}'''$ of the potential matrix elements described in (19). As opposed to what happened for the kinetic energy, these elements have no particular reason to cancel out if one (or several) of their indices is zero. We shall therefore compute the different contributions, arranging them in decreasing order of the number of their zero indices.

A noticeable simplification of the computation occurs with the choice of the trial vector, as the coherent state $|\varphi_{\mathbf{0}}\rangle$ is one of its factors. Any time one of the four indices in the potential energy term is zero, the corresponding annihilation operator may be replaced by the complex number α_0. This is because the trial ket $|\Phi_B\rangle$ is an eigenvector of the operator a_0 with eigenvalue α_0 - see relation (46). In the same way, each time one of the two indices $\mathbf{k}$ or $\mathbf{k}'$ is zero, the creation operators on the left of the product, and hence acting on the bra $\langle\Phi_B|$, can be replaced by α_0^*, since the Hermitian conjugate of relation (46) is:

$$\langle\varphi_{\mathbf{0}}|\,a_0^\dagger = \alpha_0^*\,\langle\varphi_{\mathbf{0}}| \tag{50}$$

These two operators are therefore simply replaced by numbers. Let us examine in turn all the possible cases.

(i) If the four indices $\mathbf{k}, \mathbf{k}', \mathbf{k}'', \mathbf{k}'''$ are zero, the interaction potential contributes via the constant $V_{\mathbf{0}}$ (forward scattering term), defined in (16) as the integral of the interaction potential $W_2(\mathbf{r})$; this contribution is written as:

$$\left\langle\widehat{W}_{\text{forward}}\right\rangle_{0,0} = \frac{V_{\mathbf{0}}}{2}\,\langle\varphi_{\mathbf{0}}|\,a_0^\dagger a_0^\dagger a_0 a_0\,|\varphi_{\mathbf{0}}\rangle = V_{\mathbf{0}}\frac{(N_0)^2}{2} \tag{51}$$

The corresponding term is represented in Figure 6.

There is no contribution from terms where three (and only three) indices $\mathbf{k}, \mathbf{k}', \mathbf{k}'', \mathbf{k}'''$ are zero: total momentum conservation would require the fourth index to also be zero.

(ii) If among the four indices $\mathbf{k}, \mathbf{k}', \mathbf{k}'', \mathbf{k}'''$ two are zero, one concerning an annihilation operator, the other a creation operator, the momentum conservation requires the

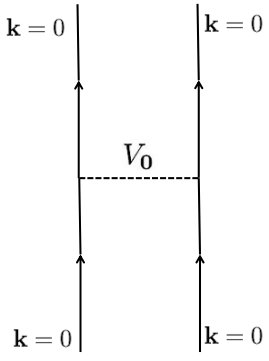


Figure 6: Diagram symbolizing the interactions between particles in the individual state $\mathbf{k}=0$*, and which remain in that state after the interaction (forward scattering term that yields the internal mean field of the condensate).*

other two operators to be $a_{\mathbf{k}}^{\dagger}$ and $a_{\mathbf{k}}$, with the same index $\mathbf{k}$. This can yield either a direct term, or an exchange term.

- the direct terms contain the average value of either the product $a_{\mathbf{k}}^{\dagger}a_0^{\dagger}a_0a_{\mathbf{k}}$, or of the product $a_0^{\dagger}a_{\mathbf{k}}^{\dagger}a_{\mathbf{k}}a_0$; it is again a forward scattering process and the potential appears via the constant $V_{\mathbf{0}}$. The two average values can be factored into two terms $\langle\varphi_{\mathbf{0}}|\, a_0^{\dagger}a_0\,|\varphi_{\mathbf{0}}\rangle = |\alpha_0|^2$ and $\langle\overline{\varphi}_{\mathbf{k}}|\,\widehat{n}_{\mathbf{k}}\,|\overline{\varphi}_{\mathbf{k}}\rangle = \sinh^2\theta_{\mathbf{k}}$; they are thus equal and the corresponding contribution is written as:

$$\left\langle \widehat{W}_{\text{direct}} \right\rangle_{0,e} = V_{\mathbf{0}}N_0 \sum_{\mathbf{k}\neq 0} \sinh^2\theta_{\mathbf{k}} \tag{52}$$

where the subscript e symbolizes the ensemble of the "excited" states, i.e. those with momentum $\hbar\mathbf{k} \neq 0$ that have a non-zero kinetic energy. Introducing the average total number of particles N_e in these excited states:

$$N_e = \sum_{\mathbf{k}\neq 0} \left\langle a_{\mathbf{k}}^{\dagger}a_{\mathbf{k}} \right\rangle = \sum_{\mathbf{k}\neq 0} \sinh^2\theta_{\mathbf{k}} \tag{53}$$

we can write:

$$\left\langle \widehat{W}_{\text{direct}} \right\rangle_{0,e} = V_{\mathbf{0}}N_0N_e \tag{54}$$

This term is simply interpreted as coming from the interaction between N_0 particles in the condensed state $\mathbf{k}=0$ and N_e particles in the other individual states.

- the exchange terms contain $a_{\mathbf{k}}^{\dagger}a_0^{\dagger}a_{\mathbf{k}}a_0$ and $a_0^{\dagger}a_{\mathbf{k}}^{\dagger}a_0a_{\mathbf{k}}$. We are now dealing with a momentum transfer process, and the potential now appears via the constant $V_{\mathbf{k}}$ obtained by inserting $\mathbf{q}=\mathbf{k}$ in (16). Otherwise, the computation is the same as for the direct

terms: the two average values can be factored, and the corresponding contribution is written:

$$\left\langle \widehat{W}_{\text{ex}} \right\rangle_{0,e} = N_0 \sum_{\mathbf{k}\neq 0} V_{\mathbf{k}} \, \sinh^2 \theta_{\mathbf{k}} \tag{55}$$

The two terms (54) and (55) correspond to mean field contributions associated with the interaction between $\mathbf{k} = 0$ particles and $\mathbf{k} \neq 0$ particles, taking into account the indistinguishability of the particles that led to the exchange term.

(iii) if the product of operators contains the annihilation operator in the state $\mathbf{k} = 0$ twice, the momentum conservation requires the product to be of the form $a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} a_0 a_0$. We are now dealing with a process where two particles in the state $\mathbf{k} = 0$ are replaced by a pair $(\mathbf{k}, -\mathbf{k})$, which amounts to creating a pair from particles initially in the condensate, as shown on the left-hand side of Figure 7; here again, the potential appears via the constant $V_{\mathbf{k}}$. The two annihilation operators introduce the factor $[\alpha_0]^2 = N_0 e^{2i\zeta_0}$ and the other two operators, an "anomalous average value" in a state $|\overline{\varphi}_{\mathbf{k}}\rangle$, which we already computed in (C-51) of Chapter XVII. We therefore get:

$$-\frac{N_0}{2} \sum_{\mathbf{k}\neq 0} V_{\mathbf{k}} \, \sinh\theta_{\mathbf{k}} \, \cosh\theta_{\mathbf{k}} \, e^{2i(\zeta_0 - \zeta_{\mathbf{k}})} \tag{56}$$

If the product of operators contains the creation operator in the state $\mathbf{k} = 0$ twice, it must necessarily be of the form $a^\dagger_0 a^\dagger_0 a_{\mathbf{k}} a_{-\mathbf{k}}$, which corresponds to the annihilation of a pair whose particles are transferred to the state $\mathbf{k} = 0$ (right-hand side of Figure 7). This product is the Hermitian conjugate of the previous one, and its average value is the complex conjugate of the previous result. The sum of these two terms is the contribution of the processes of creation and annihilation of pairs from the condensate:

$$-N_0 \sum_{\mathbf{k}\neq 0} V_{\mathbf{k}} \, \sinh\theta_{\mathbf{k}} \, \cosh\theta_{\mathbf{k}} \, \cos 2\,(\zeta_0 - \zeta_{\mathbf{k}}) \tag{57}$$

These terms come from the pairing of particles, as opposed to other terms that are related to the mean field. We shall see in Complement E_{XVII} the essential role they play in the Bose-Einstein condensation of an ensemble of bosons.

(iv) There are matrix elements of the interaction potential involving a single particle in the $\mathbf{k} = 0$ individual state, and three other particles in the $\mathbf{k} \neq 0$ states. The corresponding terms have a zero average value in the state $|\Phi_B\rangle$ because of the structure of its component $|\Psi_{\text{paired}}\rangle$, where the occupation numbers of two paired states must always vary together.

(v) We have yet to compute the contribution of terms where none of the wave vectors are zero. The computation is very similar to that of § 3, and we shall again distinguish three cases:

– **Case I**

Terms containing interactions where particles are created in the states from which they were destroyed: a direct term in $a^\dagger_{\mathbf{k}} a^\dagger_{\mathbf{k}'} a_{\mathbf{k}'} a_{\mathbf{k}}$, and an exchange term in $a^\dagger_{\mathbf{k}'} a^\dagger_{\mathbf{k}} a_{\mathbf{k}'} a_{\mathbf{k}}$. The computation is the same as in § 3, except for the fact that no minus sign occurs in

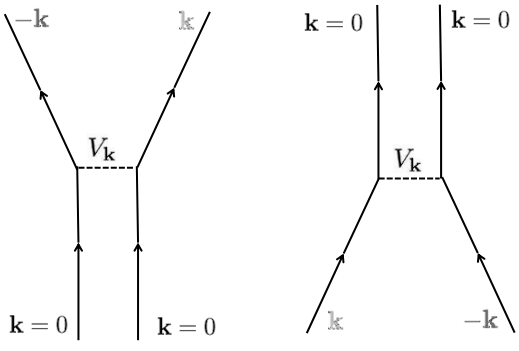


Figure 7: The diagram on the left represents a process where two particles, initially in the $\mathbf{k} = 0$ *individual state, interact and end up in states of opposite momenta. The diagram on the right represents the inverse process, where two particles of opposite momenta collide and end up in the* $\mathbf{k} = 0$ *individual state (i.e. in the condensate). As opposed to the previous terms, corresponding to the mean field of interacting particles, the terms corresponding to this diagram are introduced by the pairing process: they play a central role in the Bogolubov theory (Complement* E_{XVII}*).*

the exchange term. Result (31) therefore becomes, for bosons[7]:

$$\frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}'}\left[V_{\mathbf{0}}+V_{\mathbf{k}'-\mathbf{k}}\right]\left|v_{\mathbf{k}}\right|^2\left|v_{\mathbf{k}'}\right|^2=\frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}'}\left[V_{\mathbf{0}}+V_{\mathbf{k}'-\mathbf{k}}\right]\ \sinh^2\theta_{\mathbf{k}}\ \sinh^2\theta_{\mathbf{k}'}$$
$$\simeq\frac{V_{\mathbf{0}}}{2}\left(N_e\right)^2+\frac{1}{2}\sum_{\mathbf{k},\mathbf{k}'}V_{\mathbf{k}'-\mathbf{k}}\ \sinh^2\theta_{\mathbf{k}}\ \sinh^2\theta_{\mathbf{k}'}\qquad(58)$$

The first term is the direct term that, when $N_e \gg 1$, is interpreted as the effect of the interaction mean field between the $N_e\left(N_e-1\right)/2$ different pairs of particles (when $N_e \gg 1$). It is corrected by a second exchange term, which expresses the increased interaction between particles due to the boson bunching effect.

– **Case II**

The "pair annihilation-creation" term in $a^\dagger_{\mathbf{k}}a^\dagger_{-\mathbf{k}}a_{\mathbf{k}'}a_{-\mathbf{k}'}$, which yields here, taking

[7]As opposed to the fermion case, the contribution of the terms $\mathbf{k}=\mathbf{k}'$ is not zero but involves the average value of the operator $\widehat{n}_{\mathbf{k}}\left(\widehat{n}_{\mathbf{k}}-1\right)$ in the state $|\varphi_{\mathbf{k}}\rangle$, see § C-2-b of Chapter XVII. However, for a large system, the number of individual states $\mathbf{k}$ is very important, and this contribution is totally negligible compared to (58). This is why we neglected this term.

As for fermions, the summations over $\mathbf{k}$ do not have any restriction (no limitation to half the reciprocal space).

(C-51) and (C-52) of Chapter XVII into account:

$$\frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}}\ \langle\overline{\varphi}_{\mathbf{k}''}|\, a^{\dagger}_{\mathbf{k}''}a^{\dagger}_{-\mathbf{k}''}\,|\overline{\varphi}_{\mathbf{k}''}\rangle\,\langle\overline{\varphi}_{\mathbf{k}}|\, a_{-\mathbf{k}}a_{\mathbf{k}}\,|\overline{\varphi}_{\mathbf{k}}\rangle$$
$$= \frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{k}''} V_{\mathbf{k}''-\mathbf{k}}\ \sinh\theta_{\mathbf{k}}\cosh\theta_{\mathbf{k}}\ \sinh\theta_{\mathbf{k}''}\cosh\theta_{\mathbf{k}''}\ e^{2i(\zeta_{\mathbf{k}}-\zeta_{\mathbf{k}''})} \tag{59}$$

– **Case III**

Finally, the case where only one pair is involved leads, as in the fermion case, to a term proportional to $V_{\mathbf{0}}\langle N\rangle$, negligible compared to the term in $V_{\mathbf{0}}\langle N\rangle^2$ of (58) when the average particle number is large. We shall therefore neglect it.

4-c. Total energy

Regrouping the terms in $V_{\mathbf{0}}$ of (51), (54) and of (58), we get a total mean field term:

$$\left\langle \widehat{W}_{\text{mean field}}\right\rangle = \frac{V_{\mathbf{0}}}{2}\,(N_0+N_e)^2 \tag{60}$$

From a physical point of view, it is natural that this term be proportional to the square of the total particle number divided by 2, that is to the number of ways N particles can be associated by pairs (when $N \gg 1$). If we now include (55), (57) and (59), we get for the total energy:

$$\left\langle \widehat{H}\right\rangle_B = \sum_{\mathbf{k}\neq 0} e_k\ \sinh^2\theta_{\mathbf{k}} + \frac{V_{\mathbf{0}}}{2}\,(N_0+N_e)^2$$
$$+ N_0\sum_{\mathbf{k}\neq 0} V_{\mathbf{k}}\left[\sinh^2\theta_{\mathbf{k}} - \sinh\theta_{\mathbf{k}}\ \cosh\theta_{\mathbf{k}}\ \cos 2\,(\zeta_0-\zeta_{\mathbf{k}})\right]$$
$$+\frac{1}{2}\sum_{\mathbf{k},\mathbf{k}'\neq 0} V_{\mathbf{k}-\mathbf{k}'}\left[\sinh^2\theta_{\mathbf{k}}\ \sinh^2\theta_{\mathbf{k}'} + \sinh\theta_{\mathbf{k}}\ \sinh\theta_{\mathbf{k}'}\ \cosh\theta_{\mathbf{k}}\ \cosh\theta_{\mathbf{k}'}\cos 2\,(\zeta_{\mathbf{k}}-\zeta_{\mathbf{k}'})\right] \tag{61}$$

The second summation on the right-hand side describes the effect of momentum transfers between $\mathbf{k}\neq 0$ particles and the condensate, as well as the processes of annihilation and creation of pairs from the condensate (this last process depends on the relative phases $\zeta_0-\zeta_{\mathbf{k}}$ and, as already mentioned, arises from the pairing of particles). The last terms on the right-hand side, included in the double summation over $\mathbf{k}$ and $\mathbf{k}'$, correspond to interactions between particles in the $\mathbf{k}\neq 0$ states. Since the number of individual states is very large, we have ignored the constraint $\mathbf{k}\neq\mathbf{k}'$ of relation (59), which has a negligible effect; furthermore, as we noted in Chapter XVII, it is justified to replace the imaginary exponential by a cosine.

We have shown, in this complement, that the paired states are a useful tool for computing the average energy of an ensemble of interacting particles. In the following complements, we shall use these results successively for fermions and bosons.

Complement C_{XVII}

Fermion pairing, BCS theory

1 Optimization of the energy 1890
1-a Function to be optimized 1891
1-b Cancelling the total variation 1893
1-c Short-range potential, study of the gap 1896
2 Distribution functions, correlations 1899
2-a One-particle distribution 1899
2-b Two-particle distribution, internal pair wave function 1901
2-c Properties of the pair wave function, coherence length 1909
3 Physical discussion 1914
3-a Modification of the Fermi surface and phase locking 1914
3-b Gain in energy 1917
3-c Non-perturbative character of the BCS theory 1918
4 Excited states 1919
4-a Bogolubov-Valatin transformation 1919
4-b Broken pairs and excited pairs 1920
4-c Stationarity of the energies 1920
4-d Excitation energies 1922

We present in this complement the BCS mechanism for the pairing of fermions through attractive interactions. The three letters BCS refer to the names of J. Bardeen, L.N. Cooper and J.R. Schrieffer who proposed in 1957 [9] a theory for a physical phenomenon already observed in 1911 by H. Kamerlingh Onnes in Leiden, but as yet unexplained. This latter scientist observed that, below a certain temperature, the electrical resistivity of certain metals (mercury in his case) abruptly goes to zero as a phase transition occurs toward a so-called "superconducting" state. Along with this transition, many other spectacular effects occur, such as the expulsion of magnetic fields from the material. In this complement, we shall be concerned, with the general pairing mechanism of attractive fermions in the framework of BCS theory. We shall not, however, give any detail about the theory of metals, simply accepting the existence of an attraction between the fermions, without justifying its precise origin. In metals, this effective attraction comes from a coupling between electrons and phonons, and is therefore indirect, introducing an additional complexity to the problem. Furthermore, we shall not present any calculation of electrical resistivity, and hence not show that it can go to zero.

The BCS theory is a mean field theory, of the same type as the Hartree-Fock theory (Complements D_{XV} and E_{XV}). In this latter theory, particles are assumed to independently propagate in the mean field created by all the others; the system is described by an N-particle Fock state. Here, we shall assume that the particles form pairs, and this hypothesis will lead us to use, as a variation trial ket, the ket $|\Psi_{BCS}\rangle$ introduced in Chapter XVII; this complement is a direct application of the results of that chapter. The state

we will choose does indeed mathematically resemble a Fock state of "molecules", each composed of two particles. It should not be concluded, however, that this approximation reduces to a theory where each molecule is considered as an identifiable object moving in the mean field created by all the others. This naive picture is correct in the limit where the molecules are very strongly bound, but we shall see that it is totally inappropriate for loosely bound pairs such as those in the BCS theory. As already underlined in the introduction to Chapter XVII, the use of paired states brings a lot of flexibility to the mean field approach, as it allows modulating the binary correlation function between particles, and then to adapt it to interactions.

We start (§ 1) by minimizing the energy to determine the optimal quantum state in the family considered. In § 2, we discuss some physical properties of the optimized BCS wave function, mainly in terms of one- or two-particle correlation functions, but also in terms of what is called "non-diagonal order" (Complements A_{XVI} and A_{XVII}). Finally, in § 3, we shall study in more detail the physical content of the BCS pairing mechanism allowing the optimization of the energy of a fermion system, and in particular the role of phase locking (spontaneous symmetry breaking). For the sake of simplicity, we assume throughout this complement that the temperature is zero, but the BCS method can also be extended to the study of non-zero temperatures. This will lead to the study of excited states (§ 4), as will be briefly mentioned in § 4-d).

Shortly before the BCS theory was established, Cooper proposed a model including two attractive fermions. He showed that the exclusion of their wave functions from the interior of a Fermi sphere led to a bound state having certain properties similar to those described later by the complete BCS theory. This theory can be considered to be a generalization to N particles of the Cooper model, highlighting the collective effects leading to the properties of the BCS ground state. The Cooper model will be studied in Complement D_{XVII}, and its analogies with the N-particle theory will be underlined. In the present complement, we present the BCS theory, starting directly from the general results of Chapitre XVII; we shall also use the average energy values calculated in § 3 of Complement B_{XVII}.

We obviously cannot give here a detailed account of superconductivity theory and its various resulting effects, which would require an entire book. Limiting a large part of the computations to zero temperature situations already implies that numerous phenomena are outside the scope of this complement. To learn more about the subject, the reader can consult reference [8].

1. Optimization of the energy

Relation (B-11) of Chapter XVII yields the expression of the paired state[1] $|\Psi_{BCS}\rangle$:

$$|\Psi_{BCS}\rangle = \exp\left\{\sum_{\mathbf{k}} g_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right\}|0\rangle = \prod_{\otimes \mathbf{k}} |\varphi_{\mathbf{k}}\rangle \tag{1}$$

[1] This state is a superposition of components containing different numbers of particles. As already mentioned in Chapter XVII, one could also choose a variational state where the particle number is perfectly determined ([8], § 5-4 and Appendix C of Chap 5), but that would make the calculations a bit more complex.

The ket $|\Psi_{\rm BCS}\rangle$ was then normalized by separately normalizing each ket $|\varphi_{\mathbf{k}}\rangle$, which became the kets $|\overline{\varphi}_{\mathbf{k}}\rangle$:

$$|\overline{\varphi}_{\mathbf{k}}\rangle = \left[u_{\mathbf{k}} + v_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle \tag{2}$$

where the two functions $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are related by:

$$v_{\mathbf{k}} = u_{\mathbf{k}}\ g_{\mathbf{k}} \tag{3}$$

and satisfy:

$$|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2 = 1 \tag{4}$$

In that chapter, $g_{\mathbf{k}}$ was introduced as the Fourier transform of the wave function $\chi(\mathbf{r})$ of the "diatomic molecule" used to build the paired state; until now, this state was not specified. Here, we shall consider the $g_{\mathbf{k}}$ as variational parameters. Choosing $g_{\mathbf{k}} = 0$ leads to $v_{\mathbf{k}} = 0$ and $|u_{\mathbf{k}}| = 1$: in that case, the two individual states $\mathbf{k},\uparrow$ and $-\mathbf{k},\downarrow$ are neither occupied nor paired. They will be, however, if $g_{\mathbf{k}}$ is not zero. In general, the number of non-zero $g_{\mathbf{k}}$ is, a priori, arbitrary (finite or infinite). We can, for example, limit their number by setting a maximum value k_c for the modulus of $\mathbf{k}$, and consider this maximum value as a supplementary variational parameter defining the trial ket.

We were led, in that same Chapter XVII, to set $u_{\mathbf{k}} = \cos\theta_{\mathbf{k}}\ e^{-i\zeta_{\mathbf{k}}}$ and $v_{\mathbf{k}} = \sin\theta_{\mathbf{k}}\ e^{i\zeta_{\mathbf{k}}}$, relations that imply that $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ have opposite phases (a situation always possible to obtain by changing the global phase of the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$, which has no physical consequences). In the present complement, it will be more convenient to assume that the phase of $|\overline{\varphi}_{\mathbf{k}}\rangle$ is chosen in order to make $u_{\mathbf{k}}$ real and positive, and we will set:

$$\begin{aligned} u_{\mathbf{k}} &= \cos\theta_{\mathbf{k}} \\ v_{\mathbf{k}} &= \sin\theta_{\mathbf{k}}\ e^{2i\zeta_{\mathbf{k}}} \end{aligned} \tag{5}$$

Relation (C-19) of Chapter XVII yields the particle number in the state $\left|\overline{\Psi}_{\rm BCS}\right\rangle$:

$$\langle N\rangle = 2\sum_{\mathbf{k}} |v_{\mathbf{k}}|^2 = 2\sum_{\mathbf{k}} \sin^2\theta_{\mathbf{k}} \tag{6}$$

1-a. Function to be optimized

The average particle number in the state $|\Psi_{\rm BCS}\rangle$ may be changed by varying the $\mathbf{k}$ dependence of the $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$: as an example, choosing $|u_{\mathbf{k}}| = 1$ and $v_{\mathbf{k}} = 0$ for any value of $\mathbf{k}$, the average number $\langle N\rangle$ will be zero; on the other hand, if $|u_{\mathbf{k}}|$ is very small and $|v_{\mathbf{k}}|$ equals 1 for a great number of $\mathbf{k}$ values, the average total particle number $\langle N\rangle$ can attain arbitrarily large values. As the energy minimization operation makes sense only for a fixed value of $\langle N\rangle$, we shall determine that value with the Lagrangian multiplier μ (chemical potential; see Appendix VI, § 1-c). We will optimize the $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ choices by introducing the variations $\mathrm{d}u_{\mathbf{k}}$ and $\mathrm{d}v_{\mathbf{k}}$ and cancelling the variation of the average value $A = \langle H\rangle - \mu\langle N\rangle$. The volume L^3 of the physical system and its chemical potential μ are supposed to be fixed; we can choose one of two equivalent sets of variables to be determined, either the $u_{\mathbf{k}}$ and the $v_{\mathbf{k}}$, or the $\theta_{\mathbf{k}}$ and the $\zeta_{\mathbf{k}}$.

Relation (40) of Complement B_{XVII} yields $\langle H \rangle$, whereas $\langle N \rangle$ is given by (6). We then have:

$$\begin{aligned} A &= \langle H \rangle - \mu \langle N \rangle \\ &= 2 \sum_{\mathbf{k}} (e_k - \mu) \ \ |v_{\mathbf{k}}|^2 + \frac{V_{\mathbf{0}}}{4} \langle N \rangle^2 + \sum_{\mathbf{k} \neq \mathbf{k}'} [V_{\mathbf{0}} - V_{\mathbf{k}'-\mathbf{k}}] \, |v_{\mathbf{k}}|^2 \, |v_{\mathbf{k}'}|^2 \\ &\qquad + \sum_{\mathbf{k},\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \ v_{\mathbf{k}}^* \ u_{\mathbf{k}} \ v_{\mathbf{k}'} \ u_{\mathbf{k}'} \end{aligned} \tag{7}$$

with, according to (5):

$$|v_{\mathbf{k}}|^2 = \sin^2 \theta_{\mathbf{k}} \qquad ; \qquad v_{\mathbf{k}}^* \ u_{\mathbf{k}} = \sin \theta_{\mathbf{k}} \ \cos \theta_{\mathbf{k}} \ e^{-2i\zeta_{\mathbf{k}}} \tag{8}$$

In the above expression for A, e_k is the kinetic energy of a free particle with momentum $\hbar\mathbf{k}$:

$$e_k = \frac{\hbar^2 k^2}{2m} \tag{9}$$

and the $V_{\mathbf{k}}$ are the Fourier transforms of the interaction potential $W_2(\mathbf{r})$:

$$V_{\mathbf{k}} = \frac{1}{L^3} \int \mathrm{d}^3 r \ e^{-i\mathbf{k}\cdot\mathbf{r}} \ W_2(\mathbf{r}) \tag{10}$$

As this potential is rotationally invariant, the function $V_{\mathbf{k}}$ only depends on the modulus k of $\mathbf{k}$, and it must be real (Appendix I, § 2-e); as the potential is attractive, we can assume all the $V_{\mathbf{k}}$ to be negative.

We saw in Chapter XVII that the first term of (7) corresponds to the kinetic energy, the second to the mean field (diagonal term) for particles with opposite spins, the third one to the similar term for particles with identical spins (in which the direct and exchange terms cancel each other for a short-range potential). Finally, the fourth term, on the second line (which contains $V_{\mathbf{k}-\mathbf{k}'}$) plays a particularly important role in what follows; it comes from the pair annihilation-creation diagram schematized in Figure 4 of Complement B_{XVII}. It is often called the "pairing term".

We use relation (6) to write::

$$\mathrm{d}\left[\langle N \rangle\right]^2 = 2 \langle N \rangle \ \mathrm{d} \langle N \rangle = 4 \langle N \rangle \sum_{\mathbf{k}} \mathrm{d} \, |v_{\mathbf{k}}|^2 \tag{11}$$

We then get:

$$\begin{aligned} \mathrm{d}A &= 2 \sum_{\mathbf{k}} \xi_k \ \mathrm{d} \, |v_{\mathbf{k}}|^2 \\ &\qquad + \sum_{\mathbf{k},\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \left\{ v_{\mathbf{k}'} u_{\mathbf{k}'} \ \mathrm{d}\left[u_{\mathbf{k}} v_{\mathbf{k}}^*\right] + v_{\mathbf{k}'}^* u_{\mathbf{k}'} \ \mathrm{d}\left[u_{\mathbf{k}} v_{\mathbf{k}}\right] \right\} \end{aligned} \tag{12}$$

where the variable ξ_k is the kinetic energy with respect to the chemical potential, and corrected by the interaction effects [2]:

$$\xi_k = e_k - \mu + \frac{V_{\mathbf{0}}}{2} \langle N \rangle + \sum_{\mathbf{k}'} (V_{\mathbf{0}} - V_{\mathbf{k}'-\mathbf{k}}) \ \ |v_{\mathbf{k}'}|^2 \tag{13}$$

[2] A factor 2 appears in front of the summation over $\mathbf{k}'$ as the variations of $|v_{\mathbf{k}}|^2$ and $|v_{\mathbf{k}'}|^2$ in (7) must be added, but is included in the factor 2 in front of the summation containing ξ_k.

Comment:

For the applications of the BCS theory, the choice of the interaction potential to be used in the equations is not necessarily self-evident.

This is especially true in the superconductivity theory of metals, where the fermions involved are electrons which, isolated, interact via a repulsive Coulomb potential. In a metal, however, the direct repulsive interaction between electrons is mostly screened and they interact indirectly via the crystalline network deformations (phonons, see Complement J_V). This phenomenon leads to a long-range attractive component in their effective interaction, and explains why pairing between electrons is possible. This effective interaction depends on the phonon characteristics, and in particular on the Debye frequency of the solid under study.

This is also true in the theory of an ultra-cold diluted fermionic gas, where we do not use directly the interatomic potential in the equations. This interatomic potential contains, at short-range, a strongly repulsive part (often assimilated to a "hard core") and, at an intermediate distance, a strongly attractive well, permitting the formation of a large number of molecular states. Now when the gas under study is very dilute, the three-body collisions leading to these molecules are very rare, meaning these molecular states play practically no role; only the long distance effects of the potential have a real importance. In other words, the essential role is played by the asymptotic properties of the stationary collision states, as described by the scattering amplitude (Chapter VIII, relation B-9) and the associated phase shifts (Chapitre VIII, § C). The potential used in the BCS computations will therefore be an "effective potential". Furthermore, as the collisions occur at very low energy, this effective potential only depends on the phase shift δ_0 associated with $l = 0$. This phase shift is generally characterized by a "scattering length" a_0 defined as $\delta_0 \simeq -ka_0$ when $k \longrightarrow 0$; the effective potential will be attractive if this scattering length is negative.

As this complement deals mainly with the quantum mechanism for BCS pairing, and not with the determination of a valid potential, we shall not examine this point further and assume a pertinent choice of the interaction potential has been made.

1-b. Cancelling the total variation

It is obvious in (7) that the first three terms on the right-hand side depend only on the moduli of the $v_{\mathbf{k}}$; only the last term (annihilation-creation) depends on the phases $\zeta_{\mathbf{k}}$. Now the function A must be minimal if we vary the $\zeta_{\mathbf{k}}$ without changing the $\theta_{\mathbf{k}}$, i.e. when we vary only this last term:

$$\sum_{\mathbf{k},\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \ \sin\theta_{\mathbf{k}} \cos\theta_{\mathbf{k}} \ \sin\theta_{\mathbf{k}'} \cos\theta_{\mathbf{k}'} \ e^{2i(\zeta_{\mathbf{k}'}-\zeta_{\mathbf{k}})} \tag{14}$$

We assumed that all the potential matrix elements were negative, whereas relation (C-5) of Chapter XVII shows that the products $\sin\theta_{\mathbf{k}} \cos\theta_{\mathbf{k}}$ are positive. The lowest value of this sum will be obtained when all the terms in the summation over $\mathbf{k}$ and $\mathbf{k}'$ have the same phase in order to add coherently. This condition is called the "phase locking condition", and will be discussed in more detail in § 3-a-β. The minimum obtained does not depend on the absolute phase of the $v_{\mathbf{k}}$, but only on their relative phases. One can then simply choose all the $\zeta_{\mathbf{k}}$ to be equal to zero, which means that all the $v_{\mathbf{k}}$ are real and positive. This is the choice we shall adopt from now on.

In relation (12), the terms in $\mathrm{d}\left[u_{\mathbf{k}} v_{\mathbf{k}}^{*}\right]$ and $\mathrm{d}\left[u_{\mathbf{k}} v_{\mathbf{k}}\right]$ now become equal (the $\mathbf{k}$ and $\mathbf{k}'$ are dummy indices), and we have:

$$\mathrm{d}A = 2\sum_{\mathbf{k}} \xi_k \ \mathrm{d}\left(v_{\mathbf{k}}\right)^2 + 2\sum_{\mathbf{k},\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'} \ v_{\mathbf{k}'} u_{\mathbf{k}'} \ \mathrm{d}\left(u_{\mathbf{k}} v_{\mathbf{k}}\right) \tag{15}$$

We must now vary the $\theta_{\mathbf{k}}$. For that purpose, we introduce the quantities $\Delta_{\mathbf{k}}$, having the dimension of an energy, as:

$$\Delta_{\mathbf{k}} = \sum_{\mathbf{k}'} \left(-V_{\mathbf{k}-\mathbf{k}'}\right) \ u_{\mathbf{k}'} \ v_{\mathbf{k}'} \tag{16}$$

The $\Delta_{\mathbf{k}}$ are real since the $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are real, and positive since we assumed the interaction potential matrix elements to be negative. They are called "gaps" and play an important role in the BCS theory. It will be easier to discuss this role in the case of a very short-range potential; this will be done in § 1-c. The choice of the word "gap" will also be explained later (in § 4, see in particular Figure 7). The variation of A can now be written as:

$$\mathrm{d}A = 4\sum_{\mathbf{k}} \xi_k \ v_{\mathbf{k}} \ \mathrm{d}v_{\mathbf{k}} - 2\sum_{\mathbf{k}} \Delta_{\mathbf{k}} \ \left[u_{\mathbf{k}} \mathrm{d}v_{\mathbf{k}} + v_{\mathbf{k}} \mathrm{d}u_{\mathbf{k}}\right] \tag{17}$$

The variations of $\mathrm{d}u_{\mathbf{k}}$ and $\mathrm{d}v_{\mathbf{k}}$ are, however, not independent since relation (4) requires (for $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ real) that:

$$2u_{\mathbf{k}} \ \mathrm{d}u_{\mathbf{k}} + 2v_{\mathbf{k}} \ \mathrm{d}v_{\mathbf{k}} = 0 \tag{18}$$

This means we can replace $\mathrm{d}u_{\mathbf{k}}$ by $-v_{\mathbf{k}}\mathrm{d}v_{\mathbf{k}}/u_{\mathbf{k}}$. The right-hand side of (17) then becomes:

$$4\sum_{\mathbf{k}} \xi_k \ v_{\mathbf{k}} \ \mathrm{d}v_{\mathbf{k}} - 2\sum_{\mathbf{k}} \Delta_{\mathbf{k}} \ \left[u_{\mathbf{k}} - \frac{\left(v_{\mathbf{k}}\right)^2}{u_{\mathbf{k}}}\right] \mathrm{d}v_{\mathbf{k}} \tag{19}$$

Cancelling the variation of A with respect to all the $v_{\mathbf{k}}$ leads to:

$$2\xi_k v_{\mathbf{k}} - \Delta_{\mathbf{k}} \ \left[u_{\mathbf{k}} - \frac{\left(v_{\mathbf{k}}\right)^2}{u_{\mathbf{k}}}\right] = 0 \tag{20}$$

Multiplying by $u_{\mathbf{k}}$, we get:

$$2\xi_k \ u_{\mathbf{k}} v_{\mathbf{k}} = \Delta_{\mathbf{k}} \left[\left(u_{\mathbf{k}}\right)^2 - \left(v_{\mathbf{k}}\right)^2\right] \tag{21}$$

or else:

$$\xi_k \ \sin 2\theta_{\mathbf{k}} = \Delta_{\mathbf{k}} \ \cos 2\theta_{\mathbf{k}} \tag{22}$$

One can then compute the sine and the cosine, since:

$$\left[\cos 2\theta_{\mathbf{k}}\right]^2 = \frac{\left[\cos 2\theta_{\mathbf{k}}\right]^2}{\left[\cos 2\theta_{\mathbf{k}}\right]^2 + \left[\sin 2\theta_{\mathbf{k}}\right]^2} = \frac{\left(\xi_k\right)^2}{\left(\xi_k\right)^2 + \left(\Delta_{\mathbf{k}}\right)^2} \tag{23}$$

and we obtain:

$$\cos 2\theta_{\mathbf{k}} = \frac{\pm\xi_k}{\sqrt{(\xi_k)^2 + (\Delta_{\mathbf{k}})^2}} = \pm\frac{\xi_k}{E_k}$$
$$\sin 2\theta_{\mathbf{k}} = \frac{\pm\Delta_{\mathbf{k}}}{\sqrt{(\xi_k)^2 + (\Delta_{\mathbf{k}})^2}} = \pm\frac{\Delta_{\mathbf{k}}}{E_k} \tag{24}$$

where we have set:

$$E_k = \sqrt{(\xi_k)^2 + (\Delta_{\mathbf{k}})^2} \tag{25}$$

We finally obtain:

$$[u_{\mathbf{k}}]^2 = \frac{1}{2}\left[1 + \cos 2\theta_{\mathbf{k}}\right] = \frac{1}{2}\left[1 \pm \frac{\xi_k}{E_k}\right]$$
$$[v_{\mathbf{k}}]^2 = \frac{1}{2}\left[1 - \cos 2\theta_{\mathbf{k}}\right] = \frac{1}{2}\left[1 \mp \frac{\xi_k}{E_k}\right] \tag{26}$$

They are many possibilities for rendering stationary the difference of average values, $A = \langle H\rangle - \mu\langle N\rangle$, depending on the sign $\pm$ chosen in each equation, and for each value of $\mathbf{k}$. This multiplicity of solutions is not surprising since the stationarity is obtained not only for the ground state, but also for all the possible excited states of the physical system; those will be discussed in § 4, and we will see that the stationarity conditions (26) include the possibility of "excited pairs" (§ 4-c-β). For the moment, we shall concentrate on the search for the ground state and look for the absolute minimum of the average value (7).

Let us examine which signs must be chosen in (26) to get the ground state, i.e. the lowest possible value of A in expression (7). The chemical potential of an ideal fermion gas in its ground state is positive, equal to the Fermi level, and proportional to the particle density to the power 2/3 (Complement C_{XIV}, § 1-a). In the presence of a weak attractive potential, the factor $(e_k - \mu)$ in the first term on the right-hand side of (7) is negative when the modulus of $\mathbf{k}$ is small, and positive when $e_k \gg \mu$. In the first case, to minimize A, it is better to choose values of $[v_{\mathbf{k}}]^2$ as large as possible, and hence the $-$ sign in the second equation (26) since $\xi_{\mathbf{k}}$ is negative in this case. On the other hand, when $e_k \gg \mu$, it is better to choose values of $[v_{\mathbf{k}}]^2$ as small as possible, and it is again the $-$ sign in the second equality that must be chosen. As a result, it is the $-$ sign that must be taken in the second equality (26), and hence the $+$ sign in the first one.

As we know that $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are positive, we finally obtain for the ground state:

$$u_{\mathbf{k}} = \sqrt{\frac{1}{2}\left(1 + \frac{\xi_k}{E_k}\right)}$$
$$v_{\mathbf{k}} = \sqrt{\frac{1}{2}\left(1 - \frac{\xi_k}{E_k}\right)} \tag{27}$$

Inserting these results in (21), we verify that the stationarity relations are fulfilled, independently of the sign of ξ_k. They apply as long as the self-consistent condition derived

from the $\Delta_{\mathbf{k}}$ definition (16) is satisfied:

$$\Delta_{\mathbf{k}} = -\frac{1}{2}\sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\sqrt{1-\left(\frac{\xi_{k'}}{E_{k'}}\right)^2} = -\frac{1}{2}\sum_{\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\frac{\Delta_{\mathbf{k}'}}{\sqrt{(\xi_{k'})^2+(\Delta_{\mathbf{k}'})^2}} \tag{28}$$

As we now show, this condition takes on a simpler form for a very short-range interaction potential.

The above computation shows that starting from any function $\chi(\mathbf{r})$, or (which amounts to the same thing) from functions $g_{\mathbf{k}}$ considered to be entirely free variables, the optimization procedure yields values for $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$; this in turn fixes the optimal $g_{\mathbf{k}}$ and determines the function $\chi(\mathbf{r})$ for building the paired state described in (1).

1-c. Short-range potential, study of the gap

The matrix elements of the interaction potential V_k were defined in Complement B_{XVII} as the Fourier transforms of the potential – see relations (16) and (18) of that complement. For a regular potential of range b, the matrix element V_k necessarily varies when k changes by a quantity of the order of $1/b$; in particular, $V_k \to 0$ when $k \gg 1/b$. However, in many physical applications of the BCS theory, the k wave vectors involved remain very small compared to $1/b$, and a useful approximation is to ignore the variations of the V_k. We therefore consider them all equal to the same constant $-V$:

$$V_{\mathbf{k}} = -V \tag{29}$$

The minus sign was introduced to make V a positive number for an attractive potential; this number is inversely proportional to the volume $\mathcal{V}$, as shown by relation (10). Definition (13) of $\xi_{\mathbf{k}}$ now takes on a simplified form:

$$\xi_k = e_k - \mu - \frac{V}{2}\langle N\rangle \tag{30}$$

that can be inserted in the relations (27) to get the functions $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$.

Relation (16) also has a simpler version; all the $\Delta_{\mathbf{k}}$ take on the same value Δ:

$$\Delta = V\sum_{\mathbf{k}'} u_{\mathbf{k}'}v_{\mathbf{k}'} \tag{31}$$

In this case, there exists only one value of the gap, and since the $u_{\mathbf{k}'}$ and $v_{\mathbf{k}'}$ are real, this value is also real. We shall see in what follows that Δ plays a particularly important role, especially in the dispersion curve characterizing the system excitations (see for example Figure 7 which shows the existence of an energy minimum equal to Δ).

All the previous formulas apply, provided we replace the $\Delta_{\mathbf{k}}$ by Δ. Relations (24) then become, with the sign choice leading to the ground state:

$$\begin{aligned}\cos 2\theta_{\mathbf{k}} &= \frac{\xi_k}{\sqrt{(\xi_k)^2+\Delta^2}} = \frac{\xi_k}{E_k}\\ \sin 2\theta_{\mathbf{k}} &= \frac{\Delta}{\sqrt{(\xi_k)^2+\Delta^2}} = \frac{\Delta}{E_k}\end{aligned} \tag{32}$$

Equalities (27) are unchanged. When $k \to \infty$ the second relation (32) shows that, taking (25) and (30) into account, the value of $\theta_{\mathbf{k}}$ goes to zero as:

$$\theta_{\mathbf{k}} \sim \Delta/2E_k \sim \Delta/2\xi_k \tag{33}$$

It will be useful in what follows to know the asymptotic behaviors of the functions $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$, whose values as a function of $\theta_{\mathbf{k}}$ are given by (5). When $k \to \infty$, relation (33) shows that $u_{\mathbf{k}}$ goes to 1 whereas $v_{\mathbf{k}}$ goes to zero as:

$$v_{\mathbf{k}} \underset{k\to\infty}{\sim} \frac{\Delta}{2\xi_k} + 0(\frac{\Delta^2}{\xi_k^2}) \sim \frac{1}{k^2} + 0(\frac{1}{k^4}) \tag{34}$$

This ensures the convergence of the summations (C-19) and (C-26) of Chapter XVII giving the average values of $\widehat{N}$ and of its square.

α. Self-consistency condition and divergences

The self-consistent condition (28) now takes on the simpler form:

$$\Delta = \frac{V}{2}\sum_{\mathbf{k}} \sqrt{1 - \left(\frac{\xi_k}{E_k}\right)^2} = \frac{V}{2}\sum_{\mathbf{k}} \frac{\Delta}{\sqrt{(\xi_k)^2 + \Delta^2}} \tag{35}$$

that is:

$$1 = \frac{V}{2}\sum_{\mathbf{k}} \frac{1}{\sqrt{(\xi_k)^2 + \Delta^2}} \tag{36}$$

which is an implicit equation expressing the gap Δ in terms of μ (as this latter parameter appears in the definition of ξ_k).

Choosing a large volume L^3, we can replace the discrete summation over $\mathbf{k}$ by an integral. We shall assume, as mentioned in § 1, that the $g_{\mathbf{k}}$ in the variational ket (1) are zero when the modulus of $\mathbf{k}$ is larger than a given cutoff value k_c. Under these conditions, the implicit equation for the gap becomes:

$$1 = \frac{V}{2}\frac{L^3}{(2\pi)^3}\int_0^{k_c} \mathrm{d}^3k \frac{1}{\sqrt{(\xi_k)^2 + \Delta^2}} \tag{37}$$

where k_c is the upper limit of the wave vectors introduced in § 1; remember that V is inversely proportional to the volume, and hence the right-hand side of this equation does not depend on that volume. The integral will diverge if k_c is infinite, since when $k \to \infty$, the function to be integrated behaves as $1/\xi_k \sim 1/e_k \sim 1/k^2$. The value obtained for Δ therefore depends on the value chosen for k_c; this upper limit then plays an important role.

β. Calculation of the gap

We note e_c the equivalent, in terms of energy, of the cutoff frequency k_c:

$$e_c = \frac{\hbar^2 k_c^2}{2m} \tag{38}$$

We now choose the energy e_k as the integration variable. We then have to consider the density of states $D(e_k)$, obtained[3] by taking the differential of the definition of e_k:

$$D(e_k) = \frac{L^3}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\sqrt{e_k} \tag{39}$$

Relation (37) then becomes:

$$1 = \frac{V}{2}\int_0^{e_c} D(e_k)\ \mathrm{d}e_k \frac{1}{\sqrt{(e_k-\overline{\mu})^2+\Delta^2}} \tag{40}$$

where, to simplify relation (30), we introduced a chemical potential $\overline{\mu}$ relative to the mean field energy[4]:

$$\overline{\mu} = \mu + \frac{V}{2}\langle N\rangle \tag{41}$$

In relation (40), the function to be integrated over e_k contains a fraction that is maximum for $e_k = \overline{\mu}$; this fraction takes on significant values in an energy band of width Δ centered, in $\mathbf{k}$ space, on the surface of the "Fermi sphere" (see Complement C_{XIV}) whose radius k_F obeys:

$$\frac{\hbar^2 k_F^2}{2m} = \mu \tag{42}$$

As for the density of states $D(e_k)$, it takes on low values in the vicinity of the center of that sphere, but increasingly larger ones outside. The inside of the sphere barely contributes to the summation, the main contribution coming from the outside, in between the Fermi surface and the cutoff energy e_{k_c}. We can then find in this region an intermediate value D_0 for the density of states that can replace $D(e_k)$ without changing the integral, with:

$$D(\mu) \leq D_0 \leq D(e_c) \tag{43}$$

The density of states can be removed from the integral, and we get:

$$1 = \frac{VD_0}{2}\int_0^{e_c} \mathrm{d}e_k \frac{1}{\sqrt{(e_k-\overline{\mu})^2+\Delta^2}} \tag{44}$$

As V is inversely proportional to the volume, whereas, according to (39), the density of states is proportional to it, this relation is independent of the volume.

In the physics of superconducting metals, the attractive interaction between the electrons is mediated by the motions of the crystal's ions, i.e. by the phonons of the network. A cutoff energy naturally appears in the matrix elements of the interaction potential, the Debye energy $\hbar\omega_D$ of the phonons. One often uses a simple model where,

[3] This density of states is defined in Complement C_{XIV}, and given by formula (8) of that complement. In our case, as we do not have to take into account the two spin states, the density of states is half the one computed in that complement.

[4] With the sign convention we chose for V in (29), this mean field energy is equal to $-V\langle N\rangle/2$ per particle.

in (28), the potential matrix elements $V_{\mathbf{k}-\mathbf{k}'}$ are zero as soon as the difference in energy $|e_k - e_{k'}|$ is larger than $\hbar\omega_D$, supposed to be much smaller than $\overline{\mu}$; otherwise, they are all equal to a constant V. The same computations as those that led to (44), yield in this case the gap equation[5] for levels close to the Fermi surface:

$$1 \simeq \frac{VD_F}{2}\int_{\overline{\mu}-\hbar\omega_D}^{\overline{\mu}+\hbar\omega_D} \mathrm{d}e_k \frac{1}{\sqrt{(e_k-\overline{\mu})^2+\Delta^2}} = VD_F \ \mathrm{arsinh}\left(\frac{\hbar\omega_D}{\Delta}\right) \tag{45}$$

where D_F is the density of states on the Fermi surface. If, furthermore, we make the approximation $VD_F \ll 1$, we get:

$$\Delta = \frac{\hbar\omega_D}{\sinh\left(1/VD_F\right)} \simeq 2\hbar\omega_D \exp\left(-\frac{1}{VD_F}\right) \tag{46}$$

This important equation is called the "BCS gap equation".

It is worth noting that this expression cannot be expanded in a power series of V when the interaction goes to zero: all the derivatives of Δ with respect to V are zero for $V = 0$. Consequently, this expression cannot be obtained in the framework of a perturbation theory in powers of the interaction (this point will be discussed in § 3-c).

2. Distribution functions, correlations

Inserting expressions (27) in the trial ket, we obtain the optimal state vector $|\Psi_{\mathrm{BCS}}\rangle$ that best describes the ground state. We now examine the physical implications of that optimized quantum state, concerning the properties of the one- or two-particle distribution functions. These properties will be used later on in this complement to understand the origin of the energy lowering due to condensation into pairs of particles.

2-a. One-particle distribution

As we are going to show, the properties of the one-particle distribution are fairly close to those of an ideal gas.

α. Momentum space

Once the gap Δ is obtained, we can use relations (30) and (32) to determine the values of $\theta_{\mathbf{k}}$ for each value of $\mathbf{k}$; relation (C-16) of Chapter XVII then yields the average number of particles in each pair of states. As the two states composing the pair play the same role, the average number of particles in each of the states is simply half that number, that is $\sin^2\theta_{\mathbf{k}}$. Figure 1 plots, as a function of e_k, the variation of the distribution function $\langle n_{\mathbf{k},\nu}\rangle$ obtained, which is the momentum distribution function of a particle, once the variables $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ have been optimized. For an ideal gas, we saw in Complement $\mathrm{B_{XV}}$ that it is a Fermi-Dirac distribution; at zero temperature (as is the case here, since we are studying the ground state), this distribution is a "step function" equal to 1 for $e_k < \mu$ and to zero for $e_k > \mu$ (dotted curve). In our case, the transition

[5] There are two equal contributions to the integral on the right-hand side, one from the values of e_k above $\overline{\mu}$, the others from those below; this is why the factor 1/2 has disappeared from the second equality.

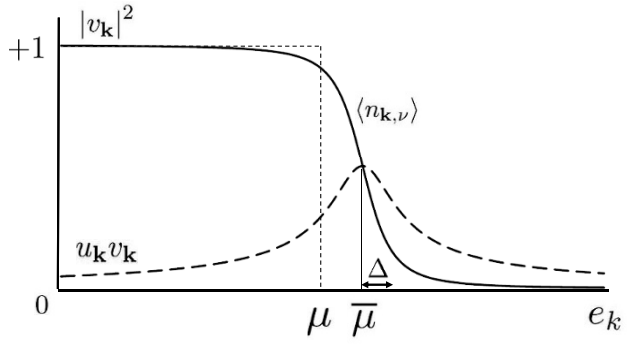


Figure 1: Plots of the one-particle distribution function $\langle n_{\mathbf{k},\nu} \rangle = |v_{\mathbf{k}}|^2$ *in the BCS state, as a function of the energy* e_k*. In the absence of interactions, this function is equal to* 1 *for* $e_k < \mu$*, and zero for* $e_k > \mu$ *(dotted line step function). In the presence of interactions, due to the pairing of fermions the curve is rounded off over an energy domain of the order of the gap* Δ *(double arrow on the horizontal axis), and the variations occur around the value* $\overline{\mu}$ *(value of* μ *shifted by the mean field effect). The dashed curve plots the product* $u_{\mathbf{k}} v_{\mathbf{k}}$ *as a function of the same variable* e_k*. This function is largest around* $e_k = \overline{\mu}$*, in a domain spreading over a few* Δ*.*

between 0 and 1 occurs around $\overline{\mu}$, hence for a value of chemical potential shifted by the mean field effect as indicated by relation (41); this energy shift due to the mean field is natural. What is more striking is that the curve no longer presents a discontinuous step, but varies progressively over an energy domain whose width is of the order of the gap Δ. The interaction effect depopulates certain pairs of states in favor of other pairs having higher kinetic energies. Certain fermions are promoted from the inside of the Fermi surface towards the outside, this effect occurring over a depth of the order of Δ. In $\mathbf{k}$ space, the perturbation introduced by the attractions is localized in the neighborhood of that surface; the fermions situated close to the center of the Fermi sphere are not concerned, whereas those close to that surface gain an energy of the order of the gap.

β. Position space

Relations (B-22) and (B-23) of Chapter XVI yield the one-particle correlation function in position space:

$$G_1(\mathbf{r}, \nu; \mathbf{r}', \nu') = \left\langle \Psi_\nu^\dagger(\mathbf{r}) \Psi_{\nu'}(\mathbf{r}') \right\rangle = \langle \mathbf{r}', \nu' | \, \rho_I \, | \mathbf{r}, \nu \rangle \tag{47}$$

where ρ_I is the one-particle density operator. Taking into account formula (A-14) of Chapter XVI, applied to normalized plane waves, we can write:

$$G_1(\mathbf{r},\nu;\mathbf{r}',\nu') = \frac{1}{L^3}\sum_{\mathbf{k},\mathbf{k}'} e^{i\left(\mathbf{k}'\cdot\mathbf{r}'-\mathbf{k}\cdot\mathbf{r}\right)} \left\langle a^\dagger_{\mathbf{k},\nu} a_{\mathbf{k}',\nu'} \right\rangle$$
$$= \frac{1}{L^3}\sum_{\mathbf{k},\mathbf{k}'} e^{i\left(\mathbf{k}'\cdot\mathbf{r}'-\mathbf{k}\cdot\mathbf{r}\right)} \ \langle\Psi_{\mathrm{BCS}}|\, a^\dagger_{\mathbf{k},\nu} a_{\mathbf{k}',\nu'}\, |\Psi_{\mathrm{BCS}}\rangle \tag{48}$$

where $a^\dagger_{\mathbf{k},\nu}$ is the creation operator in the individual state of momentum $\hbar\mathbf{k}$ and spin ν, and $a_{\mathbf{k}',\nu'}$ the annihilation operator in the state of momentum $\hbar\mathbf{k}'$ and spin ν'. Now, in the state $|\Psi_{\mathrm{BCS}}\rangle$, the occupation number of each momentum pair are either 0 or 2, which means that the average values of the product of these operators is zero whenever each of them concerns a different pair, or if the two individual states are in the same pair, but are different from each other. We now have:

$$G_1(\mathbf{r},\nu;\mathbf{r}',\nu') = \frac{1}{L^3}\delta_{\nu\nu'}\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\left(\mathbf{r}'-\mathbf{r}\right)} \ \langle\Psi_{\mathrm{BCS}}|\, a^\dagger_{\mathbf{k},\nu} a_{\mathbf{k},\nu}\, |\Psi_{\mathrm{BCS}}\rangle$$
$$= \frac{1}{L^3}\delta_{\nu\nu'}\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\left(\mathbf{r}'-\mathbf{r}\right)} \langle\hat{n}_{\mathbf{k},\nu}\rangle \tag{49}$$

with:

$$\langle n_{\mathbf{k},\nu}\rangle = \langle\Psi_{\mathrm{BCS}}|\, a^\dagger_{\mathbf{k},\nu} a_{\mathbf{k},\nu}\, |\Psi_{\mathrm{BCS}}\rangle = |v_{\mathbf{k}}|^2 \tag{50}$$

The function G_1 is proportional to the Fourier transform of the average population $\langle n_{\mathbf{k},\nu}\rangle$ of the individual state $\mathbf{k},\nu$. As this average population is a function whose width is of the order of the Fermi wave vector k_F, the function G_1 goes to zero when $|\mathbf{r}-\mathbf{r}'| \gg 1/k_F$, i.e. as soon as the difference in positions is no longer microscopic: in this system, there exists no "long-range non-diagonal order" of the one-particle correlation function. For $\mathbf{r}=\mathbf{r}'$, we get:

$$G_1(\mathbf{r},\nu;\mathbf{r},\nu') = \frac{\delta_{\nu\nu'}}{2L^3}\left\langle\hat{N}\right\rangle = \delta_{\nu\nu'}\frac{n_{\mathrm{BCS}}}{2} \tag{51}$$

where n_{BCS} is the numerical particle density:

$$n_{\mathrm{BCS}} = \frac{\left\langle\hat{N}\right\rangle}{L^3} \tag{52}$$

The function $G_1(\mathbf{r},\nu;\mathbf{r},\nu)$ has no spatial dependence; the factor $1/2$ reflects the fact that the total density n_{BCS} is shared equally among the two spin states. The results are the same as for an ideal gas.

2-b. Two-particle distribution, internal pair wave function

As opposed to the one-particle distribution, the two-particle distribution is strongly affected by the BCS mechanism, which is to be expected since it is a pairing process.

α. Momentum space, peak in the distribution

Relation (C-19) of Chapter XV yields the expression for the matrix elements of the two-particle density operator ρ_{II} in an unspecified basis. In the momentum representation, they are written:

$$\langle 1:\mathbf{k}'',\nu_3;2:\mathbf{k}''',\nu_4|\,\rho_{II}\,|1:\mathbf{k},\nu_1;2:\mathbf{k}',\nu_2\rangle = \left\langle a^{\dagger}_{\mathbf{k},\nu_1}a^{\dagger}_{\mathbf{k}',\nu_2}a_{\mathbf{k}''',\nu_4}a_{\mathbf{k}'',\nu_3}\right\rangle \tag{53}$$

We shall mainly consider the diagonal elements, characterizing the correlations between the momenta of two particles:

$$\langle 1:\mathbf{k},\nu;2:\mathbf{k}',\nu'|\,\rho_{II}\,|1:\mathbf{k},\nu;2:\mathbf{k}',\nu'\rangle = \left\langle a^{\dagger}_{\mathbf{k},\nu}a^{\dagger}_{\mathbf{k}',\nu'}a_{\mathbf{k}',\nu'}a_{\mathbf{k},\nu}\right\rangle \tag{54}$$

In this expression, the creation operators repopulate precisely the same states as those that have been depopulated by the annihilation operators.

For $\nu'=\nu$, in order to obtain a non-zero result, we must have $\mathbf{k}'\neq\mathbf{k}$ (otherwise we get the square of a fermionic operator, which is zero). We get:

$$\langle 1:\mathbf{k},\nu;2:\mathbf{k}',\nu|\,\rho_{II}\,|1:\mathbf{k},\nu;2:\mathbf{k}',\nu\rangle = |v_{\mathbf{k}}|^2\,|v_{\mathbf{k}'}|^2 \qquad (\text{if } \mathbf{k}'\neq\mathbf{k}) \tag{55}$$

which means there exists no correlations between the momenta.

For $\nu'=-\nu$, when $\mathbf{k}'$ is different from $-\mathbf{k}$, two different pairs are involved, and we again get a product[6]:

$$\langle 1:\mathbf{k},\nu;2:\mathbf{k}',-\nu|\,\rho_{II}\,|1:\mathbf{k},\nu;2:\mathbf{k}',-\nu\rangle = |v_{\mathbf{k}}|^2\,|v_{\mathbf{k}'}|^2 \qquad (\text{if } \mathbf{k}'\neq-\mathbf{k}) \tag{56}$$

On the other hand, if $\mathbf{k}'=-\mathbf{k}$, only one pair is concerned, destroyed and then reconstructed by the operators (as before, this is a contribution from the diagram in Figure 4 of Complement B_{XVII}; the computation then involves a single state $|\overline{\varphi}_{\mathbf{k}}\rangle$, and we get:

$$\langle 1:\mathbf{k},\nu;2:-\mathbf{k},-\nu|\,\rho_{II}\,|1:\mathbf{k},\nu;2:-\mathbf{k},-\nu\rangle = |v_{\mathbf{k}}|^2 \tag{57}$$

This result is not the limit of the previous one when $\mathbf{k}'\to-\mathbf{k}$, which would be $|v_{\mathbf{k}}|^4$; the value we obtain is larger, since $|v_{\mathbf{k}}|^2\leq 1$.

This shows that for all the k values that do not correspond to a pair of opposite momenta, the density operator is simply a product, involving no correlations between the particles' momenta. This confirms what we found in the study of the one-particle density operator: all the $\mathbf{k}$ states having a momentum smaller than that of the Fermi level are populated, with a rounding off of the functions due to the pairing phenomenon. On the other hand, for opposite values of both the momenta and the spin values, as is the case for a pair, we observe a discontinuity of the diagonal correlation function: it jumps from $|v_{\mathbf{k}}|^4$ to the larger value $|v_{\mathbf{k}}|^2$. The corresponding discontinuity $\delta n^{(2)}_{\mathbf{k}}$ can be written as:

$$\begin{aligned}\delta n^{(2)}_{\mathbf{k}} &= |v_{\mathbf{k}}|^2 - |v_{\mathbf{k}}|^4 = |v_{\mathbf{k}}|^2\left[1-|v_{\mathbf{k}}|^2\right]\\ &= |u_{\mathbf{k}}v_{\mathbf{k}}|^2\end{aligned} \tag{58}$$

[6] If $\nu=-$, to get this result we use the fact that the function $v(\mathbf{k})$ is even. Now if $\mathbf{k}'=\mathbf{k}$, two pairs are still involved, labeled by opposite values of momentum (remember that we chose the convention where each pair is labeled by the momentum of the spin + particle); but, here again, the parity of $\nu(\mathbf{k})$ leads to $|v(\mathbf{k})|^4$, in agreement with (56).

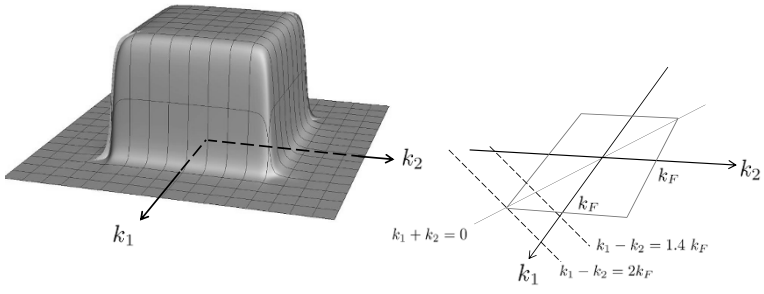


Figure 2: The left part of the figure shows the distribution function of the momenta of two particles, assuming that the two momenta $\hbar\mathbf{k}_1$ et $\hbar\mathbf{k}_2$ are parallel (or antiparallel); a value $\Delta/E_F = 1/10$ has been chosen. In the descending part of the surface, and in the two corners where $\mathbf{k}_1 + \mathbf{k}_2$ vanishes, one can distinguish a small crest indicating a partial Bose-Einstein condensation. In order to see this effect more clearly, we cut the surface along vertical planes whose trace is indicated by the dashed lines in the right part of the figure. This leads to the curves of Figure 3.

We shall see below (§ 2-b-β) that $\delta n_{\mathbf{k}}^{(2)}$ is none other than the square of the $\mathbf{k}$ component of the pair wave function. This discontinuity is significant for the values of $\mathbf{k}$ for which the product $u_{\mathbf{k}} v_{\mathbf{k}}$ takes on its largest values; Figure 1 shows that it corresponds to a region around the Fermi surface, with an energy bandwidth of the order of the gap Δ.

The momentum distribution function depends on the 6 components of the two momenta, which does not allow a simple graphic representation. To simplify, we are going to assume the two particles' momenta $\hbar\mathbf{k}_1$ and $\hbar\mathbf{k}_2$ are parallel (or antiparallel), so that the distribution we wish to represent becomes a surface in three-dimensional space: we plot k_1 along one axis, k_2 along the second, and the probability along the third perpendicular axis. We then obtain the surface shown in the left part of Figure 2, where it has been assumed that $\Delta/E_F = 1/10$.

To explore this surfact in more detail, we plot in Figure 3 the curves we obtain by cutting this surface by vertical planes parallel to the first bisector of the k_1 and k_2 axes; we assume the difference $k_1 - k_2$ to be constant (dashed lines in the right part of Figure 2) and use, as the variable, the sum of these two momenta. The horizontal axis in Figure 3 then represents the dimensionless variable q:

$$q = \frac{k_1 + k_2}{2k_F} \tag{59}$$

As q varies, the corresponding point in the plane k_1, k_2 moves along a straight line.

For a fixed value of $\kappa = k_1 - k_2$, we must set $q = 0$ for the wave vector components to have opposite values: $k_1 = \kappa/2$ and $k_2 = -\kappa/2$. On the left-hand side of Figure 3, the difference κ is chosen equal to 1.4 k_F; we obtain an almost square curve, rounded off by the fact that Δ is not zero (as was the case in Figure 1), and which presents

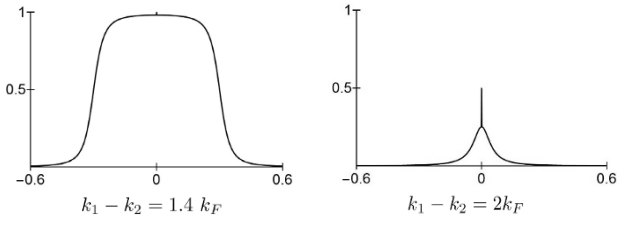


Figure 3: Plots of the distribution function for two particles of parallel (or antiparallel) momenta $\mathbf{k}_1$ *and* $\mathbf{k}_2$*, as a function of the dimensionless variable* $q = (k_1 + k_2)/2k_F$*; the figure was plotted with the choice* $\Delta/E_F = 1/10$*.*
For the curve on the left-hand side, the difference $(k_1 - k_2)$ *is chosen equal to* $1.4\ k_F$*. The curve looks like a bell shaped function, practically constant for small values of q, and decreasing for larger values following a rounded slope similar to the one in Figure 1 (and all the more steep as the* Δ *value is chosen smaller). No singularity of the distribution is clearly visible (except for a minuscule peak at the origin).*
For the curve on the right-hand side, the difference in momenta is chosen equal to $2\hbar k_F$*; when q is close to zero, the two momenta take on values that both fall into the rounded part of the one-particle distribution. A singularity at* $q = 0$ *is now clearly visible, signaling an accumulation of "molecules" in a state where their center of mass does not move. The height of the central peak corresponds to the population of the discrete level having a zero total momentum, and its width reduces to zero as it is a discrete level.*

a barely visible peak. But, as we saw before, the effects of the pairing are important when $k_1 = -k_2 \simeq \pm k_F$, i.e. when $\kappa \simeq \pm 2k_F$. On the right-hand side of Figure 3, the momentum difference is chosen equal to $2\hbar k_F$, so that the momenta can both fall in the rounded part of the distributions. We observe, superposed on a "pedestal", a narrow peak indicating an additional population in the level having a zero total momentum. The value of that population is given by the height of the peak, whose width is strictly zero for discrete levels. The singularity of this momentum distribution is then clearly visible.

Therefore, a singularity appears in the momentum distribution of particle pairs, whose centers of mass present a condensation in momentum space. This is, however, a partial condensation: as opposed to the boson case, the condensation peak appears on a pedestal due the presence of a majority of non-condensed pairs. Actually, the only pairs involved are those whose two components have energies falling around the Fermi level E_F, in a bandwidth of the order of the gap Δ. Despite these restrictions, it is nevertheless true that the condensation phenomenon into attractive BCS pairs has properties related to Bose-Einstein condensation for repulsive bosons. The link between that condensation and the appearance of an order parameter for the pair field is discussed in § 2-b-γ of Complement A_{XVII}.

β. Position space, correlations described by the pair wave functions

We did not find any effects of the interactions on G_1. But, as pointed out before, since the BCS theory relies on pairing, one expects to find more interesting properties concerning the two-particle correlation functions. They will be studied now, limiting ourselves to the "diagonal" correlation function, as defined by relation (B-33) of Chapter XVI, including the spin variables as in (B-36). This function is written as:

$$\begin{aligned} G_2(\mathbf{r},\nu;\mathbf{r}',\nu') &= \left\langle \Psi_\nu^\dagger(\mathbf{r})\Psi_{\nu'}^\dagger(\mathbf{r}')\Psi_{\nu'}(\mathbf{r}')\Psi_\nu(\mathbf{r})\right\rangle \\ &= \langle 1:\mathbf{r},\nu;2:\mathbf{r}',\nu'|\ \rho_{II}\,|1:\mathbf{r},\nu;2:\mathbf{r}',\nu'\rangle \end{aligned} \tag{60}$$

(ρ_{II} is the two-particle density operator), or else, as before:

$$G_2(\mathbf{r},\nu;\mathbf{r}',\nu') = \frac{1}{L^6}\sum_{\mathbf{k},\mathbf{k}',\mathbf{k}'',\mathbf{k}'''} e^{i\left[(\mathbf{k}-\mathbf{k}''')\cdot\mathbf{r}+(\mathbf{k}'-\mathbf{k}'')\cdot\mathbf{r}'\right]}\left\langle a_{\mathbf{k}''',\nu}^\dagger a_{\mathbf{k}'',\nu'}^\dagger a_{\mathbf{k}',\nu'}a_{\mathbf{k},\nu}\right\rangle \tag{61}$$

This expression includes the average values of the product of four operators, whose computation is similar to the one explained in § 3 of Complement $\mathrm{B_{XVII}}$ for the average interaction energy, except that, in our case, the spin indices are fixed rather than appearing as summation indices. Figure 4 schematizes with a diagram each term of (61): the incoming arrows represent particles that disappear (annihilation operator action), the outgoing arrows represent those that will appear (creation operator action); each value of $\mathbf{k}$ is associated with a position value $\mathbf{r}$, via an exponential $e^{i\mathbf{k}\cdot\mathbf{r}}$ for the incoming arrows, or $e^{-i\mathbf{k}\cdot\mathbf{r}}$ for the outgoing ones, as well as with a value of the spin ν.

Parallel spins: if $\nu=\nu'$, the two destruction operators necessarily concern pairs with different $\mathbf{k}$. To restore the populations of these two couples of states to an even value, the only possibility is to again give each one its initial value; otherwise the result will be zero. We must have, either $\mathbf{k}=\mathbf{k}'''$ and $\mathbf{k}'=\mathbf{k}''$ (direct term), or $\mathbf{k}=\mathbf{k}''$ and $\mathbf{k}'=\mathbf{k}'''$ (exchange term). In the first case, we obtain (after two anticommutations whose sign changes cancel each other) a result[7] independent of the position variables:

$$\frac{1}{L^6}\sum_{\mathbf{k},\mathbf{k}'}\langle\overline{\varphi}_\mathbf{k}|\,a_{\mathbf{k},\nu}^\dagger a_{\mathbf{k},\nu}\,|\overline{\varphi}_\mathbf{k}\rangle\,\langle\overline{\varphi}_{\mathbf{k}'}|\,a_{\mathbf{k}',\nu}^\dagger a_{\mathbf{k}',\nu}\,|\overline{\varphi}_{\mathbf{k}'}\rangle = \frac{1}{L^6}\sum_{\mathbf{k},\mathbf{k}'}|v_\mathbf{k}|^2\,|v_{\mathbf{k}'}|^2 \tag{62}$$

and in the second case (after only one anticommutation):

$$\begin{aligned} -\frac{1}{L^6}\sum_{\mathbf{k},\mathbf{k}'}e^{i(\mathbf{k}-\mathbf{k}')\cdot(\mathbf{r}-\mathbf{r}')}\,&\langle\overline{\varphi}_\mathbf{k}|\,a_{\mathbf{k},\nu}^\dagger a_{\mathbf{k},\nu}\,|\overline{\varphi}_\mathbf{k}\rangle\,\langle\overline{\varphi}_{\mathbf{k}'}|\,a_{\mathbf{k}',\nu}^\dagger a_{\mathbf{k}',\nu}\,|\overline{\varphi}_{\mathbf{k}'}\rangle \\ &= -\frac{1}{L^6}\sum_{\mathbf{k},\mathbf{k}'}e^{i(\mathbf{k}-\mathbf{k}')\cdot(\mathbf{r}-\mathbf{r}')}\,|v_\mathbf{k}|^2\,|v_{\mathbf{k}'}|^2 \end{aligned} \tag{63}$$

Regrouping these two contributions, and using (6), we get:

$$G_2(\mathbf{r},\nu;\mathbf{r}',\nu) = \left(\frac{\langle N\rangle}{2L^3}\right)^2\left[1-\left[F(\mathbf{r}-\mathbf{r}')\right]^2\right] \tag{64}$$

[7] If $\nu=-$, we must change the sign of $\mathbf{k}$ and $\mathbf{k}'$ in $v_\mathbf{k}$ and $v_{\mathbf{k}'}$ but, as before, it does not change the result since we can change the sign of summation variables.

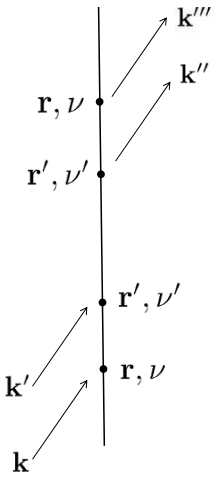


Figure 4: This diagram symbolizes each term involved in the binary correlation function. The two incoming arrows on the bottom represent particles that will disappear during the interaction under the action of the two annihilation operators; the two outgoing arrows on the top represent particles that appeared during the interaction under the action of the two creation operators. All the arrows are associated with an imaginary exponential of the position, with a positive argument for the incoming arrows, and a negative one for the outgoing arrows. The indices ν label the spins.

with an exchange term containing the (real) function:

$$F(\mathbf{r}) = \frac{2}{\langle N \rangle} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \left| v_{\mathbf{k}} \right|^2 \tag{65}$$

This result has the same form as relation (22) of Complement A_{XVI}, taking into account the fact that the population of each spin state is half of $\langle N \rangle$. It shows that the correlation function for two parallel spins exhibits an "exchange hole" very similar to the one plotted in Figure 2 of that complement, but with a slightly different shape, since here the functions $\left| v_{\mathbf{k}} \right|^2$ are no longer exactly discontinuous step functions. The width of this exchange hole is of the order of the inverse of k_F, the Fermi wave number related to the Fermi energy E_F by $E_F = \hbar^2 k_F^2 / 2m$.

Opposite spins: If $\nu \neq \nu'$, it is possible for the two annihilation or creation operators to act on the same pair of states; we are then dealing with a pair annihilation-creation term (term of type **II** according to the classification presented in § D-1-a of Chapter XVII). Figure 5 symbolizes the three types of diagrams playing a role in the computation of the correlation function for opposite spins: **I** (forward scattering), **II** (pair-pair) and **III** (special cases). The computation of their sum has been performed in § D-2 of that same chapter, and led to the following result:

$$G_2(\mathbf{r}, \nu; \mathbf{r}', -\nu) \simeq \left[\frac{\langle N \rangle}{2L^3} \right]^2 + \left| \phi_{\text{pair}}(\mathbf{r} - \mathbf{r}') \right|^2 \tag{66}$$

We have used the following definition of the (non-normalized) "pair wave function"[8] $\phi_{\text{pair}}(\mathbf{r}-\mathbf{r}')$:

$$\phi_{\text{pair}}(\mathbf{r}) = \frac{1}{L^3}\sum_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{r}} = \frac{\Delta}{2L^3}\sum_{\mathbf{k}} \frac{1}{\sqrt{(\xi_k)^2+\Delta^2}} e^{-i\mathbf{k}\cdot\mathbf{r}} \tag{67}$$

which is simply the pair wave function already introduced in relation (D-14) of Chapter XVII. We find again relations (38) and (39) of Complement A_{XVII}, where this wave function was obtained by a different method involving the two-particle field operator. The presence of the second term on the right-hand side of (66) is thus due to the non-zero average value of the pair field, introduced in that complement (non-zero order parameter).

We have just shown that, contrary to what happens in an ideal gas (Complement A_{XVI}, § 2-b), two particles with opposite spins may be spatially correlated. This correlation is described by the modulus squared of the wave function $\phi_{\text{pair}}(\mathbf{r}-\mathbf{r}')$, defined by its spatial Fourier transform $u_{\mathbf{k}} v_{\mathbf{k}}$. This new wave function, different from the one we used at the beginning to build the N-particle trial wave function, was introduced in § D-2 of Chapter XVII, as well as in Complement A_{XVII}, starting from the pair field operator. The spatial correlation it characterizes is purely dynamic, as it does not exist in the absence of interactions. Its physical consequences, in terms of potential energy, will be discussed in § 3-a-α.

Physical discussion: on the right-hand side of (66), the first term does not contain any spatial dependence; it simply corresponds to the correlation function of an ensemble of totally independent particles. The second term, on the other hand, depends on the position differences $\mathbf{r}-\mathbf{r}'$; we now discuss its physical origin in terms of quantum interference.

This second term comes from the contribution of the pair annihilation-creation terms for which we have, in relation (61), $k=-k'$ and $k''=-k'''$. Let us show that "cutting them in half", they look like interference terms. They include average values of operator products that, when $\nu'=-\nu$, can be written:

$$\begin{aligned}\left\langle a^{\dagger}_{\mathbf{k}''',\nu} a^{\dagger}_{-\mathbf{k}''',-\nu} a_{-\mathbf{k},-\nu} a_{\mathbf{k},\nu}\right\rangle &= \langle\Psi_{\text{BCS}}|\, a^{\dagger}_{\mathbf{k}''',\nu} a^{\dagger}_{-\mathbf{k}''',-\nu} a_{-\mathbf{k},-\nu} a_{\mathbf{k},\nu}\,|\Psi_{\text{BCS}}\rangle \\ &= \left\langle \widetilde{\Psi}(\mathbf{k}''')\,\middle|\,\widetilde{\Psi}(\mathbf{k})\right\rangle \end{aligned} \tag{68}$$

where $\left|\widetilde{\Psi}(\mathbf{k})\right\rangle$ is defined as:

$$\left|\widetilde{\Psi}(\mathbf{k})\right\rangle = a_{-\mathbf{k},-\nu} a_{\mathbf{k},\nu}\,|\Psi_{\text{BCS}}\rangle \tag{69}$$

Relation (66) then becomes:

$$G_2(\mathbf{r},\nu;\mathbf{r}',-\nu) \simeq \left[\frac{\langle N\rangle}{2L^3}\right]^2 + \frac{1}{L^6}\sum_{\mathbf{k},\mathbf{k}'''} e^{i(\mathbf{k}-\mathbf{k}''')\cdot(\mathbf{r}-\mathbf{r}')} \left\langle \widetilde{\Psi}(\mathbf{k}''')\,\middle|\,\widetilde{\Psi}(\mathbf{k})\right\rangle \tag{70}$$

[8] The factor $1/L^3$ appearing in (67) permits defining a pair wave function independent of the dimension L of the physical system, in the limit of large L where the sum over the $\mathbf{k}$ becomes an integral over d^3k multiplied by $(L/2\pi)^3$. As a result, the square of that wave function is not homogeneous to the inverse of a volume, as is generally the case for a particle wave function, but to $1/L^6$. Actually, it should be considered as a two-particle wave function, product of a constant wave function $1/L^{3/2}$ of the center of mass of the pair (assumed to have a zero momentum) and a wave function describing its relative position variable.

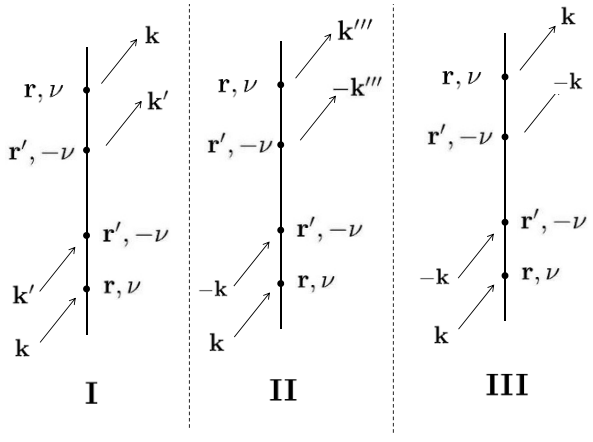


*Figure 5: Diagrams symbolizing various contributions to the binary correlation function for opposite spins ($\nu \neq \nu'$). The diagram of type **I** corresponds to a process where two particles with opposite spins are destroyed and then re-created in exactly the same individual states (forward scattering). The type **II** diagram corresponds to the case where two particles of the same pair are destroyed, and then two particles are created in the states of another pair (pair annihilation-creation process). Finally, type **III** diagram is a special case of the previous diagrams, and yields a negligible contribution. It is the type **II** diagram that introduces the spatial dependance of the correlation function.*

The position dependent term in the correlation function can be interpreted as resulting from the interference between a process where two particles of the same pair $(\mathbf{k}, -\mathbf{k})$ are annihilated, and a second process where the two annihilated particles are from another pair $(\mathbf{k}''', -\mathbf{k}''')$; these two processes are schematized in Figure 6.

According to (1) and (69), we have:

$$\left|\widetilde{\Psi}\left(\mathbf{k}\right)\right\rangle = v_{\mathbf{k}}\left|n_{\mathbf{k}}=0; n_{-\mathbf{k}}=0\right\rangle \prod_{\otimes \mathbf{k}' \neq \mathbf{k}} \left|\overline{\varphi}_{\mathbf{k}'}\right\rangle \tag{71}$$

If $\mathbf{k} \neq \mathbf{k}'''$, the two states $\left|\widetilde{\Psi}\left(\mathbf{k}\right)\right\rangle$ and $\left|\widetilde{\Psi}\left(\mathbf{k}'''\right)\right\rangle$ are neither identical, nor orthogonal; they actually have identical components on all the pairs of states different from $(\mathbf{k}, -\mathbf{k})$ and $(\mathbf{k}''', -\mathbf{k}''')$, but these two pairs have the same component only for states where the 4 populations are zero. We can then write:

$$\left\langle\widetilde{\Psi}\left(\mathbf{k}'''\right)\right|\left.\widetilde{\Psi}\left(\mathbf{k}\right)\right\rangle = v^{*}_{\mathbf{k}'''} u_{\mathbf{k}'''} u^{*}_{\mathbf{k}} v_{\mathbf{k}} \tag{72}$$

Inserting this result in (70) yields relation (66), whose spatial dependence does come from the interference between the two processes schematized in Figure 6.

One could also use relation (68) of Complement A_{XVII} to express the product $\Psi_{\nu'}(\mathbf{r}')\Psi_{\nu}(\mathbf{r})$ as a function of a sum of pair annihilation operators. This is another way of understanding the role of pairs in the determination of the binary correlation function expression (66).

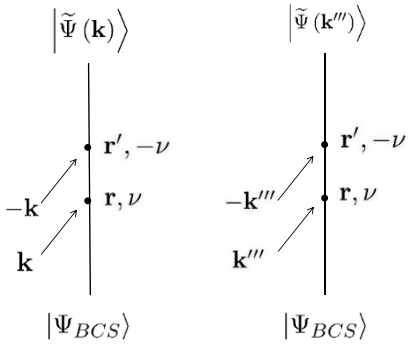


Figure 6: Diagram symbolizing two pair annihilation processes from the initial state $|\Psi_{BCS}\rangle$, leading respectively to the states $\left|\widetilde{\Psi}\,(\mathbf{k})\right\rangle$ and $\left|\widetilde{\Psi}\,(\mathbf{k}''')\right\rangle$. As these two states are not orthogonal, an interference effect occurs that is at the origin of the position dependent part of the binary correlation function.

2-c. Properties of the pair wave function, coherence length

The pair wave function plays an important role in the BCS theory, and not only for the binary correlation functions, as we already mentioned. Its range determines the coherence length of the system, and its norm is also related to the number of quanta present in the field of condensed pairs (Complement $\mathrm{A_{XVII}}$).

α. Form of the pair wave function

As the functions $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ only depend on the modulus of $\mathbf{k}$, we can apply the Fourier transform formulas for this case – see Appendix I, relation (52). Replacing the discrete summation by an integral, the pair wave function (67) becomes:

$$\phi_{\mathrm{pair}}\,(\mathbf{r}) = \left(\frac{1}{2\pi}\right)^3 \int \mathrm{d}^3k\; e^{-i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}} v_{\mathbf{k}} = \frac{1}{(2\pi)^2}\;\frac{\Delta}{r}\int_0^\infty k\mathrm{d}k \frac{1}{\sqrt{(e_k-\overline{\mu})^2+\Delta^2}}\;\sin kr \tag{73}$$

Therefore,the pair wave function is real. Figure 1 gives a plot of $u_{\mathbf{k}}v_{\mathbf{k}}$ as a function of the energy e_k; it presents a maximum in the vicinity of the surface of the Fermi sphere, with a peak whose width is of the order of Δ. More details on the role this pair wave function plays in the correlation functions are given in Complement $\mathrm{A_{XVII}}$.

The Fourier transform of $\phi_{\mathrm{pair}}(\mathbf{r})$ is thus concentrated around values of the modulus of $\mathbf{k}$ of the order of the Fermi wave vector k_F. Its spread δk is such that the corresponding variation of energy e_k is of the order of the gap Δ, which leads to the condition:

$$\frac{\hbar^2}{2m} 2k_F\;\delta k \simeq \Delta \quad \text{that is} \quad \frac{\delta k}{k_F} \simeq \frac{\Delta}{E_F} \tag{74}$$

This wave function oscillates[9] as a function of the position $\mathbf{r}$, at a spatial frequency approximately equal to the wave vector at the Fermi surface. These oscillations are damped over a length of the order of ξ_{pair} defined by (the arbitrary factor $2/\pi$ is introduced to match the usual definition found in the literature):

$$\xi_{\text{pair}} = \frac{2}{\pi \delta k} = \frac{2\hbar^2 k_F}{\pi m \Delta} = \frac{1}{k_F}\frac{4E_F}{\pi \Delta} \tag{75}$$

which is of the order of the distance between fermions, multiplied by the ratio E_F/Δ, very large compared to 1. Each fermion pair extends over a relatively large volume, leading to a strong overlap between pairs. In a superconductor, the length ξ_{pair} is called the "coherence length"[10]; it characterizes the capacity of the physical system to adapt to spatial constraints, and plays a role analogous to the "healing length" in systems of condensed bosons (Complement C_{XV}, § 4-b).

We have shown that the pairing significantly modifies the correlation functions for opposite spins: the particles now become correlated, whereas this was absolutely not the case for an ideal gas. It is a positive correlation, leading to a bunching tendency (the opposite of a Pauli exclusion); this explains the decrease in the interaction energy of the particles (§ 2-c-γ). On the other hand, the pairing has no significant effect on the correlation function of particles having the same spin direction; it remains similar to the correlation function of an ideal gas, with an exchange hole whose width is of the order of $1/k_F$. Relation (75) indicates that the width of this exchange hole is much smaller than the distance ξ_{pair} over which the modifications of the correlation function for opposite spins occur.

As mentioned before, the pair wave function has little in common with the initial wave function $\chi(\mathbf{r}_1 - \mathbf{r}_2)$ used in Chapter XVII to build the N-particle variational ket, since (3) shows that the Fourier transform of χ is $g_{\mathbf{k}} = v_{\mathbf{k}}/u_{\mathbf{k}}$. This is not surprising: when building a trial ket by the repetitive action of the same pair creation operator, we do not simply juxtapose those pairs. The antisymmetrization effects are dominant, and in each term of the expansion to the power P of operator $\sum_k g_{\mathbf{k}}\, a^\dagger_{\mathbf{k},\uparrow} a^\dagger_{-\mathbf{k},\downarrow}$ in relation (1), the result is zero each time the same value of $\mathbf{k}$ is repeated (the square of a fermionic creation operator is zero; two fermions cannot occupy the same individual state). This is why the antisymmetrization effects completely remodel the pairs formed in the system of N identical particles.

β. Norm of the pair wave function

According to (67), the component on $|\mathbf{k}\rangle$ of the ket $|\phi_{\text{pair}}\rangle$ associated with the wave function $\phi_{\text{pair}}(\mathbf{r})$ is written as:

$$\langle \mathbf{k} \,|\phi_{\text{pair}}\rangle = \frac{1}{L^{3/2}} u_{\mathbf{k}} v_{\mathbf{k}} \tag{76}$$

[9] The existence of this oscillation is confirmed by the fact that its integral over the entire space is practically zero; this integral is proportional to $u_0 v_0$, i.e. $u_0\sqrt{1-u_0^2}$, which is indeed practically zero since $u_0 \simeq 1$.

[10] The coherence length ξ_{pair} should not be confused with the (London) "penetration depth" that characterizes the magnetic field exclusion from a superconductor, and that depends on the charges of the particles.

(the functions $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are even). The square of this ket's norm is therefore:

$$\langle \phi_{\text{pair}} \,|\phi_{\text{pair}}\rangle = \frac{1}{L^3}\sum_{\mathbf{k}} |u_{\mathbf{k}} v_{\mathbf{k}}|^2 \tag{77}$$

Replacing the discrete summation by an integral, we get:

$$\langle \phi_{\text{pair}} \,|\phi_{\text{pair}}\rangle = \frac{1}{L^3}\left(\frac{L}{2\pi}\right)^3 \int \mathrm{d}^3 k \;\; |u_{\mathbf{k}} v_{\mathbf{k}}|^2 = \frac{\Delta^2}{4L^3}\int_0^\infty D\left(e_k\right) \;\; \frac{\mathrm{d}e_k}{\left(e_k - \overline{\mu}\right)^2 + \Delta^2} \tag{78}$$

where, in the second equality, we chose e_k as the integral variable, introducing the density of states $D\left(e_k\right)$ defined in (39). The function to be integrated converges since $D\left(e_k\right)$ varies as $\sqrt{e_k}$; this function is concentrated around $e_k = \overline{\mu}$, spreading over a width $\Delta \ll \overline{\mu}$. We can then, to a good approximation, replace $D\left(e_k\right)$ by its value D_F at the Fermi energy, and extend the lower bound of the integral to $-\infty$. As the integral of a Lorentz function is known (Appendix II, § 1-b):

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}e_k}{\left(e_k - \overline{\mu}\right)^2 + \Delta^2} = \frac{\pi}{\Delta} \tag{79}$$

we get:

$$\langle \phi_{\text{pair}} \,|\phi_{\text{pair}}\rangle = \frac{\pi}{4L^3} D_F \Delta \tag{80}$$

We showed in § 2-a-β of Complement $\mathrm{A_{XVII}}$ that the norm $\sqrt{\langle \phi_{\text{pair}} \,|\phi_{\text{pair}}\rangle}$ yields the average value of the field $\langle \Phi_{\text{pair}}\left(\mathbf{R}\right)\rangle$, hence the value of the order parameter. In § 2-b-β of that same complement we showed that the square of the norm, equal to $\langle \phi_{\text{pair}} \,|\phi_{\text{pair}}\rangle$, yields the large distance behavior of the average value $\left\langle \Phi^\dagger_{\text{pair}}\left(\mathbf{R}\right) \Phi_{\text{pair}}\left(\mathbf{R}'\right)\right\rangle$; the quantity $\langle \phi_{\text{pair}} \,|\phi_{\text{pair}}\rangle$ is thus related[11] to the field pair intensity (or, in other words, to the total number of quanta in that field).

In addition, we just saw in the above § 2-b-α that a peak in the momentum distribution signaled the presence of a Bose-Einstein condensation. Inserting (76) into (58) shows that this peak height is:

$$\delta n^{(2)}_{\mathbf{k}} = |u_{\mathbf{k}} v_{\mathbf{k}}|^2 = L^3 \left|\langle \mathbf{k} \,|\phi_{\text{pair}}\rangle\right|^2 \tag{81}$$

The total particle number associated with this peak is:

$$\sum_{\mathbf{k}} \delta n^{(2)}_{\mathbf{k}} = L^3 \langle \phi_{\text{pair}} \,|\phi_{\text{pair}}\rangle = \frac{\pi}{4} D_F \Delta \tag{82}$$

Consequently, the square of the norm, $\langle \phi_{\text{pair}} \,|\phi_{\text{pair}}\rangle$, multiplied by the volume, is also the total particle number in the condensation peak found in § 2-b-α, which confirms the previous interpretation.

[11] The pair field operators Φ_{pair} and $\Phi^\dagger_{\text{pair}}$ do not exactly satisfy the boson commutation relations (Complement $\mathrm{A_{XVII}}$); the operator $\Phi^\dagger_{\text{pair}}\Phi_{\text{pair}}$ is thus not, stricto sensu, an operator giving the number of quanta in the pair field.

γ. Link to the interaction energy

The energy term on the third line of (7) can be written, in the zero range potential approximation (29):

$$\sum_{\mathbf{k},\mathbf{k}'} V_{\mathbf{k}-\mathbf{k}'}\ v_{\mathbf{k}}^{*}\ u_{\mathbf{k}}\ v_{\mathbf{k}'}\ u_{\mathbf{k}'} = -V \sum_{\mathbf{k}} v_{\mathbf{k}}^{*}\ u_{\mathbf{k}} \sum_{\mathbf{k}'} v_{\mathbf{k}'}\ u_{\mathbf{k}'}$$
$$= -VL^{6} \left|\phi_{\text{pair}}(0)\right|^{2} \tag{83}$$

This result yields an energy proportional to V and to the probability that the two components of a pair, described by the wave function $\phi_{\text{pair}}(\mathbf{r})$, are found at the same point; this makes sense since the pair size is very large compared to the interaction potential range.

δ. Non-diagonal order

When studying Bose-Einstein condensation for bosons, we showed in § 3 of Complement A_{XVI} that the one-particle non-diagonal correlation function did not go to zero at large distance, when a significant fraction of the particles occupy the same individual state. Nevertheless, in § 2-a of the present complement, we found that this was not the case for a system of paired fermions, where the non-diagonal order goes to zero over a microscopic distance. This can be understood from a physical point of view, since, in the present case, there is no accumulation of particles in the same individual quantum state. On the other hand, we saw in § 2-b-α that the center of mass of the pairs of particles is subject to a phenomenon of partial accumulation, reminiscent of a Bose-Einstein condensation. It is thus natural to examine the properties of the non-diagonal functions relative to pairs, and compute the "non-diagonal position" average value:

$$\left\langle \Psi_{\uparrow}^{\dagger}(\mathbf{r})\Psi_{\downarrow}^{\dagger}(\mathbf{r})\Psi_{\downarrow}(\mathbf{r}')\Psi_{\uparrow}(\mathbf{r}') \right\rangle \tag{84}$$

With two positions $\mathbf{r}'$ on the right, and two positions $\mathbf{r}$ on the left, this expression is the exact transposition to two particles of the one-particle non-diagonal function: a couple of particles with opposite spins are annihilated at point $\mathbf{r}'$, and then recreated at a different point $\mathbf{r}$. In a more general way, we are going to evaluate the 4-point average value:

$$\left\langle \Psi_{\uparrow}^{\dagger}(\mathbf{r}_1)\Psi_{\downarrow}^{\dagger}(\mathbf{r}_2)\Psi_{\downarrow}(\mathbf{r}_3)\Psi_{\uparrow}(\mathbf{r}_4) \right\rangle \tag{85}$$

which, following the same computation steps as for the one-particle functions, can be written as:

$$\left\langle \Psi_{\uparrow}^{\dagger}(\mathbf{r}_1)\Psi_{\downarrow}^{\dagger}(\mathbf{r}_2)\Psi_{\downarrow}(\mathbf{r}'_2)\Psi_{\uparrow}(\mathbf{r}'_1) \right\rangle$$
$$= \frac{1}{L^{6}} \sum_{\mathbf{k},\mathbf{k}',\mathbf{k}'',\mathbf{k}'''} e^{i\left[\left(\mathbf{k}\cdot\mathbf{r}'_1+\mathbf{k}'\cdot\mathbf{r}'_2\right)-\left(\mathbf{k}''\cdot\mathbf{r}_2+\mathbf{k}'''\cdot\mathbf{r}_1\right)\right]} \left\langle a_{\mathbf{k}''',\uparrow}^{\dagger} a_{\mathbf{k}'',\downarrow}^{\dagger} a_{\mathbf{k}',\downarrow} a_{\mathbf{k},\uparrow} \right\rangle \tag{86}$$

In this equality, the matrix elements have already been evaluated in § 2-b-β. We are going to show that this expression can be written as:

$$\left\langle \Psi^{\dagger}_{\uparrow}(\mathbf{r}_1)\Psi^{\dagger}_{\downarrow}(\mathbf{r}_2)\Psi_{\downarrow}(\mathbf{r}'_2)\Psi_{\uparrow}(\mathbf{r}'_1)\right\rangle \\ = G_1(\mathbf{r}_1,\uparrow;\mathbf{r}'_1,\uparrow)\; G_1(\mathbf{r}_2,\downarrow;\mathbf{r}'_2,\downarrow) + \phi^{*}_{\text{pair}}(\mathbf{r}_1-\mathbf{r}_2)\; \phi_{\text{pair}}\left(\mathbf{r}'_1-\mathbf{r}'_2\right) \tag{87}$$

where the one-particle non-diagonal distribution G_1 has been defined in relation (49) for the case $\nu = \nu'$, and the pair wave function ϕ_{pair} in relation (67). This equality is the same as relation (72) of Complement A_{XVII}, but is now obtained by another method.

Demonstration:

To compute expression (86), we distinguish several cases, as already explained on several occasions:

(**I**) Forward scattering terms; if the annihilation operators do not act on two states of the same pair ($(\mathbf{k} \neq -\mathbf{k}')$, this term will be non-zero only if $\mathbf{k}''' = \mathbf{k}$ and $\mathbf{k}'' = \mathbf{k}'$, in which case it is written:

$$\frac{1}{L^6}\sum_{\mathbf{k},\mathbf{k}'} e^{i\left[\mathbf{k}\cdot\left(\mathbf{r}'_1-\mathbf{r}_1\right)+\mathbf{k}'\cdot\left(\mathbf{r}'_2-\mathbf{r}_2\right)\right]}\,|v_{\mathbf{k}}|^2\,|v_{\mathbf{k}'}|^2 = G_1(\mathbf{r}_1,\uparrow,\mathbf{r}'_1,\uparrow)\; G_1(\mathbf{r}_2,\downarrow,\mathbf{r}'_2,\downarrow) \tag{88}$$

As we already mentioned, for example in Chapter XVII (§ D-1-a), the constraints on the summation indices can be ignored if the size of the system, L, is macroscopic; the two summations then become independent.

(**II**) Terms corresponding to the annihilation-creation of different pairs; if $\mathbf{k} = -\mathbf{k}'$ and $\mathbf{k}'' = -\mathbf{k}'''$ but $\mathbf{k} \neq \mathbf{k}'''$ (annihilation-creation of different pairs), we get the contribution:

$$\frac{1}{L^6}\sum_{\mathbf{k},\mathbf{k}''} e^{i\left[\mathbf{k}\cdot\left(\mathbf{r}'_1-\mathbf{r}'_2\right)-\mathbf{k}''\cdot\left(\mathbf{r}_2-\mathbf{r}_1\right)\right]}\; u^{*}_{\mathbf{k}}v_{\mathbf{k}}u_{\mathbf{k}''}v^{*}_{\mathbf{k}''} = \phi_{\text{pair}}(\mathbf{r}'_1-\mathbf{r}'_2)\; \phi^{*}_{\text{pair}}(\mathbf{r}_2-\mathbf{r}_1) \tag{89}$$

As, in addition, the function $\phi_{\text{pair}}(\mathbf{r})$ is even, we indeed obtain the second term of the right-hand side of (87).

(**III**) If $\mathbf{k} = -\mathbf{k}'$ and $\mathbf{k}'' = -\mathbf{k}'''$, and furthermore $\mathbf{k} = \mathbf{k}'''$ (annihilation-creation of the same pair), we get:

$$\frac{1}{L^6}\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\left(\mathbf{r}'_1+\mathbf{r}_2-\mathbf{r}'_2-\mathbf{r}_1\right)}\,|v_{\mathbf{k}}|^2 \tag{90}$$

This term is negligible as it is proportional to $\langle N\rangle / L^6$ when all the positions are the same, whereas the term (**I**) is proportional to $\langle N\rangle^2 / L^6$.

Let us now assume the positions can be grouped two by two: $\mathbf{r}_1$ and $\mathbf{r}_2$ are close to each other, as are $\mathbf{r}'_1$ and $\mathbf{r}'_2$, but that the two groups' positions are further away from each other. Under these conditions, the first term in G_1 on the right-hand side of (87), which has a microscopic range in $(\mathbf{r}_1 - \mathbf{r}'_1)$ and $(\mathbf{r}_2 - \mathbf{r}'_2)$, becomes very small. We are then left with the product of the pair wave functions. It follows that the non-diagonal

correlation function is simply the product of pair wave functions calculated at the relative positions[12].

In the particular case where $\mathbf{r}_1 = \mathbf{r}_2 = \mathbf{r}$ and $\mathbf{r}'_2 = \mathbf{r}'_1 = \mathbf{r}'$, we get the pair correlation function (84), which obeys:

$$\left\langle \Psi^\dagger_\uparrow(\mathbf{r})\Psi^\dagger_\downarrow(\mathbf{r})\Psi_\downarrow(\mathbf{r}')\Psi_\uparrow(\mathbf{r}') \right\rangle \underset{\mathbf{r}-\mathbf{r}'\to\infty}{\longrightarrow} \left|\phi_{\text{pair}}(0)\right|^2 \tag{91}$$

This non-zero long distance limit signals the existence of a non-diagonal order for the two-particle density operator. It comes, as was already the case for the pair wave function, from contribution (**II**), meaning from terms corresponding to the annihilation-creation of different pairs. This situation is reminiscent of what we encountered in the case of a condensed boson gas; but in the present case, the non-diagonal order concerns the pairs and not the individual particles.

3. Physical discussion

In an ideal gas of fermions, and as we saw in Complement A_{XVI}, there already exist strong correlations between the particles, simply due to their indistinguishability (a purely statistical effect). In the presence of attractive interactions, the BCS mechanism introduces additional correlations (dynamic correlations) that lower the total system energy. We are going to show that this decrease in energy comes from a slight imbalance between an increase in kinetic energy and a decrease of the potential energy, the latter one slightly surpassing the first one.

For clarity, we shall discuss this using the short-range potential approximation (§ 1-c) where all the matrix elements of the interaction potential are replaced by a constant $-V$ (V being positive); all the $\Delta_{\mathbf{k}}$ are then equal to the same gap Δ.

3-a. Modification of the Fermi surface and phase locking

The energy written in (7) first includes a kinetic energy term, then a mean field term expressed in terms of the average particle number. If that average number is constant, this term is independent of the quantum state of the system, and hence not related to the BCS pairing mechanism. By contrast, the last term in (7), which is the one we optimized in the variational calculation, is far more interesting; we call it the "pairing term", and use the words "pairing potential energy" or else "condensation energy" for its optimized value E_{paired}. As the $u_{\mathbf{k}}$ and the $v_{\mathbf{k}}$ are real, E_{paired} can be written as:

$$E_{\text{paired}} = -V\left[\sum_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}}\right]^2 \tag{92}$$

where the $u_{\mathbf{k}}$ and the $v_{\mathbf{k}}$ take the optimized values given in (27). We see that to get a large condensation energy, the sum $\sum_{\mathbf{k}} u_{\mathbf{k}} v_{\mathbf{k}}$ must take the largest possible value.

[12] As mentioned in note 8, the center of mass variables are not included in (87) since we assumed all the pairs to be at rest. If this were not the case, the long-range factorization of the non-diagonal order would be expressed as the product of a function of the two variables $(\mathbf{r}'_1 - \mathbf{r}'_2)$ and $(\mathbf{r}'_1 + \mathbf{r}'_2)/2$ by the complex conjugate of that same function of the two variables $(\mathbf{r}_1 - \mathbf{r}_2)$ and $(\mathbf{r}_1 + \mathbf{r}_2)/2$ (in other words by the product of a function of $\mathbf{r}'_1$ and $\mathbf{r}'_2$ by the complex conjugate of that same function of the two variables $\mathbf{r}_1$ and $\mathbf{r}_2$).

α. Compromise between the kinetic and potential energies

In an ideal gas, the ground state is the one for which all the individual states of energy lower than the Fermi level (chemical potential $\mu = E_F$) are occupied by one particle, and all the states above are totally empty. In the **k**-space, the particles each occupy a state inside one of the two Fermi spheres of radius k_F (with $E_F = \hbar^2 k_F^2/2m$), one associated with the + spin state, and one associated with the − spin state. Using the ket (1), such a state simply corresponds to the case where:

$$\begin{array}{llll} u_{\mathbf{k}} = 0 & \text{and} & v_{\mathbf{k}} = 1 & \text{for } k < k_F \\ u_{\mathbf{k}} = 1 & \text{and} & v_{\mathbf{k}} = 0 & \text{for } k > k_F \end{array} \qquad \text{(ideal gas)} \tag{93}$$

Whatever the value of **k**, one or the other of the functions $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ will be zero, and so is the product $u_{\mathbf{k}} v_{\mathbf{k}}$; the condensation energy of an interacting system remains equal to zero as long as the state of the system does not differ from the ideal gas state. A condensation energy can only be obtained via a deformation of the Fermi distribution.

It is the attractive interactions that actually distort this distribution to create an overlap between regions where both functions $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are different from zero, as can be seen in Figure 1. This allows minimizing the pairing energy (92), but involves a transfer of particles from the inside of the Fermi sphere to the outside, hence toward states of higher kinetic energy; this has a cost in terms of kinetic energy. The optimization we performed amounts to looking for the most favorable balance between the gain in potential energy and the cost in kinetic energy. The condensation energy is proportional to the square of the integral of the dashed line curve in Figure 1, which presents a maximum in the vicinity of the Fermi energy E_F. The largest contributions come from energies close to E_F, over a width of the order of a few Δ - but the figure also shows that the contributions to the condensation energy spread relatively far from the Fermi surface (the curve only decreases as the inverse of the energy's distance from its maximum). The Fermi surface, which was perfectly defined for an ideal gas, becomes blurred over a certain energy domain.

The two-particle correlation function expresses in more detail this optimization of the attractive potential energy. Relation (64) shows that, for parallel spins, no significant change of the correlation function occurs, when compared to that of an ideal gas (for which an exchange hole is already present in the binary correlation function) – hence no significant change of the corresponding interaction energy. This lack of effect comes from the fact that the BCS wave function only pairs particles with opposite spins. On the other hand, when the spin directions are opposite, relation (66) shows that the probability of presence of two particles at a short distance from each other is increased; the larger the pair wave function modulus at the origin ($\mathbf{r} = \mathbf{r}'$), the higher this increase will be. It directly yields the gain in the attractive potential energy.

In other words, the BCS gain in energy comes from the fact that, because of the interactions, the system changes its wave function to optimize its pairing potential energy. It develops correlations that go beyond that of an ideal gas; they are referred to as "dynamic correlations", as opposed to the statistical correlations (coming solely from the indistinguishability of the particles, such as those studied in Complement A_{XVI}). This produces a deformation of the ideal gas Fermi distribution that, instead of presenting an abrupt transition between the occupied and empty states (perfectly well defined Fermi sphere), presents at its edge a more progressive transition region (blurred Fermi sphere). The system's state vector then becomes a superposition of states where the particle

number in each pair of states $(\mathbf{k}, -\mathbf{k})$ fluctuates. The potential energy term that drives the BCS mechanism is the "pair annihilation-creation" term computed in § 3-a-β of Complement B_{XVII}, and that is schematized on its Figure 4. It includes a sum of terms containing non-diagonal matrix elements of the form:

$$\left\langle n_{(\mathbf{k},-\mathbf{k})} = 2;\ n_{(\mathbf{k}'',-\mathbf{k}'')} = 0 \right| W_2 \left| n_{(\mathbf{k},-\mathbf{k})} = 0;\ n_{(\mathbf{k}'',-\mathbf{k}'')} = 2 \right\rangle \tag{94}$$

(the occupation numbers of all the other pairs remaining the same); between the ket and the bra, a pair $(\mathbf{k}'', -\mathbf{k}'')$ is replaced by another one $(\mathbf{k}, -\mathbf{k})$. The BCS energy gain is due to the summation over all these non-diagonal terms; they are sensitive to the coherence of the state vector between these two components (where the numbers of pairs fluctuate in a correlated way) and hence to their relative phase.

β. Phase locking and cooperative effects

In the computation of the § 1-b, the minimization of the energy led us to choose the phases of all the $v_{\mathbf{k}}$ to be equal, and they simply disappeared from the following calculations. To discuss the physical process at work, it is useful to reintroduce them with their non-specified values before the optimization, as they appeared in (14); when all the matrix elements of the interaction potential are equal, the average value of the pairing energy is written:

$$E_{\text{paired}} = -V \sum_{\mathbf{k},\mathbf{k}'} \sin\theta_{\mathbf{k}} \cos\theta_{\mathbf{k}}\ \sin\theta_{\mathbf{k}'} \cos\theta_{\mathbf{k}'}\ e^{2i(\zeta_{\mathbf{k}'} - \zeta_{\mathbf{k}})} \tag{95}$$

We mentioned above that adding to the phase $\zeta_{\mathbf{k}}$ of each $v_{\mathbf{k}}$ any given common phase χ, left all the results unchanged. The energy is invariant with respect to a symmetry of the wave function, the one that concerns the global phase of the $v_{\mathbf{k}}$. It is often called the "$U(1)$ symmetry", referring to the $U(1)$ unitary symmetry group of rotations around a circle, isomorphic to the group of phase changes for a complex number.

On the other hand, changing one by one the phases of the $v_{\mathbf{k}}$, leads to an obvious reduction (in absolute value) of the pairing energy (95): in the complex plane, vectors that were perfectly aligned, now take different directions and the modulus of their sum is reduced. We saw that this term is at the origin of the gain in energy provided by the BCS mechanism; it is clearly linked to the acquisition of a common phase by all the pairs of individual states. This is an example of a phenomenon called in physics "spontaneous symmetry breaking": whatever the phase of each $v_{\mathbf{k}}$, which can take on any value, it is essential that it be the same for all, otherwise most of the gain in energy will be lost. In a similar way, in the ferromagnetic transition in a solid, the space direction along which the spins will align is not a priori fixed, but it is essential that it be the same[13] for all the spins.

Note finally the cooperative character of the energy gain obtained, which, mathematically, corresponds to the presence of a double summation over $\mathbf{k}$ and $\mathbf{k}'$. Starting

[13] For an ensemble of spins parallel to any given direction, each spin is in a state where the relative phase between the components on $|+\rangle$ and $|-\rangle$ is the same. For the BCS mechanism, it is the phase between the components where the occupation number of the couple of states $\mathbf{k}, -\mathbf{k}$ is 0 or 2 that takes on a value independent of $\mathbf{k}$. The corresponding energy lowering results from an interference effect between states where two pairs $\mathbf{k}$ and $\mathbf{k}'$ have respective occupation numbers $n_{\mathbf{k}} = 2, n_{\mathbf{k}'} = 0$ and $n_{\mathbf{k}} = 0, n_{\mathbf{k}'} = 2$; therefore it cannot be directly expressed in terms of pair populations.

from a perfectly phase locked situation, destroying the phase locking of a single pair leads to an energy loss proportional to the number of pairs that remained phase locked; the individual energy of a single pair is not what is at stake. On the other hand, starting from a situation where the phases of all the pairs are random, changing a single phase $\zeta_{\mathbf{k}}$ barely modifies the average energy. We are in the presence of a cooperative effect: the greater the number of pairs that are already phase locked, the higher the tendency for a new pair to become phase locked; this tendency can be seen, in a way, as resulting from a mean field created in a cumulative way by all the other pairs. Here again we see the analogy with a ferromagnetic material where, the greater the number of spins already aligned, the higher the gain in energy with the alignment of a new spin.

3-b. Gain in energy

We now compute the gain in energy resulting from the pairs' formation. We first insert in (7) relations (27), which yield the optimal values of the $u_{\mathbf{k}}$ and the $v_{\mathbf{k}}$, and use the definition (25) for E_k; we also take all the potential matrix elements to be equal to the same constant $-V$. Since we then have:

$$|v_{\mathbf{k}}|^2 = \frac{1}{2}\left(1 - \frac{\xi_k}{E_k}\right) = \frac{1}{2}\frac{E_k - \xi_k}{E_k} \tag{96}$$

and:

$$u_{\mathbf{k}} v_{\mathbf{k}} = \frac{1}{2}\sqrt{1 - \left(\frac{\xi_k}{E_k}\right)^2} = \frac{\Delta}{2E_k} \tag{97}$$

we get:

$$\begin{aligned}&\left\langle \widehat{H} - \mu \widehat{N} \right\rangle_{\mathrm{BCS}} \\ &= -\frac{V}{4}\langle N\rangle^2 + \sum_{\mathbf{k}} \left[\xi_k \, \frac{E_k - \xi_k}{E_k} - \frac{V\Delta^2}{4} \sum_{\mathbf{k},\mathbf{k}'} \frac{1}{\sqrt{(\xi_k)^2 + \Delta^2}} \frac{1}{\sqrt{(\xi_{k'})^2 + \Delta^2}} \right]\end{aligned} \tag{98}$$

The first term on the right-hand side is the mean field term (as before, we have neglected 1 compared to the total number of particles). The second one corresponds to the kinetic energy, and the third one, to the interaction between pairs:

$$\frac{-V\Delta^2}{4} \sum_{\mathbf{k},\mathbf{k}'} \frac{1}{\sqrt{(\xi_k)^2 + \Delta^2}} \frac{1}{\sqrt{(\xi_{k'})^2 + \Delta^2}} = -\frac{\Delta^2}{2} \sum_{\mathbf{k}} \frac{1}{\sqrt{(\xi_k)^2 + \Delta^2}} \tag{99}$$

where to get the second equality we have used relation (36) to eliminate the summation over $\mathbf{k}'$. Using again the definition (25) for E_k, we get:

$$\left\langle \widehat{H} - \mu \widehat{N} \right\rangle_{\mathrm{BCS}} = -\frac{V}{4}\langle N\rangle^2 + \sum_{\mathbf{k}} \frac{1}{\sqrt{(\xi_k)^2 + \Delta^2}} \left[\xi_k \left(\sqrt{(\xi_k)^2 + \Delta^2} - \xi_k \right) - \frac{\Delta^2}{2} \right] \tag{100}$$

On the right-hand side of this expression, the first term, corresponding to the mean field, is of no particular interest. The second one accounts for the change in energy due to the dynamic correlations introduced by the interactions; it characterizes the BCS mechanism.

We first check the convergence of its summation over $\mathbf{k}$, for a fixed value of Δ. This is not the case for each term in the parenthesis, which tends toward a constant when $k \to \infty$, leading to a summation in $1/\xi_k \sim 1/k^2$ that diverges. We are now going to show that the divergent terms cancel each other. We can write:

$$\sqrt{(\xi_k)^2 + \Delta^2} - \xi_k \sim \frac{\Delta^2}{2\xi_k} - \frac{3}{8}\frac{\Delta^4}{(\xi_k)^3} + .. \tag{101}$$

and thus:

$$\frac{1}{\sqrt{(\xi_k)^2 + \Delta^2}}\left[\xi_k\left(\sqrt{(\xi_k)^2 + \Delta^2} - \xi_k\right) - \frac{\Delta^2}{2}\right] \sim -\frac{3}{8}\frac{\Delta^4}{(\xi_k)^3} + .. \tag{102}$$

We have just shown that the divergent terms in the infinite summation of $\mathbf{k}$ in relation (100) cancel each other between the kinetic and interaction terms; the function in the summation goes to zero, for large k values, as $1/(\xi_k)^3 \sim 1/k^6$, which ensures the convergence and does not require the introduction of a cutoff frequency k_c (apart from the one we had to introduce before to ensure a finite value for Δ). This was also the case for the total number of particles. We thus see that once we have introduced an upper boundary (cutoff) k_c in the integral determining the gap Δ, all the other important physical quantities remain finite, without having to reimpose this cutoff frequency.

The precise determination of the energy requires, in general, the computation of somewhat complex integrals. It will not be detailed here, but yields the result:

$$\left\langle \widehat{H} - \mu\widehat{N} \right\rangle_{\text{BCS}} = \left\langle \widehat{H} - \mu\widehat{N} \right\rangle_0 - \frac{V}{4}\left[\langle N \rangle\right]^2 - \frac{1}{2}\Delta^2 D_F \tag{103}$$

(remember that D_F is the density of individual states at the surface of the Fermi sphere, and is proportional to the volume $\mathcal{V} = L^3$). Finally, the energy gain resulting from the BCS pairing is given by:

$$\delta E = -\frac{1}{2}\Delta^2 D_F \tag{104}$$

It can be shown that the values of ξ_k that contribute most to the energy are those that are lower than or comparable to the gap Δ; the energy change linked to the pairing phenomenon is mainly located in the vicinity of the Fermi surface. This result is often interpreted by saying that, in an ensemble of $\Delta \times D_F$ pairs, each pair gains an energy of the order of Δ, which explains the Δ^2 dependence of (104); while this image is simple, it has its shortcomings (see note 13).

3-c. Non-perturbative character of the BCS theory

Generally, the most basic way to take the interactions into account is to use a first order perturbation theory (Chapter XI), where the energy correction is the average value, in the initial non-perturbed state, of the perturbation Hamiltonian. Applied here, the first order correction to the energy is obtained by inserting the values (93) into (7).

The first term (kinetic energy) on the right-hand side of (7) is unchanged, and the third one remains zero since, according to (93), the product $u_{\mathbf{k}} v_{\mathbf{k}}$ is always zero for any value of $\mathbf{k}$. We are left with the second term, which produces a mean field correction. To the next perturbation order, the effect of the potential is to change the ground state by transferring pairs of particles, initially both inside the Fermi sphere, toward individual states whose momenta fall outside the sphere (all the while keeping the total momentum constant); this changes at the same time the average kinetic energy (which is increased) and the interaction potential energy. The computations become more and more complex as the perturbation order increases. And above all, it is clear that this approach to higher and higher perturbation orders cannot account for the existence of the gap obtained in (46): as the function $\Delta(V)$ has all its derivatives with respect to V equal to zero at $V = 0$, it can not be expanded as a series in V.

The BCS theory is a non-perturbative method that solves this difficulty. However, it is not an exact method since it is a variational method, but the chosen wave function is sufficiently well adapted to allow the inclusion of important physical effects, without using any perturbation theory.

4. Excited states

Up to now, we only studied the ground state of the system of attractive fermions. As soon as the temperature is no longer zero, excited states of the system begin to be populated. In this last section, we shall give a survey of the BCS theory predictions concerning the excited states. A study of the BCS theory at non-zero temperature can be found in more specialized books.

4-a. Bogolubov-Valatin transformation

Relations (E-3) and (E-4) of Chapter XVII define the Bogolubov-Valatin transformations of the creation and annihilation operators of spin 1/2 fermions. With the notation of the present complement where the spin directions are explicit, they become:

$$\begin{aligned} b_{\mathbf{k}} &= u_{\mathbf{k}}\, a_{\mathbf{k},\uparrow} - v_{\mathbf{k}}\, a^{\dagger}_{-\mathbf{k},\downarrow} \\ b_{-\mathbf{k}} &= u_{\mathbf{k}}\, a_{-\mathbf{k},\downarrow} + v_{\mathbf{k}}\, a^{\dagger}_{\mathbf{k},\uparrow} \end{aligned} \tag{105}$$

which, by conjugation, yield the definitions of the Hermitian conjugate operators $b^{\dagger}_{\mathbf{k}}$ and $b^{\dagger}_{-\mathbf{k}}$:

$$\begin{aligned} b^{\dagger}_{\mathbf{k}} &= u^{*}_{\mathbf{k}}\, a^{\dagger}_{\mathbf{k},\uparrow} - v^{*}_{\mathbf{k}}\, a_{-\mathbf{k},\downarrow} \\ b^{\dagger}_{-\mathbf{k}} &= u^{*}_{\mathbf{k}}\, a^{\dagger}_{-\mathbf{k},\downarrow} + v^{*}_{\mathbf{k}}\, a_{\mathbf{k},\uparrow} \end{aligned} \tag{106}$$

For each value of $\mathbf{k}$, we get a general transformation of the four initial creation and annihilation operators a or $a^{\dagger}$ into four new operators b and $b^{\dagger}$. We showed in Chapter XVII that these operators obey the usual anticommutation relations of fermions.

We also saw in that chapter that the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$ defined in (2):

$$|\overline{\varphi}_{\mathbf{k}}\rangle = \left[u_{\mathbf{k}} + v_{\mathbf{k}}\, a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle \tag{107}$$

is an eigenvector of the two operators $b_{\mathbf{k}}$ and $b_{-\mathbf{k}}$ with a zero eigenvalue:

$$\begin{aligned} b_{\mathbf{k}} \left|\overline{\varphi}_{\mathbf{k}}\right\rangle &= u_{\mathbf{k}} v_{\mathbf{k}} \left[a^{\dagger}_{-\mathbf{k},\downarrow} - a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle = 0 \\ b_{-\mathbf{k}} \left|\overline{\varphi}_{\mathbf{k}}\right\rangle &= u_{\mathbf{k}} v_{\mathbf{k}} \left[-a^{\dagger}_{\mathbf{k},\uparrow} + a^{\dagger}_{\mathbf{k},\uparrow}\right] |0\rangle = 0 \end{aligned} \tag{108}$$

It follows that the variational ket $|\Psi_{\mathrm{BCS}}\rangle$ of the ground state, written in (1), is an eigenvector, for any value of $\mathbf{k}$, of all the operators $b_{\mathbf{k}}$ and $b_{-\mathbf{k}}$, with a zero eigenvalue. It is therefore also an eigenket of all the operators $b^{\dagger}_{\mathbf{k}} b_{\mathbf{k}}$ and $b^{\dagger}_{-\mathbf{k}} b_{-\mathbf{k}}$ with a zero eigenvalue, which is a minimum eigenvalue for operators defined as positive or zero. Furthermore, we showed that the repeated action of the creation operators $b^{\dagger}_{\mathbf{k}}$ and $b^{\dagger}_{-\mathbf{k}}$ permits obtaining other states, which are also eigenvectors of the operators $b^{\dagger}_{\mathbf{k}} b_{\mathbf{k}}$ and $b^{\dagger}_{-\mathbf{k}} b_{-\mathbf{k}}$. We are going to show that these operators $b^{\dagger}_{\mathbf{k}} b_{\mathbf{k}}$ and $b^{\dagger}_{-\mathbf{k}} b_{-\mathbf{k}}$ can be interpreted as corresponding to the occupation numbers of the excitations present in the physical system.

4-b. Broken pairs and excited pairs

Letting $b^{\dagger}_{\mathbf{k}}$ act on (107), we get:

$$\begin{aligned} b^{\dagger}_{\mathbf{k}} \left|\overline{\varphi}_{\mathbf{k}}\right\rangle &= \left[|u_{\mathbf{k}}|^2 a^{\dagger}_{\mathbf{k},\uparrow} - |v_{\mathbf{k}}|^2 a_{-\mathbf{k},\downarrow} a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow}\right] |0\rangle = \left[|u_{\mathbf{k}}|^2 a^{\dagger}_{\mathbf{k},\uparrow} + |v_{\mathbf{k}}|^2 a^{\dagger}_{\mathbf{k},\uparrow}\right] |0\rangle \\ &= a^{\dagger}_{\mathbf{k},\uparrow} |0\rangle \end{aligned} \tag{109}$$

which is a ket normalized to unity, and obviously non-zero (as opposed to the one resulting from the action of $b_{\mathbf{k}}$). Similarly, if we consider the action of $b^{\dagger}_{-\mathbf{k}}$, we get another non-zero ket:

$$b^{\dagger}_{-\mathbf{k}} \left|\overline{\varphi}_{\mathbf{k}}\right\rangle = a^{\dagger}_{-\mathbf{k},\downarrow} |0\rangle \tag{110}$$

These two new normalized kets are orthogonal to the initial ket $|\overline{\varphi}_{\mathbf{k}}\rangle$, since they correspond to an occupation number equal to 1, whereas the occupation numbers of $|\overline{\varphi}_{\mathbf{k}}\rangle$ are 0 and 2. In these states, a pair has been replaced by a single particle, not belonging to any pair; they are called "broken pair" states.

As the squares of the operators $b^{\dagger}_{\mathbf{k}}$ and $b^{\dagger}_{-\mathbf{k}}$ are zero, the repeated application of any of the two does not allow constructing new orthogonal states; however this can be achieved with their cross product. Letting $b^{\dagger}_{\mathbf{k}}$ act on the ket (110), we get the ket $|\overline{\varphi}_{\mathbf{k}}\rangle^{e}$:

$$\left|\overline{\varphi}_{\mathbf{k}}\right\rangle^{e} = b^{\dagger}_{\mathbf{k}} b^{\dagger}_{-\mathbf{k}} \left|\overline{\varphi}_{\mathbf{k}}\right\rangle = \left[u^{*}_{\mathbf{k}} a^{\dagger}_{\mathbf{k},\uparrow} a^{\dagger}_{-\mathbf{k},\downarrow} - v^{*}_{\mathbf{k}}\right] |0\rangle \tag{111}$$

which is another normalized ket, and orthonormal, as can be easily checked, to $|\overline{\varphi}_{\mathbf{k}}\rangle$; letting now the two operators $b^{\dagger}_{\mathbf{k}}$ and $b^{\dagger}_{-\mathbf{k}}$, in the inverse order, act on $|\overline{\varphi}_{\mathbf{k}}\rangle$, we get the same ket, $|\overline{\varphi}_{\mathbf{k}}\rangle^{e}$, within a change of sign. The components of the two states $|\overline{\varphi}_{\mathbf{k}}\rangle$ and $|\overline{\varphi}_{\mathbf{k}}\rangle^{e}$ contain occupation numbers equal to 0 or 2; the ket $|\overline{\varphi}_{\mathbf{k}}\rangle^{e}$ is called "excited pair" state. To go from $|\overline{\varphi}_{\mathbf{k}}\rangle$ to $|\overline{\varphi}_{\mathbf{k}}\rangle^{e}$, we simply switch $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$, change the sign of $v_{\mathbf{k}}$, and finally take their complex conjugate (this is true for the general case, but in the BCS pairing case, as $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ take on real values, this last step becomes unnecessary).

4-c. Stationarity of the energies

Let us now show that the energies of these new states are stationary with respect to the variational parameters.

α. Broken pair

According to (109), the action of $b^\dagger_{\mathbf{k}}$ on the ground state $|\Psi_{\rm BCS}\rangle$ leads to the ket:

$$b^\dagger_{\mathbf{k}}\,|\Psi_{\rm BCS}\rangle = \left[a^\dagger_{\mathbf{k},\uparrow}\,|0\rangle\right] \otimes |\Psi'_{\rm BCS}\rangle \tag{112}$$

where $|\Psi'_{\rm BCS}\rangle$ is just the ket $|\Psi_{\rm BCS}\rangle$ whose $\mathbf{k}$ pair component has been removed from the product:

$$|\Psi'_{\rm BCS}\rangle = \prod_{\otimes \mathbf{k}'\neq\mathbf{k}} |\overline{\varphi}_{\mathbf{k}}\rangle \tag{113}$$

The energy average value in the state (112) is the sum of three terms:

(i) the kinetic energy $e_k = \hbar^2k^2/2m$ associated with the state $a^\dagger_{\mathbf{k},\uparrow}\,|0\rangle$

(ii) the energy associated with the state $|\Psi'_{\rm BCS}\rangle$; the computation of that energy is the same as for $|\Psi_{\rm BCS}\rangle$, including the pair interaction energy, except for the fact that one less pair is now involved in the calculation. This slightly modifies the value of Δ, and hence the optimal value of the parameters $\theta_{\mathbf{k}}$; however, since the relative variation of Δ is inversely proportional to the number of particles, we shall ignore this slight modification.

(iii) finally, the interaction energy between the particle in the individual state $a^\dagger_{\mathbf{k},\uparrow}\,|0\rangle$ and the particles described by $|\Psi'_{\rm BCS}\rangle$; the pair structure of that state means that the only contributions are a direct term in:

$$\begin{aligned}\frac{1}{2}\sum_{\mathbf{k}'\neq\mathbf{k},\nu} V_{\mathbf{0}}\left\langle a^\dagger_{\mathbf{k},\uparrow}a^\dagger_{\mathbf{k}',\nu}a_{\mathbf{k}',\nu}a_{\mathbf{k},\uparrow} + a^\dagger_{\mathbf{k}',\nu}a^\dagger_{\mathbf{k},\uparrow}a_{\mathbf{k},\uparrow}a_{\mathbf{k}',\nu}\right\rangle &= V_{\mathbf{0}}\sum_{\mathbf{k}',\nu}\langle\widehat{n}_{\mathbf{k}',\nu}\rangle \\ &= V_{\mathbf{0}}N'\end{aligned} \tag{114}$$

(where N' is the average particle number in the state $|\Psi'_{\rm BCS}\rangle$) and an exchange term in:

$$\frac{1}{2}\sum_{\mathbf{k}'\neq\mathbf{k}} V_{\mathbf{k}'-\mathbf{k}}\left\langle a^\dagger_{\mathbf{k},\uparrow}a^\dagger_{\mathbf{k}',\uparrow}a_{\mathbf{k},\uparrow}a_{\mathbf{k}',\uparrow} + a^\dagger_{\mathbf{k}',\uparrow}a^\dagger_{\mathbf{k},\uparrow}a_{\mathbf{k}',\uparrow}a_{\mathbf{k},\uparrow}\right\rangle = -\sum_{\mathbf{k}'\neq\mathbf{k}} V_{\mathbf{k}'-\mathbf{k}}\,\langle\widehat{n}_{\mathbf{k}',\uparrow}\rangle \tag{115}$$

Note again that, for an interaction that does not act on the spins, the exchange is only possible with particles having the same spin. With the short-range potential approximation (29):

$$V_{\mathbf{0}} = V_{\mathbf{k}} = -V \tag{116}$$

the term (114) becomes equal to $-VN'$, the term (115) to $VN'/2$, and their sum simply yields a constant $-VN'/2$.

The parameters defining the variational state (113) are the set of the $\theta_{\mathbf{k}'}$ and the $\zeta_{\mathbf{k}'}$ for $\mathbf{k}'\neq\mathbf{k}$ (the dependence on the θ_k parameter, which characterized the broken pair, is no longer present). These parameters are the ones that make the energy stationary, since we neglected the slight variation of the gap linked to the disappearance of a pair; they play no role either in (i) or in (iii).

We have just confirmed that $b^\dagger_{\mathbf{k}}\,|\Psi_{\rm BCS}\rangle$ renders the energy stationary (in the framework of the variational approximation we are using). A symmetry argument shows that this is obviously also the case for the state $b^\dagger_{-\mathbf{k}}\,|\Psi_{\rm BCS}\rangle$.

β. Excited pair

In the stationary relations (26), the change of $|\overline{\varphi}_{\mathbf{k}}\rangle$ into $|\overline{\varphi}_{\mathbf{k}}\rangle^e$ simply amounts to exchanging the signs ± (the $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ are real); the components of the ket $|\overline{\varphi}_{\mathbf{k}}\rangle^e$ are part of the stationary solutions we have discarded in writing (27). We thus confirm that the excited pair corresponds to a stationary energy; it is actually the highest possible energy for the pair of states $(\mathbf{k}, -\mathbf{k})$.

4-d. Excitation energies

In the 4-dimensional state space associated with each pair of states $(\mathbf{k}, -\mathbf{k})$, the creation operators $b^\dagger$ acting on the $|\overline{\varphi}_{\mathbf{k}}\rangle$ permit building a new basis of 4 orthonormal states, whose average energies are stationary. They can be considered to be the approximate eigenvectors of the system Hamiltonian. We now compute the corresponding eigenvalues.

In the case of the broken pair, the excited state $|\Psi'_{\mathrm{BCS}}\rangle$ does not contain the same number of particles as $|\Psi_{\mathrm{BCS}}\rangle$; it does not make sense to directly compare their energies. To make a valid comparison, we must take into account the presence of a particle reservoir whose energy increases by μ (chemical potential) each time it absorbs a particle. In other words, we must evaluate the variations of the average value $\left\langle \widehat{H} - \mu\widehat{N} \right\rangle$.

α. Broken pair

We now show that the variation of the average value $\left\langle \widehat{H} - \mu\widehat{N} \right\rangle$ associated with the breaking of the pair is simply the energy E_k defined in (25).

To do so, we compute the variation of this average value when the state $|\Psi_{\mathrm{BCS}}\rangle$ is replaced by expression (112). Several terms come into play:

(i) The variation of the average value of the kinetic energy is the difference between the energy e_k of a particle and that of the population $|v_{\mathbf{k}}|^2$ of the pair $(\mathbf{k}, -\mathbf{k})$, with a kinetic energy of $2e_k$; this difference is therefore $e_k\left(1 - 2|v_{\mathbf{k}}|^2\right)$.

(ii) As for the potential energy, the passage from $|\Psi_{\mathrm{BCS}}\rangle$ to the ket $|\Psi'_{\mathrm{BCS}}\rangle$ changes the average particle number $\langle N\rangle$ from $2|v_{\mathbf{k}}|^2$ (initial population of the pair) to 0, so that the variation of $\langle N\rangle$ is $\delta\langle N\rangle = -2|v_{\mathbf{k}}|^2$. The mean field term $-V\langle N\rangle^2/4$ in (7) varies by $-V\langle N\rangle\delta\langle N\rangle/2$, that is $V\langle N\rangle|v_{\mathbf{k}}|^2$. The following term in (7) is zero for a short-range potential. Finally, the breaking of a pair has an impact on the binding energy in the last term of (7); if we change the dummy summation index k into k', and k' into k'', the terms that will change correspond to the terms $k = k'$ and the terms $k = k''$, which double each other. The breaking of the pair leads to an increase of energy equal to:

$$2Vu_{\mathbf{k}}v_{\mathbf{k}}\sum_{\mathbf{k}'}u_{\mathbf{k}'}v_{\mathbf{k}'} = 2\Delta\, u_{\mathbf{k}}v_{\mathbf{k}} = \Delta\sqrt{1 - \frac{\xi_k^2}{(E_k)^2}} = \frac{\Delta^2}{E_k} \tag{117}$$

where we have used (31) and (27).

(iii) We saw that the unpaired particle has a potential energy $-VN'/2 \simeq -V\langle N\rangle/2$. This energy must be added to the variation of the mean field term calculated above, to give a contribution equal to $-V\langle N\rangle\left(1 - 2|v_{\mathbf{k}}|^2\right)/2$.

We now sum all the previous contributions and add the variation of $-\mu \left\langle \widehat{N} \right\rangle$ which yields a term $-\mu \left(1 - 2\left|v_{\mathbf{k}}\right|^2\right)$. The total variation is then:

$$\delta \left\langle \widehat{H} - \mu \widehat{N} \right\rangle = \left(e_k - \frac{V \langle N \rangle}{2} - \mu \right) \left(1 - 2\left|v_{\mathbf{k}}\right|^2\right) + \frac{\Delta^2}{E_k} \tag{118}$$

or else, taking (13) into account:

$$\begin{aligned} \delta \left\langle \widehat{H} - \mu \widehat{N} \right\rangle &= \xi_k \left[1 - \left(1 - \frac{\xi_k}{E_k} \right) \right] + \frac{\Delta^2}{E_k} \\ &= \frac{\xi_k^2 + \Delta^2}{E_k} = E_k \end{aligned} \tag{119}$$

We find the expected result[14].

β. Excited pair

We now assume that in the product that yields $|\Psi_{\text{BCS}}\rangle$, the ket $|\overline{\varphi}_{\mathbf{k}}\rangle$ is replaced by the orthogonal ket $|\overline{\varphi}_{\mathbf{k}}\rangle^e$ written in (111), and which describes an "excited pair". We are going to show that the variation of the average value $\left\langle \widehat{H} - \mu \widehat{N} \right\rangle$ associated with that excitation is $2E_k$, twice the excitation associated with the breaking of the pair.

To show this, we must, here again, add several variations. The first one comes from the kinetic energy and yields $2e_k \left(\left|u_{\mathbf{k}}\right|^2 - \left|v_{\mathbf{k}}\right|^2\right)$, that is $2e_k \left(1 - 2\left|v_{\mathbf{k}}\right|\right)^2\big)$ – see relation (4). The second one is the mean field term introduced by the fact that the average value of the total particle number varies by $2\left(\left|u_{\mathbf{k}}\right|^2 - \left|v_{\mathbf{k}}\right|^2\right)$, which leads to a potential energy variation $-V \langle N \rangle \left(1 - 2\left|v_{\mathbf{k}}\right|^2\right)$. We must also account for a variation of the pair binding energy, which comes from the sign change of the product $u_{\mathbf{k}} v_{\mathbf{k}}$, which doubles the term (117). Because of the change in the average particle number, the term in $\mu \langle N \rangle$ gives a contribution of $2\mu \left(\left|u_{\mathbf{k}}\right|^2 - \left|v_{\mathbf{k}}\right|^2\right)$, that is $2\mu \left(1 - 2\left|v_{\mathbf{k}}\right|^2\right)$. Finally, all the terms found for the broken pair are just doubled here and we indeed find $2E_k$.

γ. Spectrum of the elementary excitations

We now have the energy of the three excited states associated with each pair of states: the energy E_k (doubly degenerate since it corresponds to both kets $b_{\mathbf{k}}^\dagger |\overline{\varphi}_{\mathbf{k}}\rangle$ and $b_{-\mathbf{k}}^\dagger |\overline{\varphi}_{\mathbf{k}}\rangle$) and the energy $2E_k$ (non-degenerate). The value of these energies is given in (25). Figure 7 plots the dispersion relation of these elementary excitations (energy of these excitations as a function of their momentum). The solid line corresponds to the spectrum associated with the breaking of a pair $\mathbf{k}$, $-\mathbf{k}$, during which a particle disappears, as in relations (109) and (110); as the spectrum associated with the excitation of a pair can be simply obtained by the multiplication by a factor 2, it is not plotted in the figure. The dashed lines plot those same energies for an ideal gas (no interaction) for which $\Delta = 0$. The interaction effect creates a "gap" Δ, which yields a minimum value for the excitation energy that otherwise can go to zero for an ideal gas.

Upon the breaking of a pair, we saw that, on average, the system's total population changes by $1 - 2\left|v_{\mathbf{k}}\right|^2$. The curve in Figure 7 has a different interpretation, depending on

[14] The calculation would be the same for an ideal gas; in the particular case where $\Delta = 0$ and according to (25), we would get: $E_k = |\xi_k|$.

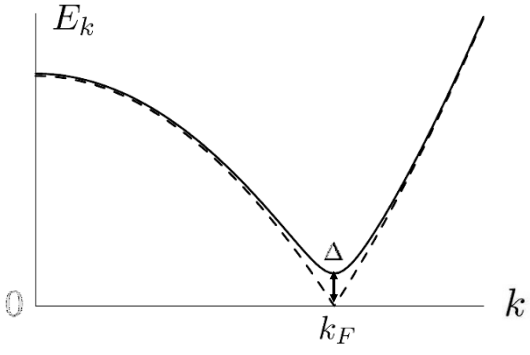


Figure 7: The solid line plots the variation as a function of k of the excitation energy $E_k = \sqrt{(\xi_k)^2 + \Delta^2}$, with $\xi_k = e_k - \mu$ and $e_k = \hbar^2 k^2/2m$ (for the sake of simplicity, we assumed $\mu \simeq \overline{\mu}$, so as to neglect the difference between $\overline{\mu}$ and μ in the expression of ξ_k). This energy presents a minimum equal to the gap Δ when k is equal the the Fermi wave vector $k = k_F$ (wave vector for which $e_k = \mu$). The dashed line curve is the same function, but for $\Delta = 0$ (zero gap), hence for an ideal gas.
The BCS theory predicts that the energy associated with the breaking of a pair is E_k, and the energy associated with the excitation of a pair is twice that amount, i.e. $2E_k$. The minimum of the energy E_k is therefore the minimum of energy that must be supplied to the system in its BCS ground state to produce one of the previously computed excitations. As explained in the text, the region on the left-hand side of the curve corresponds to "hole type" excitations (the excitation leads to a particle loss for the physical system) and the region on the right-hand side to "particle type" excitations (the physical system gains a particle).
For clarity, the figure does not take into account the mean field effects; these would lower all the energies by the same negative quantity, and change the chemical potential μ into $\overline{\mu}$ defined by (41), so that ξ_k would become $\xi_k = e_k - \overline{\mu}$. These effects slightly shift the curve plotting E_k to the left.

which part of the curve we analyze. On the left-hand side (decreasing function), the solid line curve nearly perfectly matches the dashed curve, meaning $|v_{\mathbf{k}}|^2$ is practically equal to 1 (see Figure 1). In that case, $1 - 2|v_{\mathbf{k}}|^2 \simeq -1$: a particle disappears in the course of the excitation, which is said to be of the "hole type". Its energy μ is the energy needed to push one particle towards the reservoir that fixes the chemical potential, diminished by its initial kinetic energy e_k, and to which we must add the mean field correction[15]; the excitation energy is therefore $\overline{\mu} - e_k$, a value that corresponds to the left-hand side of the curve. As for the right-hand side part (increasing function), the constant $|v_{\mathbf{k}}|^2$ is practically zero: the excitation adds a particle to the system, and is said to be of the "particle type". Its energy is equal to $e_k - \mu$, energy necessary to promote a particle from its energy μ to a state of kinetic energy e_k (with, as before, a mean field correction

[15]This correction changes the initial energy e_k into $e_k - V\langle N\rangle/2$, and hence $\mu - e_k$ into $\overline{\mu} - e_k$.

that changes μ into $\overline{\mu}$). Finally, for the central part of the curve, we have a "mixed" excitation, of both hole and particle type; it is the region of the spectrum where the solid line parts the most from the dashed line, and where the BCS mechanism, which creates the gap Δ plays an essential role.

From these four energy levels associated with each pair of states, quantum statistical mechanics allows obtaining a density operator describing the thermal equilibrium of the system at temperature T, as well as all the various thermodynamic functions. The corresponding development will not be exposed in this complement. We shall simply mention that it allows extending the validity of a certain number of results obtained at $T = 0$, by simply introducing a gap $\Delta(T)$ that depends on the temperature. At zero temperature, $\Delta(0)$ is still given by (46), but the gap decreases as T increases, and goes to zero for a certain critical temperature T_c. This cancelling of the gap corresponds to a phase transition: as a system of attractive fermions is cooled down, when it reaches a certain temperature the pair condensation phenomenon occurs, which leads to a number of physical consequences. For example, the system's specific heat first takes on values higher than in the absence of transition, then abruptly (exponentially) goes to zero as $T \to 0$.

Conclusion

In conclusion, the choice of a variational basis of paired states sheds new light on the behavior of an ensemble of attractive fermions. We focused on the case of weak attractions, corresponding to electrons in superconducting metals; in such a situation, $(\Delta/E_F) \ll 1$ and relation (75) shows that the pair range ξ_{pair} is very large compared to the distance between fermions. The one-particle distribution shown in Figure 1 is then very similar to the step function obtained for an ideal gas at zero temperature, the step being nevertheless rounded off over an energy band of width equal to Δ. In other words, the BCS pairing only slightly modifies the Fermi sphere of a perfect gas. Studying the properties of the optimal state, we were able to expose a number of important phenomena: existence of spatial dynamic correlations, which explain the increase of the average attraction (negative) energy between fermions of opposite spins; phase locking accounting for the cooperative aspect of the pairing (increase of the attractive energy overcoming the increase of kinetic energy, hence leading to a decrease of the system total energy); existence of a pair wave function describing fermions of opposite spins, and reminiscent of the Cooper pairs (Complement $\mathrm{D_{XVII}}$); appearance of a "gap" in the elementary excitation spectrum explaining the robustness of the system's ground state.

Another interesting limiting case concerns strong attractive interactions where the pair range becomes very small compared to the distance between particles. "Molecules" are then really formed with a binding energy $-e_L$ large (in absolute value) compared to the Fermi energy E_F. Saying that the size of the bound state is small compared to the distance between particles amounts to saying that its momentum distribution width is large compared to the Fermi wave vector k_F. This means that due to the attractive potential, the occupation of the individual states are spread over a large number of different momenta, which dilutes the effects of Pauli exclusion principle; these effects thus become negligible whereas they are essential in the BCS case. Instead of being positive, the chemical potential is now negative, close to $-e_L$. Relations (24) then show that the θ_k (and hence the populations of the individual states $\mathbf{k}$) always remain small, nevertheless

extending up to energies e_k of the order of e_L. In that special case, the pair wave function ϕ_{pair}, with its Fourier components $u_k v_k = \sin\theta_k \cos\theta_k e^{2i\zeta_k}$, practically coincides with the wave function having as Fourier components $(v_k/u_k) = \text{tg}\theta_k e^{2i\zeta_k}$, and which was initially used for building the paired state. As the molecules contain two strongly bound fermions, they behave as composite bosons (Complement A_{XVII}, § 3), which may undergo Bose-Einstein condensation. Paired states enable us to see the continuous passage from one limiting case (BCS situation with a weakly perturbed Fermi sphere) to the other (condensation of strongly bound "molecules"). A detailed discussion of this continuous passage and its physical consequences is given in §4-6 of reference [10] and in [11].

In this complement, we emphasized the physical interpretation of the results obtained in the detailed calculations we presented; this will give the reader the necessary base for studying the experimental aspects of superconductivity, which are not presented here. Among the many aspects of superconductivity that have not been studied in this complement, we can list: transport phenomena and the disappearance of the electrical resistance; behavior in the presence of a magnetic field (Meissner-Ochsenfeld effect); experimental study of the elementary excitation spectrum and gap measurements via different methods (tunnel effect, magnetic resonance); Josephson effect.

The interested reader can refer, for example, to the books of M. Tinkham [12], of R.D. Parks [13], or of A.J. Leggett already quoted [8]. The book of Combescot and Shiau [14] presents a good overview of the four main theoretical methods for studying superconductivity, the BCS variational method discussed in this complement being the first of this list.

Complement D_{XVII}

Cooper pairs

1 **Cooper model** **1927**
2 **State vector and Hamiltonian** **1927**
3 **Solution of the eigenvalue equation** **1929**
4 **Calculation of the binding energy for a simple case** **1929**

In this complement, we present the "Cooper model" which was a first step towards the complete BCS theory. It yields some results of that theory without having to deal with the difficulties inherent to an N-body problem. With this simplified model, we study the properties of two attractive fermions whose wave function, in the momentum representation, is excluded from the Fermi sphere. The model will be presented in § 1, where we show the existence of a bound state, occurring only because of the existence of that sphere. Furthermore, we shall see that the mathematical expression for the corresponding binding energy is reminiscent of the expression for the gap value Δ obtained in the BCS theory.

1. Cooper model

Among a large ensemble of identical fermions, we focus our attention on two of them, supposed to attract each other, in order to study their two-body wave function and energy levels. The presence of all the other fermions is simply accounted for by a Fermi sphere that, because of the Pauli exclusion principle, requires the components of that wave function to be zero inside that sphere. Such an approach is obviously not very rigorous: isolating two fermions among a large number of other indistinguishable fermions does not make much sense. Furthermore, it is hard to imagine why two of them would interact via an attractive potential, whereas all the others determining the Fermi sphere would be without interaction. However, the mathematical form of the results obtained with this model presents interesting similarities with the variational method where all the fermions are treated equally; it is thus useful to study this model.

2. State vector and Hamiltonian

Consider two attractive fermions in a singlet spin state $|S=0\rangle$:

$$|S=0\rangle = \frac{1}{\sqrt{2}}\left[|1:\uparrow\rangle\,|2:\downarrow\rangle - |1:\downarrow\rangle\,|2:\uparrow\rangle\right] \tag{1}$$

The relative motion of their position variables is described by the orbital ket $|\Psi_{\text{orb}}\rangle$, and their center of mass is described by a zero momentum ket $|\Phi_{\mathbf{K}=\mathbf{0}}\rangle$. Their state vector is:

$$|\Psi\rangle = |\Phi_{\mathbf{K}=\mathbf{0}}\rangle \otimes |\Psi_{\text{orb}}\rangle \otimes |S=0\rangle \tag{2}$$

The state $|\Psi_{\text{orb}}\rangle$ is characterized by a wave function $\Psi_{\text{orb}}(\mathbf{r})$:

$$\Psi_{\text{orb}}(\mathbf{r}) = \langle \mathbf{r} | \Psi_{\text{orb}} \rangle \tag{3}$$

where:

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \tag{4}$$

is the difference between the positions of the two particles (relative position). As the singlet state is even with respect to the exchange of the two particles, their fermionic character requires the wave function $\Psi_{\text{orb}}(\mathbf{r})$ to be even with respect to particle exchange, i.e. with respect to a sign change of $\mathbf{r}$:

$$\Psi_{\text{orb}}(-\mathbf{r}) = \Psi_{\text{orb}}(\mathbf{r}) \tag{5}$$

We assume the operator describing the attractive interaction between the two particles to be independent of the spin. As in § B-2 of Chapter VII, we separate in the the two-particle Hamiltonian H the motion of the center of mass from the relative motion, and assume the center of mass is at rest. We are then left with a Hamiltonian H_{rel} that only acts on the space of the relative motion variables, and can be written:

$$H_{\text{rel}} = \frac{\hat{\mathbf{p}}^2}{m} + V(\hat{\mathbf{r}}) \tag{6}$$

where $\hat{\mathbf{r}} = \hat{\mathbf{r}}_1 - \hat{\mathbf{r}}_2$ is the operator associated with the relative position of the two particles; $\hat{\mathbf{p}}$ is the operator associated with the momentum of the relative motion, defined as a function of the momenta $\hat{\mathbf{p}}_1$ and $\hat{\mathbf{p}}_2$ of the two particles:

$$\hat{\mathbf{p}} = \frac{\hat{\mathbf{p}}_1 - \hat{\mathbf{p}}_2}{2} \tag{7}$$

As mentioned above, we assume the presence of an ensemble of non-interacting fermions, whose Fermi level is E_F. We must then solve the eigenvalue equation:

$$H_{\text{rel}} |\Psi_{\text{orb}}\rangle = (E + 2E_F) |\Psi_{\text{orb}}\rangle \tag{8}$$

when $|\Psi_{\text{orb}}\rangle$ does not have any component inside the Fermi sphere of radius k_F, related to the Fermi level E_F by:

$$E_F = \frac{\hbar^2 (k_F)^2}{2m} \tag{9}$$

In relation (8), E is the eigen-energy with respect to twice the Fermi level. It is indeed natural to take $2E_F$ as an energy reference; this is the minimal energy to be given to the two fermions under study, for their wave function $|\Psi_{\text{orb}}\rangle$ to have zero components inside the Fermi sphere, in the absence of interaction. With this convention for the energy origin, E simply reflects the effect of the attractive interaction.

3. Solution of the eigenvalue equation

We now expand $|\Psi_{\text{orb}}\rangle$ on the normalized eigenvectors $|\mathbf{k}\rangle$ (plane waves) of the momentum $\hat{\mathbf{p}}$:

$$|\Psi_{\text{orb}}\rangle = \sum_{\mathbf{k}} g_{\mathbf{k}} |\mathbf{k}\rangle \tag{10}$$

Projected onto $|\mathbf{k}\rangle$, the eigenvalue equation (8) reads:

$$\left(2e_k \; g_{\mathbf{k}} + \sum_{\mathbf{k}'} \langle \mathbf{k}| \widehat{V} |\mathbf{k}'\rangle \;\; g_{\mathbf{k}'} \right) = (E + 2E_F)\, g_{\mathbf{k}} \tag{11}$$

where we have set, as usual:

$$e_k = \frac{\hbar^2 k^2}{2m} \tag{12}$$

The absence of components of $|\Psi_{\text{orb}}\rangle$ inside the Fermi sphere leads to the relation:

$$g_{\mathbf{k}} = 0 \qquad \text{if} \quad |\mathbf{k}| \leq k_F \tag{13}$$

while (11) becomes:

$$[E + 2\,(E_F - e_k)]\, g_{\mathbf{k}} = \sum_{|\mathbf{k}'|>k_F} V_{\mathbf{k}\mathbf{k}'} \; g_{\mathbf{k}'} \tag{14}$$

The matrix elements of the interaction operator $\widehat{V}$ are noted $V_{\mathbf{k}\mathbf{k}'}$:

$$V_{\mathbf{k}\mathbf{k}'} = \langle \mathbf{k}| \widehat{V} |\mathbf{k}'\rangle \tag{15}$$

4. Calculation of the binding energy for a simple case

Let us further simplify the model and assume that the potential matrix elements $V_{\mathbf{k}\mathbf{k}'}$ are such that:

$$\begin{aligned} &V_{\mathbf{k}\mathbf{k}'} = -V \quad \text{if} \;\; |\mathbf{k}|\,, |\mathbf{k}'| \leq k_F + \Delta k \\ &V_{\mathbf{k}\mathbf{k}'} = 0 \qquad\qquad\qquad\quad \text{otherwise} \end{aligned} \tag{16}$$

where Δk defines a wave vector domain $\Delta k \ll k_F$; these matrix elements can therefore be factored. Note that the minus sign in front of the constant V was introduced to ensure that, for our present attractive potential, the constant V is positive. When $|\mathbf{k}| \leq k_F + \Delta k$, the summation on the right-hand side of (14) becomes a constant $-SV$ independent of $\mathbf{k}$, with:

$$S = \sum_{k_F<|\mathbf{k}'|\leq k_F+\Delta k} g_{\mathbf{k}'} \tag{17}$$

whereas, if $|\mathbf{k}| > k_F + \Delta k$, this summation is zero. The solution of this equation is simply:

$$\begin{aligned} &g_{\mathbf{k}} = \frac{-SV}{E + 2\,(E_F - e_k)} \qquad \text{if } k_F < |\mathbf{k}| \leq k_F + \Delta k \\ &g_{\mathbf{k}} = 0 \qquad\quad \text{otherwise} \end{aligned} \tag{18}$$

We must add the self-consistent condition we get from inserting this solution in the definition (17) of S:

$$S = \sum_{k_F<|\mathbf{k}'|\leq k_F+\Delta k} \frac{-SV}{E + 2\left(E_F - e_{k'}\right)} \tag{19}$$

that is, changing the sign of the denominator:

$$\frac{1}{V} = \sum_{k_F<|\mathbf{k}'|\leq k_F+\Delta k} \frac{1}{2\left(e_{k'} - E_F\right) - E} \tag{20}$$

This condition is also an implicit equation for obtaining the energy E. Assuming the system is enclosed in a cubic box with a very large edge length L, the discrete summation can be replaced by an integral, and we get:

$$\frac{1}{V} = \frac{L^3}{2\pi^2} \int_{k_F}^{k_F+\Delta k} \frac{k^2 dk}{2\left(e_k - E_F\right) - E} \tag{21}$$

We now choose, as the integral variable, the variable x:

$$x = e_k - E_F \tag{22}$$

As $\mathrm{d}e_k = \hbar^2 k \mathrm{d}k/m$, this integral now includes a density of states $D\left(e_k\right)$:

$$D\left(e_k\right) = \frac{L^3}{2\pi^2} k^2 \frac{\mathrm{d}k}{de_k} = \frac{L^3}{2\pi^2} \frac{m}{\hbar^2} k \tag{23}$$

or:

$$D\left(e_k\right) = \frac{L^3}{2\pi^2} \frac{m}{\hbar^3} \sqrt{2me_k} \tag{24}$$

The implicit equation for E then becomes:

$$\frac{1}{V} = \int_0^{\Delta E} \mathrm{d}x \frac{D\left(E_F + x\right)}{2x - E} \tag{25}$$

where the upper bound ΔE is defined by:

$$E_F + \Delta E = \frac{\hbar^2}{2m} \left(k_F + \Delta k\right)^2 \tag{26}$$

As we assumed $\Delta k \ll k_F$, we can replace[1] in (25) the density of states $D\left(E_F + x\right)$ by $D\left(E_F\right)$, no longer dependent on the variable x, and that we simply note D_F:

$$D_F = D\left(E_F\right) \tag{27}$$

We can then perform the integration, which yields:

$$\begin{aligned} \frac{1}{V} &= \frac{D_F}{2} \left\{\ln\left(2x - E\right)\right\}_0^{\Delta E} = \frac{D_F}{2} \ln\left[\frac{2\Delta E - E}{-E}\right] \\ &= -\frac{D_F}{2} \ln\left[\frac{-E}{2\Delta E - E}\right] \end{aligned} \tag{28}$$

[1] Replacing k by k_F in (23), one can easily compute an order of magnitude for the density of states at the Fermi level. We find $D\left(E_F\right) \sim \langle N \rangle / E_F$, hence a value proportional to the average particle number.

We then have:

$$\frac{-E}{2\Delta E - E} = e^{-2/VD_F} \tag{29}$$

The solution of this equation for E is:

$$E = -2\Delta E \ \frac{e^{-2/VD_F}}{1 - e^{-2/VD_F}} \tag{30}$$

which, when $VD(E_F) \ll 1$, can be simplified to:

$$E = -2\Delta E \ \exp\left[-2/VD_F\right] \tag{31}$$

We obtain a negative energy (with respect to $2E_F$), as expected for a bound state (the wave function can be normalized). As we cannot make a series expansion of the function $\exp(-1/x)$ in the vicinity of $x = 0$, this energy cannot be expressed as a power series of the interaction potential V, since all its derivative are zero at $x = 0$; consequently, this energy cannot be obtained by an ordinary perturbation calculation.

Note also that the energy E goes to zero (through negative values) if the density of states $D(E_F)$ goes to zero, i.e. if E_F goes to zero: the existence of the bound state is therefore linked to the presence of the Fermi sphere, whose role is to introduce a non-zero density of states. If the Fermi sphere disappears, so does the bound state.

We find results similar to those found in Complement C_{XVII} using the BCS theory, and in particular to the expression (46) of that complement, which yields the gap Δ. To obtain that expression, we had to introduce an upper bound $\hbar\omega_D$ for the variation (in absolute value) of the energies around the Fermi energy; this upper bound plays a role comparable to that played by the energy ΔE we just introduced in (25). We simply have to assume that $\hbar\omega_D = \Delta E$ for the two results to become quite similar as they only differ by a factor 2 in the exponential (the sign difference simply comes from the fact that the gap Δ was defined as a positive quantity, whereas a binding energy is negative). The interest of the Cooper model is to clearly highlight the essential role played by the density of states D_F (in the vicinity of the Fermi level) in the creation of the gap Δ in the BCS theory.

Complement E_{XVII}

Condensed repulsive bosons

1 Variational state, energy 1935
1-a Variational ket 1935
1-b Total energy 1936
1-c $N_0 \gg N_e$ approximation 1936
2 Optimization 1937
2-a Stationarity conditions 1938
2-b Solution of the equations 1939
3 Properties of the ground state 1940
3-a Particle number, quantum depletion 1940
3-b Energy 1942
3-c Phase locking; comparison with the BCS mechanism 1944
3-d Correlation functions 1946
4 Bogolubov operator method 1950
4-a Variational space, restriction on the Hamiltonian 1951
4-b Bogolubov Hamiltonian 1952
4-c Constructing a basis of excited states, quasi-particles 1954

In this complement we study the properties of an ensemble of repulsively interacting bosons[1], undergoing Bose-Einstein condensation. We know that, for an ideal gas, a system of bosons in its ground state is totally condensed: a single individual quantum state, corresponding to the lowest energy, is occupied by all the particles. In the presence of short-range interactions, and for a sufficiently diluted system, one expects its properties to remain close to that of an ideal gas, and in particular that a large fraction of the particles still occupy the same individual quantum state. We consider this to be the case for the system under study, and that the population of one individual quantum state is much larger than all the others. We shall assume that each of the states obeys the periodic boundary conditions in a box of side length L (Complement C_{XIV}), and that the state with the large population[2] is the state $\mathbf{k} = \mathbf{0}$ (whose momentum $\mathbf{p} = \hbar\mathbf{k}$ and kinetic energy are zero). Consequently, if N_0 is the average value of the population of this zero momentum level, we assume that:

$$N_0 \gg n_{\mathbf{k}} \qquad (\text{for any } \mathbf{k} \neq \mathbf{0}) \tag{1}$$

[1] We will not consider the case of attractive interactions, as they lead to an unstable physical state – see § 4-b of Complement H_{XV}.

[2] This hypothesis simplifies the writing of the equations, but is not essential; in the case where it is a state of non-zero momentum $\mathbf{k}_0$ that is highly populated, one can go to the reference frame where this momentum is zero. This amounts to adding, in the initial frame, $\mathbf{k}_0$ to all the wave vectors appearing in the equations.

where $n_{\mathbf{k}} = \langle \widehat{n}_{\mathbf{k}} \rangle$ is the average number of particles occupying an individual state $\mathbf{k} \neq \mathbf{0}$. The total particle number N is:

$$N = N_0 + \sum_{\mathbf{k} \neq \mathbf{0}} n_{\mathbf{k}} \tag{2}$$

We have already used, in Complement C_{XV}, a first approximation to study the ground state of a condensed boson system: we described the state of the N-particle system as the product of N identical individual state vectors. This led to the Gross-Pitaevskii equation. This approach implies that only one single individual state $\mathbf{k} = \mathbf{0}$ is occupied ($N_0 = N$ and $n_{\mathbf{k}} = 0$ for any $\mathbf{k}$), as for an ideal gas. This obviously cannot be exact: it is clear that the interactions introduce dynamic correlations between the particles, which cannot be accounted for by a state vector that is a simple product of individual kets (hence without correlations). Actually, the effect of the interaction potential on the ground state is to transfer at least a fraction of the particles[3] from the state $\mathbf{k} = \mathbf{0}$ to the states $\mathbf{k} \neq \mathbf{0}$; a model involving only one individual state is necessarily limited to the case where the potential effect is very weak, and hence $N_0 \simeq N$.

In Complement E_{XV}, we introduced another approximation, based on the Hartree-Fock method; it is more general than the previous one as it allows taking into account a non-zero temperature. However, it still implies that each particle moves in the mean field created by all the others, ignoring the dynamic correlations; its description of the ground state is no better than the one derived from the Gross-Pitaevskii equation. Furthermore, this latter method proved to be problematic for a boson system undergoing Bose-Einstein condensation: we noted in § 3-b-β of Complement G_{XV} that, for a system of condensed bosons, the Hartree-Fock approximation predicts, at the grand canonical equilibrium, very large fluctuations of the number of condensed particles. In the real world, these fluctuations are strongly limited by the repulsion between the particles, which clearly indicates that the predictions of the Hartree-Fock approximation concerning fluctuations are non-physical.

In the present complement, we shall try to address these two problems: on one hand, we shall take into account the dynamic correlations introduced by the interactions in the physical system; on the other, we shall not let the number of condensed particles fluctuate arbitrarily. We will use a variational method, choosing a variational state that takes into account the binary correlations between the particles, but does not introduce unrealistic fluctuations of the particle number. This variational state will be built with the help of a paired state, enabling us to directly use the results of Chapter XVII. We will add an extra component, to account for the Bose-Einstein condensation in the individual $\mathbf{k} = \mathbf{0}$ state. Obviously, this is still not an exact calculation, as it involves a variational approximation, but it allows describing a physical situation more complex than the simple Gross-Pitaevskii approximation. This approach also highlights the many analogies, but also the differences, between the pairing of condensed bosons and the pairing of fermions.

In a general way, this complement illustrates how variational methods allow changing the correlations between particle pairs. When dealing with binary interactions, as in a standard Hamiltonian, these correlations determine the average value of the potential energy (Chapter XV, § C-5-b-α). The higher order correlations (ternary, etc.) may be present and play a role in the system; but they are not directly involved in the energy.

[3] This phenomenon is traditionally called "quantum depletion" and will be discussed in more detail in § 3-a.

This is why using the paired states to optimize only the binary correlations can lead to fairly good results.

We introduce in § 1 the paired variational state depending on a certain number of parameters, and compute the corresponding average energy. In § 2, we shall search for the optimal values of these parameters that minimize this energy, using an approximation where $N_0 \simeq N$ so that we can neglect the interactions between the particles in the $\mathbf{k} \neq \mathbf{0}$ individual states. In § 3 we study the physical properties of the state thus obtained, such as the number of particles that are not in the $\mathbf{k} = \mathbf{0}$ state, the energy, and the correlation functions. We shall then develop, in § 4, a different point of view, the Bogolubov operator method. We shall choose a larger variational space, and use the results of § E in Chapter XVII to get the Bogolubov Hamiltonian, which can be directly diagonalized. This will confirm a certain number of previously obtained results. The reader only interested in the Bogolubov operator method can go directly to this paragraph, which is fairly self-contained. The conclusion of this complement will sum up the results obtained and the limits of this approximation method.

1. Variational state, energy

We are now going to directly apply the results of Complement $\mathrm{B}_{\mathrm{XVII}}$, for the choice of the variational ket as well as for the computation of its average energy.

1-a. Variational ket

The (normalized) variational ket is of the form:

$$|\Phi_B\rangle = |\varphi_{\mathbf{0}}\rangle \otimes |\Psi_{\text{paired}}\rangle = |\varphi_{\mathbf{0}}\rangle \otimes \prod_{\otimes \mathbf{k} \in D} |\overline{\varphi}_{\mathbf{k}}\rangle \tag{3}$$

where the subscript B refers to the name Bogolubov. In this expression, $|\Psi_{\text{paired}}\rangle$ is the paired state for spinless particles written in (B-8) of Chapter XVII, which is a tensor product of the normalized states (C-13):

$$|\overline{\varphi}_{\mathbf{k}}\rangle = \frac{1}{\cosh\theta_{\mathbf{k}}} \exp\left[-x_{\mathbf{k}}\ a^{\dagger}_{\mathbf{k}} a^{\dagger}_{-\mathbf{k}}\right] |0\rangle \tag{4}$$

with:

$$x_{\mathbf{k}} = \tanh\theta_{\mathbf{k}}\ e^{2i\zeta_{\mathbf{k}}} \qquad (\ \theta_{\mathbf{k}} \geq 0) \tag{5}$$

The domain D of the tensor product in (3) is half the $\mathbf{k}$-space, which prevents (as we saw in Chapter XVII, § B-2-a) the double appearance of each state $|\overline{\varphi}_{\mathbf{k}}\rangle = |\overline{\varphi}_{-\mathbf{k}}\rangle$; the origin $\mathbf{k} = \mathbf{0}$ is excluded from D.

As for $|\varphi_{\mathbf{0}}\rangle$, it is the coherent state already used in Complement $\mathrm{B}_{\mathrm{XVII}}$, relation (44):

$$|\varphi_{\mathbf{0}}\rangle = e^{-N_0/2}\ e^{\alpha_0 a^{\dagger}_0} |n_0 = 0\rangle \tag{6}$$

This state depends on a complex parameter α_0, characterized by its modulus $\sqrt{N_0}$ and its phase ζ_0:

$$\alpha_0 = \sqrt{N_0} e^{i\zeta_0} \tag{7}$$

It is the normalized eigenvector of the operator a_0 with eigenvalue α_0:

$$a_0 \left|\varphi_{\mathbf{0}}\right\rangle = \alpha_0 \left|\varphi_{\mathbf{0}}\right\rangle \tag{8}$$

The average particle number in the state $\mathbf{k} = \mathbf{0}$ is then:

$$\left\langle\varphi_0\right| a_0^\dagger a_0 \left|\varphi_0\right\rangle = \alpha_0^* \alpha_0 \left\langle\varphi_0 \right|\left.\varphi_0\right\rangle = N_0 \tag{9}$$

The width of the corresponding distribution is $\sqrt{N_0}$ (Complement G_V), hence negligible compared to N_0 (this number is supposed to be large).

The variational variables contained in the trial ket (3) are thus the set of $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$, as well as N_0 and ζ_0.

1-b. Total energy

Expression (61) of Complement B_{XVII} yields the total energy in the form:

$$\begin{aligned}\left\langle \widehat{H} \right\rangle_B &= \sum_{\mathbf{k}\neq\mathbf{0}} e_k \, \sinh^2\theta_{\mathbf{k}} + \frac{V_0}{2}\left(N_0 + N_e\right)^2 \\ &+ N_0 \sum_{\mathbf{k}\neq\mathbf{0}} V_{\mathbf{k}} \left[\sinh^2\theta_{\mathbf{k}} - \sinh\theta_{\mathbf{k}} \, \cosh\theta_{\mathbf{k}} \, \cos 2\left(\zeta_0 - \zeta_{\mathbf{k}}\right)\right] \\ &+ \frac{1}{2} \sum_{\mathbf{k},\mathbf{k}'\neq\mathbf{0}} V_{\mathbf{k}-\mathbf{k}'} \left[\sinh^2\theta_{\mathbf{k}} \, \sinh^2\theta_{\mathbf{k}'} + \sinh\theta_{\mathbf{k}} \, \sinh\theta_{\mathbf{k}'} \, \cosh\theta_{\mathbf{k}} \, \cosh\theta_{\mathbf{k}'} \cos 2\left(\zeta_{\mathbf{k}} - \zeta_{\mathbf{k}'}\right)\right]\end{aligned} \tag{10}$$

where the matrix elements $V_{\mathbf{k}}$ of the particle interaction potential are defined, as in Chapter XVII, by:

$$V_{\mathbf{k}} = \frac{1}{L^3} \int \mathrm{d}^3 r \; e^{-i\mathbf{k}\cdot\mathbf{r}} \; W_2(\mathbf{r}) \tag{11}$$

The term on the second line of (10) corresponds to the momentum exchanges between the $\mathbf{k} \neq \mathbf{0}$ particles and the $\mathbf{k} = \mathbf{0}$ condensate, as well as the pair annihilation-creation processes originating from the condensate. The terms in the last line, with a double summation over $\mathbf{k}$ and $\mathbf{k}'$, correspond to interaction effects between $\mathbf{k} \neq \mathbf{0}$ particles.

1-c. $N_0 \gg N_e$ approximation

As already pointed out in the introduction, for an ideal gas in its ground state, only one individual state is occupied, corresponding to the lowest energy; in that case, the average total particle number is equal to N_0, and all the populations $n_{\mathbf{k}}$ of the other $\mathbf{k}$ states are zero. We are going to assume that the system we study is a dilute gas where the interaction effects are limited so that N_0 remains very large compared to the sum of all the populations $n_{\mathbf{k}}$:

$$N_0 \gg N_e = \sum_{\mathbf{k}\neq\mathbf{0}} n_k \tag{12}$$

This hypothesis is more constraining than the one initially proposed in (1), since now the population N_0 must largely exceed the sum of all the other populations. Nevertheless it allows a simplification of the following computations while highlighting a certain number of general physical ideas.

Under these conditions, the interactions between particles in the $\mathbf{k} \neq \mathbf{0}$ states and particles in the $\mathbf{k} = \mathbf{0}$ condensate are dominant compared to the interactions between particles both in the $\mathbf{k} \neq \mathbf{0}$ states. The interaction term on the second line of (10), proportional to N_0, is therefore much larger than the term on the last line, which does not contain N_0. This is why we use the approximate average value:

$$\left\langle \widehat{H} \right\rangle_B \simeq \sum_{\mathbf{k}\neq\mathbf{0}} e_k \, \sinh^2\theta_{\mathbf{k}} + \frac{V_0}{2} \left(N_0 + N_e\right)^2 + N_0 \sum_{\mathbf{k}\neq\mathbf{0}} V_{\mathbf{k}} \left[\sinh^2\theta_{\mathbf{k}} - \sinh\theta_{\mathbf{k}} \, \cosh\theta_{\mathbf{k}} \, \cos 2\left(\zeta_0 - \zeta_{\mathbf{k}}\right)\right] \tag{13}$$

We have yet to determine the optimal values of the variables appearing in (13) by minimizing this energy average value with respect to each of them.

2. Optimization

The variational state $|\Phi_B\rangle$ depends on the variables N_0 and ζ_0 associated with the individual state $\mathbf{k} = \mathbf{0}$ (condensate), as well as on the angles $\theta_{\mathbf{k}}$ and the phases $\zeta_{\mathbf{k}}$ associated with all the other $\mathbf{k} \neq \mathbf{0}$ states. On the other hand, N_e is not a variational variable, but a function of the previous variables determined by relation (53) of Complement $\mathrm{B_{XVII}}$:

$$N_e = \sum_{\mathbf{k}\neq\mathbf{0}} \left\langle a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \right\rangle = \sum_{\mathbf{k}\neq\mathbf{0}} \sinh^2\theta_{\mathbf{k}} \tag{14}$$

As in Complement $\mathrm{B_{XVII}}$, we introduce a Lagrange multiplier μ (chemical potential, see Appendix VI) to fix the average total particle number; we thus impose the stationarity of the difference of two average values:

$$A = \left\langle \widehat{H} \right\rangle_B - \mu \left\langle \widehat{N} \right\rangle_B \tag{15}$$

where $\left\langle \widehat{N} \right\rangle_B$ is the average total particle number in the variational state:

$$\left\langle \widehat{N} \right\rangle_B = \left\langle a_0^\dagger a_0 \right\rangle + \sum_{\mathbf{k}\neq\mathbf{0}} \left\langle a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \right\rangle = N_0 + N_e \tag{16}$$

The function to be minimized is therefore:

$$A = \frac{V_0}{2} \left(N_0 + N_e\right)^2 - \mu N_0 + \sum_{\mathbf{k}\neq\mathbf{0}} \left(e_k - \mu\right) \, \sinh^2\theta_{\mathbf{k}} + N_0 \sum_{\mathbf{k}\neq\mathbf{0}} F_{\mathbf{k}} \tag{17}$$

with:

$$F_{\mathbf{k}} = V_{\mathbf{k}} \left[\sinh^2\theta_{\mathbf{k}} - \sinh\theta_{\mathbf{k}} \, \cosh\theta_{\mathbf{k}} \, \cos 2\left(\zeta_0 - \zeta_{\mathbf{k}}\right)\right] \tag{18}$$

2-a. Stationarity conditions

The function A must be made stationary with respect to all the variables. We shall start with the phases, then the parameters $\theta_{\mathbf{k}}$, and finally N_0.

α. Stationarity with respect to the phases: phase locking

The phases only intervene in the $F_{\mathbf{k}}$, as phase differences $\zeta_0 - \zeta_{\mathbf{k}}$. Since we have a repulsive potential, we assume $V_{\mathbf{k}}$ is positive; furthermore, as the variable $\theta_{\mathbf{k}}$ is always positive according to its definition in Chapter XVII, the product $\sinh\theta_{\mathbf{k}}\,\cosh\theta_{\mathbf{k}}$ is also always positive. Expression (18) shows that, whatever the value of $\theta_{\mathbf{k}}$, the minimization of the function A with respect to the phases ζ_0 and $\zeta_{\mathbf{k}}$ requires the cosine to be equal to 1, that is:

$$\zeta_{\mathbf{k}} = \zeta_0 \qquad \text{for any } \mathbf{k} \tag{19}$$

Consequently, the phases used to build the paired states must all be equal to the phase defining the coherent state associated with $\mathbf{k} = \mathbf{0}$. We call this equality the "*phase locking condition*".

β. Stationarity with respect to the $\theta_{\mathbf{k}}$

The stationarity of A with respect to each parameter $\theta_{\mathbf{k}}$ implies that, for any $\mathbf{k}$:

$$0 = V_0\,(N_0 + N_e)\,\frac{\partial N_e}{\partial \theta_{\mathbf{k}}} + 2\,(e_k - \mu)\ \sinh\theta_{\mathbf{k}}\,\cosh\theta_{\mathbf{k}} + N_0\frac{\partial F_{\mathbf{k}}}{\partial \theta_{\mathbf{k}}} \tag{20}$$

where the derivative of $F_{\mathbf{k}}$ is taken at the phase values that satisfy relation (19). Grouping on a first line the terms in $\sinh\theta_{\mathbf{k}}\cosh\theta_{\mathbf{k}}$ (including those coming from the derivative of $\sinh^2\theta_{\mathbf{k}}$), and on a second, those coming from the derivative of $\sinh\theta_{\mathbf{k}}\cosh\theta_{\mathbf{k}}$, we get:

$$\begin{aligned} 0 \;=\; & \left[V_0\,(N_0 + N_e) + e_k - \mu + N_0 V_{\mathbf{k}}\right] 2\,\sinh\theta_{\mathbf{k}}\,\cosh\theta_{\mathbf{k}} \\ & -N_0 V_{\mathbf{k}}\left(\cosh^2\theta_{\mathbf{k}} + \sinh^2\theta_{\mathbf{k}}\right) \end{aligned} \tag{21}$$

Relation (21) then becomes:

$$0 = \left[e_k - \mu + V_0\,(N_0 + N_e) + N_0 V_{\mathbf{k}}\right]\ \sinh 2\theta_{\mathbf{k}}\ - N_0 V_{\mathbf{k}}\,\cosh 2\theta_{\mathbf{k}} \tag{22}$$

that is:

$$\tanh 2\theta_{\mathbf{k}} = \frac{N_0 V_{\mathbf{k}}}{e_k - \mu + V_0\,(N_0 + N_e) + N_0 V_{\mathbf{k}}} \tag{23}$$

γ. Stationarity with respect to N_0

We now write the stationarity of A with respect to N_0. Taking into account relations (17) and (18), as well as the phase locking condition (19), we get:

$$0 = V_0\,(N_0 + N_e) - \mu + \sum_{\mathbf{k}\neq\mathbf{0}} V_{\mathbf{k}}\left[\sinh^2\theta_{\mathbf{k}} - \sinh\theta_{\mathbf{k}}\,\cosh\theta_{\mathbf{k}}\right] \tag{24}$$

This result shows that the chemical potential is equal to:

$$\mu = V_0 (N_0 + N_e) + \delta\mu \tag{25}$$

which is the sum of a mean field term $V_0 (N_0 + N_e)$ created by all the particles, and another term $\delta\mu$:

$$\delta\mu = \sum_{\mathbf{k}\neq\mathbf{0}} V_{\mathbf{k}} \left[\sinh^2\theta_{\mathbf{k}} - \sinh\theta_{\mathbf{k}} \cosh\theta_{\mathbf{k}}\right] \tag{26}$$

This last term is also the sum of two terms of different signs: a positive contribution coming from momentum transfer processes, leading to an increase of the repulsion energy due to the boson bunching effect; a negative contribution due to the creation or annihilation of pairs from the condensate $\mathbf{k} = \mathbf{0}$ (Figure 7 of Complement B_{XVII}), and expressing the reduction of that energy, due to the dynamic correlations induced by the interactions.

Relation (23) then becomes:

$$\tanh 2\theta_{\mathbf{k}} = \frac{N_0 V_{\mathbf{k}}}{\widetilde{e}_k + N_0 V_{\mathbf{k}}} \tag{27}$$

with:

$$\widetilde{e}_k = e_k - \delta\mu \tag{28}$$

2-b. Solution of the equations

The ground state we are looking for depends on two parameters that are externally fixed, the volume L^3 of the physical system, and the chemical potential μ that controls the total number of particles. We must determine the variables $\theta_{\mathbf{k}}$ from (27), as well as N_0 from (24). This last relation includes N_e, which is not an independent variable since it is determined by (14). We have a set of non-linear equations whose solution is not obvious, a priori: relation (27) determines the $\theta_{\mathbf{k}} (N_0)$, and hence $N_e (N_0)$, as a function of N_0. But N_0 itself is determined as a function of μ and the variables $\theta_{\mathbf{k}}$ (directly, and indirectly through N_e) by the stationarity condition (24). Inserting the $\theta_{\mathbf{k}} (N_0)$ in this relation, we get an implicit equation for N_0, reminiscent of the implicit equation for the gap Δ in the BCS theory (Complement C_{VII}, § 1-c-β).

A first approach for solving this implicit equation is to proceed by successive iterations, as in the Hartree-Fock method. We start from an approximate, reasonable value of N_0, such as the value μ/V_0 obtained by assuming that N_e and $\delta\mu$ are both zero. Using (27), we then get a first approximation for the $\theta_{\mathbf{k}}$ and for N_e, that can be inserted in (24) to get a new value for N_0. Iterating the process, one can expect, as for the Hartree-Fock non-linear equations, a convergence after a certain number of cycles.

Another approach is to not arbitrarily fix the chemical potential, but rather deduce it from the computation. We then start from an arbitrary N_0 value, yielding the values of the angles $\theta_{\mathbf{k}}$, then the value of N_e using (14); this fixes the total particle number $N = N_0 + N_e$, and the relations (25)-(26) yield the chemical potential. We shall use this simpler approach in what follows.

3. Properties of the ground state

We start by computing the total particle number. To highlight the general ideas while dealing with equations as simple as possible, we shall use a model where the potential matrix elements $V_{\mathbf{k}}$ are all equal to the same constant V_0, or else equal to zero:

$$\begin{aligned} V_{\mathbf{k}} &= V_0 \quad \text{if } |\mathbf{k}| \leq k_c \\ V_{\mathbf{k}} &= 0 \quad \text{if } |\mathbf{k}| > k_c \end{aligned} \tag{29}$$

where the "cutoff value" k_c characterizes the potential range b ($k_c \sim 1/b$). To further simplify, we shall consider, in each calculation, the case where $\delta\mu = 0$, i.e. where $\widetilde{e}_k$ is the same as e_k.

3-a. Particle number, quantum depletion

We use relation (14) to compute N_e, the average number of particles in the individual states $\mathbf{k} \neq \mathbf{0}$. To get $\sinh^2\theta_{\mathbf{k}}$, let us first compute $\cosh 2\theta_{\mathbf{k}}$ using:

$$\begin{aligned} \cosh 2\theta_{\mathbf{k}} &= \sqrt{\frac{\cosh^2 2\theta_{\mathbf{k}}}{\cosh^2 2\theta_{\mathbf{k}} - \sinh^2 2\theta_{\mathbf{k}}}} = \sqrt{\frac{1}{1 - \tanh^2 2\theta_{\mathbf{k}}}} \\ &= \frac{\widetilde{e}_k + N_0 V_{\mathbf{k}}}{\sqrt{\widetilde{e}_k \left(\widetilde{e}_k + 2N_0 V_{\mathbf{k}}\right)}} \end{aligned} \tag{30}$$

Since we have:

$$2\sinh^2\theta_{\mathbf{k}} = \cosh^2\theta_{\mathbf{k}} + \sinh^2\theta_{\mathbf{k}} - 1 = \cosh 2\theta_{\mathbf{k}} - 1 \tag{31}$$

we can write:

$$\sinh^2\theta_{\mathbf{k}} = \frac{1}{2}\left[\frac{\widetilde{e}_k + N_0 V_{\mathbf{k}}}{\sqrt{\widetilde{e}_k \left(\widetilde{e}_k + 2N_0 V_{\mathbf{k}}\right)}} - 1\right] \tag{32}$$

Inserting this relation in (14) we get:

$$N_e = \frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{0}}\left[\frac{\widetilde{e}_k + N_0 V_{\mathbf{k}}}{\sqrt{\widetilde{e}_k \left(\widetilde{e}_k + 2N_0 V_{\mathbf{k}}\right)}} - 1\right] \tag{33}$$

Let us see what becomes of this expression in the simple model where the $\widetilde{e}_k$ are equal to the e_k ($\delta\mu$ is supposed to be negligible). Replacing the summation in (33) by an integral, we get:

$$N_e = \frac{L^3}{16\pi^3}\int \mathrm{d}^3k \ \left[\frac{\frac{\hbar^2k^2}{2m} + N_0 V_{\mathbf{k}}}{\sqrt{\frac{\hbar^2k^2}{2m}\left(\frac{\hbar^2k^2}{2m} + 2N_0 V_{\mathbf{k}}\right)}} - 1\right] \tag{34}$$

Using the matrix elements of the simplified potential, relation (29), the function to be integrated only depends on the modulus k of $\mathbf{k}$, and goes to zero if $k > k_c$; this means

that, using spherical coordinates, the integral over dk only goes from 0 to k_c. We define the integral variable $\mathbf{s}$:

$$\mathbf{s} = \sqrt{\frac{\hbar^2}{2mN_0V_0}}\ \mathbf{k} = \xi\ \mathbf{k} \tag{35}$$

where ξ is the "healing length" introduced[4] in § 4-b of Complement $\mathrm{C_{XV}}$:

$$\xi = \sqrt{\frac{\hbar^2}{2mN_0V_0}} \tag{36}$$

Noting S_c the upper bound of $\mathbf{s}$ coming from the upper bound k_c of k:

$$S_c = \xi\ k_c \tag{37}$$

we can write:

$$N_e = \frac{L^3}{4\pi^2}\left(\frac{2mN_0V_0}{\hbar^2}\right)^{3/2}\int_0^{S_c} s^2\ \mathrm{d}s\ \left[\frac{s^2+1}{\sqrt{s^2\,(s^2+2)}} - 1\right] \tag{38}$$

The integral in (38) is still convergent if S_c goes to infinity since, when $s \to \infty$, one can make a limited expansion in powers of the infinitely small $x = 1/s^2$ and write:

$$\frac{s^2+1}{\sqrt{s^2\,(s^2+2)}} = \frac{1+1/s^2}{\sqrt{1+2/s^2}} = 1 + \frac{1}{2s^4} + .. \tag{39}$$

The integral also converges at the origin (the function to be integrated diverges as s^{-1}, but the differential element is $s^2\mathrm{d}s$, which eliminates the divergence). This integral can be readily calculated, and for an infinite value of S_c (very short-range potential) is equal to $\sqrt{2}/3$. We then get:

$$N_e = \frac{L^3}{3\pi^2}\left(\frac{mN_0V_0}{\hbar^2}\right)^{3/2} \tag{40}$$

We find that N_e is proportional to the volume and to the product N_0V_0 to the power 3/2. When the interaction potential V_0 is zero, we confirm that all the particles are in the individual state $\mathbf{k} = \mathbf{0}$. As V_0 starts increasing, the "non-condensed fraction" $N_e/\,(N_0 + N_e)$ varies, at the beginning, proportionally to the power 3/2 of V_0.

We have found that the effect of the interaction potential is to transfer a certain number of particles from the individual state $\mathbf{k} = \mathbf{0}$ towards the $\mathbf{k} \neq \mathbf{0}$ states. This effect is often called "quantum depletion". It has nothing to do with a thermal excitation effect that would bring some particles from their ground state towards excited states, as a result of the coupling with a thermal reservoir at a non-zero temperature. The calculations we are performing in this complement concern the ground state, and we assume the temperature is rigorously zero.

[4] In Complement $\mathrm{C_{XV}}$, we defined in (61) a constant ξ as a function of the parameter g associated with an interaction potential in $g\delta(\mathbf{r})$. Such a potential corresponds to $V_0 = g/\mathcal{V}$ (where $\mathcal{V}$ is the volume), and relation (36) is indeed equivalent to $\xi = \sqrt{\hbar^2/2mgn_0}$.

Comment:

Note however that the result (40) was established with the hypothesis $N_e \ll N_0$. It is therefore only valid if:

$$V_0 \ll \frac{\hbar^2}{mL^2} (N_0)^{-1/3} \tag{41}$$

If b is the range of the potential, and U its order of magnitude, relation (11) shows that the order of magnitude of the matrix elements V_0 is Ub^3/L^3; the previous condition is then written:

$$U \ll \frac{\hbar^2}{mb^2} \left(\frac{L^3}{b^3 N_0} \right)^{1/3} \tag{42}$$

The result is thus valid if U remains small compared to the kinetic energy of a particle localized within the potential range b, multiplied by the ratio between the average particle distance in the state $\mathbf{k} = \mathbf{0}$ and b. This requires the potential range b to be sufficiently small.

3-b. Energy

We now compute the energy, taking successively all the different terms of (13) into account.

α. Kinetic energy

The first contribution comes from the kinetic energy, which according to (32) is written:

$$E_c = \sum_{\mathbf{k}\neq\mathbf{0}} e_k \sinh^2\theta_{\mathbf{k}} = \frac{1}{2} \sum_{\mathbf{k}\neq\mathbf{0}} e_k \left[\frac{\widetilde{e}_k + N_0 V_{\mathbf{k}}}{\sqrt{\widetilde{e}_k \left(\widetilde{e}_k + 2N_0 V_{\mathbf{k}}\right)}} - 1 \right] \tag{43}$$

This term reminds us of the one we encountered in the computation of N_e in (33), but the presence of the factor e_k in the summation changes its properties. In the simplified potential model where the $V_{\mathbf{k}}$ are given by (29), and if we furthermore assume that $\delta\mu = 0$, the change of variable (35) leads to:

$$E_c = \frac{L^3}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} [N_0 V_0]^{5/2} \int_0^{S_c} \mathrm{d}s \; s^4 \left[\frac{s^2+1}{\sqrt{s^2\left(s^2+2\right)}} - 1 \right] \tag{44}$$

According to (39), when $s \to \infty$, the function to be integrated behaves as:

$$s^4 \left[\frac{s^2+1}{\sqrt{s^2\left(s^2+2\right)}} - 1 \right] = s^4 \left[1 + \frac{1}{2s^4} + .. - 1 \right] = \frac{1}{2} + 0\left(\frac{1}{s^2}\right) \tag{45}$$

and tends towards a constant. Consequently its integral over $\mathrm{d}s$ is divergent if S_c is infinite; if S_c is large but finite, the integral value depends linearly on the choice of S_c.

β. Interaction with the condensate

In relation (13), the mean field term in the energy $V_0 \left(N_0 + N_e\right)^2 /2$ is known since N_e has been obtained previously – see relation (38). We do not need to compute specifically its contribution to the average energy.

The second term in the potential energy is proportional to N_0, and corresponds to the interactions between atoms outside the condensate ($\mathbf{k} \neq \mathbf{0}$ individual states) and inside the condensate (population of the $\mathbf{k} = \mathbf{0}$ state); this term contains the sum of the $F_{\mathbf{k}}$ defined in (18). In order to evaluate that term, we need to compute the product $\sinh\theta_{\mathbf{k}}\cosh\theta_{\mathbf{k}}$:

$$\begin{aligned} \sinh\theta_{\mathbf{k}}\, \cosh\theta_{\mathbf{k}} &= \frac{1}{2}\sinh 2\theta_{\mathbf{k}} = \frac{1}{2}\sqrt{\frac{\sinh^2 2\theta_{\mathbf{k}}}{\cosh^2 2\theta_{\mathbf{k}} - \sinh^2 2\theta_{\mathbf{k}}}} \\ &= \frac{1}{2}\sqrt{\frac{\tanh^2 2\theta_{\mathbf{k}}}{1 - \tanh^2 2\theta_{\mathbf{k}}}} \end{aligned} \tag{46}$$

or else, according to (27):

$$\sinh\theta_{\mathbf{k}}\, \cosh\theta_{\mathbf{k}} = \frac{N_0 V_{\mathbf{k}}}{2\sqrt{\widetilde{e}_k \left(\widetilde{e}_k + 2N_0 V_{\mathbf{k}}\right)}} \tag{47}$$

Taking into account the phase locking condition (19) for the optimal variational state, the $F_{\mathbf{k}}$ defined in (18) are equal to:

$$\begin{aligned} F_{\mathbf{k}} &= V_{\mathbf{k}} \left[\sinh^2\theta_{\mathbf{k}} - \sinh\theta_{\mathbf{k}}\, \cosh\theta_{\mathbf{k}}\right] \\ &= \frac{V_{\mathbf{k}}}{2}\left[\frac{\widetilde{e}_k}{\sqrt{\widetilde{e}_k \left(\widetilde{e}_k + 2N_0 V_{\mathbf{k}}\right)}} - 1\right] \end{aligned} \tag{48}$$

We obtain the contribution E_{N_0} to the energy:

$$E_{N_0} = N_0 \sum_{\mathbf{k}\neq\mathbf{0}} F_{\mathbf{k}} = \frac{N_0}{2} \sum_{\mathbf{k}\neq\mathbf{0}} V_{\mathbf{k}} \left[\frac{\widetilde{e}_k}{\sqrt{\widetilde{e}_k \left(\widetilde{e}_k + 2N_0 V_{\mathbf{k}}\right)}} - 1\right] \tag{49}$$

For the simple model (29) already used before, and if $\delta\mu = 0$, this results becomes:

$$E_{N_0} = \frac{L^3}{4\pi^2} N_0 V_0 \left(\frac{2mN_0 V_0}{\hbar^2}\right)^{3/2} \int_0^{S_c} s^2 \, \mathrm{d}s \, \left[\sqrt{\frac{s^2}{s^2+2}} - 1\right] \tag{50}$$

The function to be integrated behaves, at infinity, as:

$$s^2 \left[\frac{s}{\sqrt{s^2+2}} - 1\right] \sim s^2 \left[1 - \frac{1}{s^2} + .. - 1\right] = -1 + 0\left(\frac{1}{s^2}\right) \tag{51}$$

which means the integral is not convergent when $S_c \to \infty$, but depends linearly on S_c.

γ. Ground state total energy

With the same approximation $\delta\mu = 0$, the sum of (43) and (49) yields:

$$E_c + E_{N_0} = \frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{0}}\left[\sqrt{e_k\left(e_k + 2N_0V_{\mathbf{k}}\right)} - e_k - N_0V_{\mathbf{k}}\right] \tag{52}$$

or else, using again the simplified model (29) for the interaction potential and adding (44) and (50):

$$E_c + E_{N_0} = \frac{1}{4\pi^2}\left(\frac{2mL^2}{\hbar^2}\right)^{3/2}\left(N_0V_0\right)^{5/2}\int_0^{S_c} ds\ \left[s^3\sqrt{(s^2+2)} - s^4 - s^2\right] \tag{53}$$

Relations (45) and (51) show that the function to be integrated tends towards $-1/2$ when $s \to \infty$; we have again a divergent integral when $k_c \to \infty$ (or $S_c \to \infty$). Consequently, its value is a linear function of the chosen cutoff frequency k_c. However, the fact that the limit of the function to be integrated is negative indicates that, for large values of S_c, the decrease in potential energy overcomes the increase in kinetic energy.

The ground state total energy E_{ground} is the sum of all the energies we just computed, including the mean field term:

$$E_{\text{ground}} = \frac{V_0}{2}\left(N_0 + N_e\right)^2 + E_c + E_{N_0} \tag{54}$$

where $E_c + E_{N_0}$ is given by (53).

Remarque:

The divergences appearing when $S_c \to \infty$ in our calculation of the energy are not fundamental. They occur when one assumes that the matrix elements $V_{\mathbf{k}}$ of the potential written in (29) remain constant when $k \to \infty$, while they tend to zero with a realistic potential. It is indeed possible to perform a more careful treatment of the potential, such as that mentioned in § 4.2 of Référence [15], and to obtain a finite result.

3-c. Phase locking; comparison with the BCS mechanism

We just saw in § 2-a-α how all the phases $\zeta_{\mathbf{k}}$ had to become equal to the phase ζ_0 associated with the state $\mathbf{k} = \mathbf{0}$ in order to minimize the repulsive energy between any particle in the $\mathbf{k} \neq \mathbf{0}$ states and any particle in the $\mathbf{k} = \mathbf{0}$ state. Fixing the phase differences to be zero is what we called the "phase locking condition". This situation reminds us of the "symmetry breaking" of the BCS mechanism, where a common locking of all the relative phases of all the pairs $(\mathbf{k}, -\mathbf{k})$ enabled the building of a gap Δ, through a collective effect. For bosons, the equivalent of the collective mean field created by the pairs is the one created by the condensate of particles in the $\mathbf{k} = \mathbf{0}$ state. This is why it is now the relative phase of the pair states with respect to that of the condensate that plays a role; in other words, we have now an "external" instead of an "internal" mechanism. The gain in energy will be exactly the same, whatever value is chosen for the phase ζ_0; only the relative phase $\zeta_{\mathbf{k}} - \zeta_0$ is relevant. Accordingly, and just as for fermions, the arbitrary choice of the phase leads to a symmetry breaking phenomenon. The analogy is reinforced by the fact that, for a fixed value of μ, we found in § 2-b that the value of the

particle number N_0 in the $\mathbf{k} = \mathbf{0}$ state is given by an implicit equation in N_0. Similarly, in the BCS theory, it is also an implicit equation that fixes the value of the gap Δ.

Relations (40) and (53) include non-integer powers of the interaction potential V_0, and are thus non-analytic functions of that potential. They cannot be obtained by a perturbation theory as a power series expansion of V_0, and this is another analogy with the results of Complement B_{XVII}.

We also saw in that complement that, in the BCS mechanism, it is the energies in the vicinity of the Fermi level that play the most important role. This is not the case for a system of repulsive bosons. For example, in (44), the function whose integration over s yields the kinetic energy is:

$$f_c(s) = s^4 \left[\frac{s^2+1}{\sqrt{s^2(s^2+2)}} - 1 \right] \tag{55}$$

whereas the one yielding the interaction potential energy between particles in the $\mathbf{k} \neq \mathbf{0}$ states and particles in the $\mathbf{k} = \mathbf{0}$ state is:

$$f_{N_0}(s) = s^2 \left[\frac{s^2}{\sqrt{s^2(s^2+2)}} - 1 \right] \tag{56}$$

Figure 1 plots these two functions with dashed lines, as well as their sum with a solid line. It illustrates how, in the case of repulsive bosons, the effects of minimization of the potential energy overcome those of the kinetic energy. This minimization of the repulsion necessarily comes with a modification of the particles' position correlation function, which must decrease at short distances; this interpretation will be substantiated in § 3-d-β where we study the binary correlation functions. We also note in the figure that the accumulated gain of energy is not due to a particular energy band: all the s values contribute up to the limit imposed by the upper bound of the integral.

We can further refine the analysis of the energy balance by looking at the gain of energy per individual quantum state. We must then remove from the previous relations the factor k^2 coming from the density of states, and hence remove a factor s^2 from (55) and (56). Figure 2 plots the resulting functions, which show that as long as s is small or of the order of 1, the decrease in potential energy largely overcomes the increase in kinetic energy; on the other hand, the two contributions balance each other when $s \gg 1$. According to (35), the condition $s \lesssim 1$ corresponds to:

$$k \lesssim \xi^{-1} \qquad \text{that is:} \qquad e_k \lesssim N_0 V_0 \tag{57}$$

This means that individual states of low energy provide most of the decrease of the repulsive potential energy; the corresponding energy domain is proportional both to N_0 and to the interaction matrix element V_0. Relations (27) also show that it is those energies in the system ground state that the energy minimization affects the most. From the physical point of view, it is understandable that particles having a low kinetic energy compared to the interaction energy $N_0 V_0$ are the most affected by the interactions, whereas those with a kinetic energy large compared to $N_0 V_0$ have their correlations only slightly modified by the interaction potential. However, as we noted before, even though the individual contribution of the highest energy state to the energy reduction is reduced, their large number (corresponding to a density of states proportional to k^2) means that their contribution to the total energy remains significant.

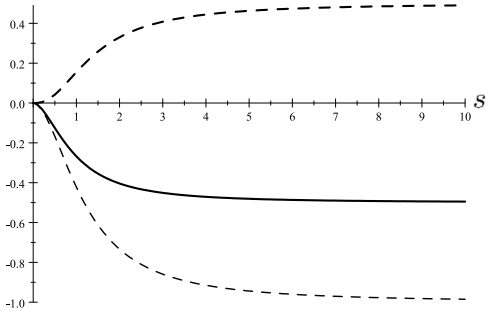


Figure 1: Plots as a function of s of the functions whose integral over s yield the kinetic energy E_c (upper dashed curve), the potential energy for the interaction with the condensate E_{N_0} (lower dashed curve), as well as their sum (solid line). The increase in the kinetic energy is overcome by the decrease in the potential energy, which ends up lowering the total energy.

3-d. Correlation functions

As the system is contained in a box and obeys periodic boundary conditions, we expect the properties of the one-body correlation functions to be translation invariant. This does not rule out a possible spatial dependence of the correlation functions, as far as the differences in positions are concerned. This is what we want to elucidate now.

α. One particle

Expanding the field operator $\Psi(\mathbf{r})$ on the annihilation operators, according to relation (A-3) of Chapter XVI, we get for the one-particle correlation function G_1

$$G_1(\mathbf{r},\mathbf{r}') = \langle\Phi_B|\,\Psi^\dagger(\mathbf{r})\Psi(\mathbf{r}')\,|\Phi_B\rangle = \frac{1}{L^3}\sum_{\mathbf{k},\mathbf{k}'} e^{i\left(\mathbf{k}'\cdot\mathbf{r}'-\mathbf{k}\cdot\mathbf{r}\right)}\,\langle\Phi_B|\,a^\dagger_{\mathbf{k}}a_{\mathbf{k}'}\,|\Phi_B\rangle \tag{58}$$

where expression (3) determines $|\Phi_B\rangle$. Since in this state the particle number in an individual state $\mathbf{k}$ is always the same as that number for the individual state $-\mathbf{k}$, the average value of $a^\dagger_{\mathbf{k}}a_{\mathbf{k}'}$ in $|\Phi_B\rangle$ will be different from zero only if $\mathbf{k}=\mathbf{k}'$. In the summation over $\mathbf{k}$, the $\mathbf{k}=\mathbf{0}$ contribution introduces a term in N_0; adding to it all the other contributions, we get:

$$G_1(\mathbf{r},\mathbf{r}') = \frac{1}{L^3}\left[N_0 + \sum_{\mathbf{k}\neq\mathbf{0}}\langle\widehat{n}_{\mathbf{k}}\rangle\, e^{i\mathbf{k}\cdot\left(\mathbf{r}'-\mathbf{r}\right)}\right] \tag{59}$$

When $\mathbf{r}=\mathbf{r}'$, the function G_1 is simply equal to the total particle density n_{tot}:

$$n_{\text{tot}} = \frac{N_0+N_e}{L^3} \tag{60}$$

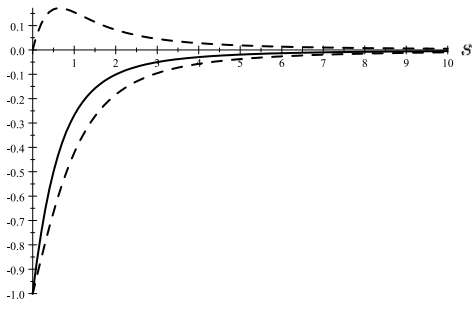


Figure 2: Plots as a function of s of the kinetic energy (upper dashed curve), the potential energy (lower dashed curve), and the total energy (solid line) per individual state. It shows that it is the lowest kinetic energy states that make the largest contribution to the lowering of the energy.

When $\mathbf{r}$ and $\mathbf{r}'$ are different, the function G_1 is the sum of two terms:

– one term corresponding to particles in the $\mathbf{k} = \mathbf{0}$ state (condensed particles), independent of the positions; this term does not decrease at large distance, but has an infinite range.

– a second term corresponding to particles in the $\mathbf{k} \neq \mathbf{0}$ state, which is the transform of the particle distribution $\langle \widehat{n}_{\mathbf{k}} \rangle$, and therefore goes to zero when $\mathbf{r}$ and $\mathbf{r}'$ move away from each other (it has a microscopic range).

We find again the Penrose-Onsager criterion according to which it is the condensed fraction of a boson system that leads to an infinite range of the non-diagonal one-body correlation function (in the case of paired fermions, we found in Complement $\mathrm{C_{XVII}}$, §§ 2-a-β and 2-b-β, that this long range does not occur for the one-body correlation function, but only for the two-body correlation function).

β. Two particles

The diagonal two-particle correlation function is written:

$$G_2(\mathbf{r},\mathbf{r};\mathbf{r}',\mathbf{r}') = \langle \Phi_B | \Psi^\dagger(\mathbf{r})\Psi^\dagger(\mathbf{r}')\Psi(\mathbf{r}')\Psi(\mathbf{r}) | \Phi_B \rangle$$
$$= \frac{1}{L^6} \sum_{\mathbf{k},\mathbf{k}',\mathbf{k}'',\mathbf{k}'''} e^{i(\mathbf{k}'''-\mathbf{k})\cdot\mathbf{r}} e^{i(\mathbf{k}''-\mathbf{k}')\cdot\mathbf{r}'} \langle \Phi_B | a_{\mathbf{k}}^\dagger a_{\mathbf{k}'}^\dagger a_{\mathbf{k}''} a_{\mathbf{k}'''} | \Phi_B \rangle \qquad (61)$$

We get the same simplifications as in § 4-b-β of Complement $\mathrm{B_{XVII}}$ in the computation of the average values of products of creation and annihilation operators: operators a_0 placed on the right each yield a factor α_0 and operators $a_0^\dagger$ placed on the left, each a factor α_0^*; the average values of the other operator products are given by the results of § C in Chapter XVII. We must distinguish between several cases, depending on the number of values, among the 4 summation indices, which are equal to $\mathbf{k} = \mathbf{0}$; we shall proceed by decreasing values of that number.

(i) If the four operators concern the $\mathbf{k} = \mathbf{0}$ state (case represented in Figure 6 of Complement B_{XVII}), we get the contribution:

$$\frac{N_0 (N_0 - 1)}{L^6} \simeq \left(\frac{N_0}{L^3} \right)^2 \tag{62}$$

which is position independent.

The case where only three of the summation indices are zero is not possible, as the corresponding term would contain the average value of an operator $a_{\mathbf{k}}$ (or of its Hermitian conjugate) in the state $|\Phi_B\rangle$, which is zero.

(ii) If one creation and one annihilation operator concern the individual $\mathbf{k} = \mathbf{0}$ state, two cases may occur and yield different types of terms:

– direct terms in $a_{\mathbf{0}}^{\dagger} a_{\mathbf{k}'}^{\dagger} a_{\mathbf{k}'} a_{\mathbf{0}}$ or $a_{\mathbf{k}'}^{\dagger} a_{\mathbf{0}}^{\dagger} a_{\mathbf{0}} a_{\mathbf{k}'}$; both contributions are equal and their sum leads to:

$$\frac{2 N_0 N_e}{L^6} \tag{63}$$

which is also position independent.

– exchange terms in $a_{\mathbf{0}}^{\dagger} a_{\mathbf{k}'}^{\dagger} a_{\mathbf{0}} a_{\mathbf{k}'}$ or $a_{\mathbf{k}'}^{\dagger} a_{\mathbf{0}}^{\dagger} a_{\mathbf{k}'} a_{\mathbf{0}}$; the corresponding terms are also equal and their sum is written:

$$\begin{aligned} &\frac{N_0}{L^6} \left\{ \sum_{\mathbf{k}' \neq \mathbf{0}} e^{i\mathbf{k}' \cdot (\mathbf{r} - \mathbf{r}')} \langle \widehat{n}_{\mathbf{k}'} \rangle + \sum_{\mathbf{k} \neq \mathbf{0}} e^{i\mathbf{k} \cdot (\mathbf{r}' - \mathbf{r})} \langle \widehat{n}_{\mathbf{k}} \rangle \right\} \\ &\quad = \frac{2 N_0}{L^6} \sum_{\mathbf{k} \neq \mathbf{0}} \langle \widehat{n}_{\mathbf{k}} \rangle \cos \left[\mathbf{k} \cdot (\mathbf{r}' - \mathbf{r}) \right] \end{aligned} \tag{64}$$

which now depends on the difference in the positions $\mathbf{r}$ and $\mathbf{r}'$. These terms reflect the existence of a bunching effect between bosons; relation (C-28) of Chapter XVII yields the value of $\langle \widehat{n}_{\mathbf{k}} \rangle$:

$$\langle \widehat{n}_{\mathbf{k}} \rangle = \sinh^2 \theta_{\mathbf{k}} \tag{65}$$

(iii) If the operators corresponding to $\mathbf{k} = \mathbf{0}$ are of the same nature (both creation or both annihilation operators), we get terms such as the ones represented on Fig 7 of Complement B_{XVII}, corresponding to the creation or annihilation of a pair from the condensate $\mathbf{k} = \mathbf{0}$:

– for a product of the type $a_{\mathbf{k}}^{\dagger} a_{-\mathbf{k}}^{\dagger} a_{\mathbf{0}} a_{\mathbf{0}}$ where $\mathbf{k}$ and $\mathbf{k}'$ are not zero but opposite (for the same reason as explained before), we get an anomalous average value in the state $|\overline{\varphi}_{\mathbf{k}}\rangle$, multiplied by the average value of the product of two operators a_0. The first average value is given by relation (C-51) of Chapter XVII, and the second yields $(\alpha_0)^2$, that is $N_0 e^{2i\zeta_0}$ according to (7).

– for a product of the type $a_{\mathbf{0}}^{\dagger} a_{\mathbf{0}}^{\dagger} a_{\mathbf{k}''} a_{-\mathbf{k}''}$ where $\mathbf{k}''$ and $\mathbf{k}'''$ are not zero but opposite, we get the complex conjugate of the previous result.

The sum of the two previous results is then written:

$$-\frac{N_0}{L^6}\left\{\sum_{\mathbf{k}\neq\mathbf{0}} e^{i\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r})}\sinh\theta_{\mathbf{k}}\,\cosh\theta_{\mathbf{k}}e^{2i(\zeta_0-\zeta_{\mathbf{k}})}\right.$$
$$\left.+\sum_{\mathbf{k}''\neq\mathbf{0}} e^{i\mathbf{k}''\cdot(\mathbf{r}'-\mathbf{r})}\sinh\theta_{\mathbf{k}''}\,\cosh\theta_{\mathbf{k}''}e^{-2i(\zeta_0-\zeta_{\mathbf{k}''})}\right\}$$
$$=-\frac{2N_0}{L^6}\sum_{\mathbf{k}\neq\mathbf{0}}\sinh\theta_{\mathbf{k}}\,\cosh\theta_{\mathbf{k}}\,\cos\left[\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r})+2\left(\zeta_0-\zeta_{\mathbf{k}}\right)\right] \tag{66}$$

(iv) Finally we have terms where none of the wave vectors is zero, corresponding to cases where the particles are in $\mathbf{k}\neq\mathbf{0}$ states before and after the interaction. They include a direct term:

$$\frac{(N_e)^2}{L^6} \tag{67}$$

which is constant, an exchange term, and finally a pair annihilation-creation term. Compared to the previous terms (which are proportional N_0), their relative value is of the order of N_e/N_0. Taking into account the exchange term and the pair creation-annihilation term leads to simple calculations of a type already performed; however, to be consistent with (12) and the corresponding energy approximations, we shall ignore those terms.

The sum of (62), (63) and (67) yield the constant N^2/L^6, to which we add (64) and (66) and get:

$$G_2\left(\mathbf{r},\mathbf{r};\mathbf{r}',\mathbf{r}'\right)\simeq\frac{N^2}{L^6}+\frac{2N_0}{L^6}\sum_{\mathbf{k}\neq\mathbf{0}}\left\{\sinh^2\theta_{\mathbf{k}}\cos\left[\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r})\right]\right.$$
$$\left.-\sinh\theta_{\mathbf{k}}\,\cosh\theta_{\mathbf{k}}\,\cos\left[\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r})+2\left(\zeta_0-\zeta_{\mathbf{k}}\right)\right]\right\} \tag{68}$$

The position dependent term on the second line shows how the relative phases introduced in $|\Phi_B\rangle$ control the relative particle position in the physical system; it confirms that the choice $\zeta_{\mathbf{k}}=\zeta_0$ does indeed decrease the probability of finding two particles close to each other.

When the phase locking condition (19) is satisfied, the position dependent contribution becomes:

$$\frac{2N_0}{L^6}\sum_{\mathbf{k}\neq\mathbf{0}}\left[\sinh^2\theta_{\mathbf{k}}-\sinh\theta_{\mathbf{k}}\cosh\theta_{\mathbf{k}}\right]\cos\left[\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r})\right] \tag{69}$$

Since $\sinh\theta_{\mathbf{k}}<\cosh\theta_{\mathbf{k}}$, the cosine in each term has a negative coefficient; it does decrease the probability $G_2\left(\mathbf{r},\mathbf{r};\mathbf{r},\mathbf{r}\right)$ of finding two particles at the same point $\mathbf{r}$: the dynamic correlations appearing in the system tend to "antibunch" the particles, and hence reduce their repulsive interactions. The final result is a compromise between a $\sinh^2\theta_{\mathbf{k}}$ term that leads to bunching (as for a non-interacting boson gas) and a antibunching term in $\sinh\theta_{\mathbf{k}}\cosh\theta_{\mathbf{k}}$ that is larger, and involves anomalous average values (creation or annihilation of particle pairs in the condensate).

Comments:

(i) The correlation function (68) is invariant with respect to the exchange of $\mathbf{r}$ and $\mathbf{r}'$, as seen from the way the terms $\mathbf{k}$ and $-\mathbf{k}$ are accounted for in the summation. Its Fourier transform only contains terms in $\cos[\mathbf{k}\cdot(\mathbf{r}'-\mathbf{r})]$, which can take on any value by an appropriate choice of the $\theta_{\mathbf{k}}$ and the $\zeta_{\mathbf{k}}$. As mentioned in the introduction, the variational state can lead to any correlation function; the results discussed above concern the optimal value of this correlation function.

(ii) On several occasions, we assumed the chemical potential correction $\delta\mu$, defined in (26), to be zero, which enabled us to replace the e_k by the $\widetilde{e}_k$. Let us now check that a non-zero value of this correction does not radically change the results we obtained.

Using the model (29) where the non-zero matrix elements $V_{\mathbf{k}}$ are all equal to the same constant V_0, expression (26) is simply written:

$$\delta\mu = V_0 \sum_{\mathbf{k}\neq\mathbf{0}} \left[\sinh^2\theta_{\mathbf{k}} - \sinh\theta_{\mathbf{k}}\,\cosh\theta_{\mathbf{k}}\right] = -V_0 \sum_{\mathbf{k}\neq\mathbf{0}} \sinh\theta_{\mathbf{k}} e^{-\theta_{\mathbf{k}}} \leq 0 \tag{70}$$

where the summation is limited to vectors $\mathbf{k}$ having a modulus less than k_c. Setting:

$$\delta\mu = -xV_0N_0 \tag{71}$$

with $x > 0$, relation (27) that fixes the $\theta_{\mathbf{k}}$ becomes:

$$\tanh 2\theta_{\mathbf{k}} = \frac{N_0V_0}{e_k + N_0V_0\,(1+x)} \tag{72}$$

The corrective effect of x and hence $\delta\mu$, is to lower the $\theta_{\mathbf{k}}$; this correction is however negligible if $e_k \gg xN_0V_0$. Consequently, the populations of the individual states $\mathbf{k}$ (which are equal to $\sinh^2\theta_{\mathbf{k}}$) decrease when $e_k \lesssim xN_0V_0$, but remain practically unchanged in the opposite case. The quantities resulting from a summation over $\mathbf{k}$, such as N_e, are then barely affected: the change in the function to be integrated only occurs for small values of s, whose contributions, in any case, are weak because of the factor s^2 in the integral (38). As for the energy, this is accentuated since the integral in (53) diverges if S_c is infinite, which means it mainly depends on the large values of s (if $S_c \gg 1$). Turning now to the correlation functions computed in § 3-d, they contain summations over $\mathbf{k}$ that lead to the same integrals; they are thus fairly insensitive to the value of $\delta\mu$. This explains why, aside from the predictions concerning the populations of small wave vectors, the approximation $\delta\mu = 0$ used in § 3 is reasonable in many cases.

To push the analysis a step further we need to compute the value of the x coefficient. This requires improving the precision of the calculations, and in particular taking into account the interactions of the particles in the $\mathbf{k} \neq \mathbf{0}$ individual states. This is beyond the scope of this complement, and we shall simply accept that x is small and note that only the small k populations are changed when $\delta\mu$ is not equal to zero.

4. Bogolubov operator method

We now present a different point of view and introduce the Bogolubov method; it is based on the search for a readily diagonalizable operator form of the Hamiltonian (or of an approximate expression of this Hamiltonian). This method not only applies to the ground state, but it also enables the study of the excited states. We shall use the results of § E in Chapter XVII to introduce new operators that simplify the diagonalization of the Hamiltonian.

4-a. Variational space, restriction on the Hamiltonian

The variational set we consider has been defined in (3); we assume:

$$|\Phi_B\rangle = |\varphi_{\mathbf{0}}\rangle \otimes |\Psi_{\text{paired}}\rangle \tag{73}$$

where $|\varphi_{\mathbf{0}}\rangle$ is the coherent state (6), and $|\Psi_{\text{paired}}\rangle$, any paired state in the Fock space spanned by all the individual states others than $\mathbf{k} = \mathbf{0}$. We call $\mathcal{E}_B(N_0)$ the ensemble of kets expressed as (73).

We now take the general Hamiltonian operator $\widehat{H}$ written in (8) of Complement $\mathrm{B_{XVII}}$, and consider its action restricted to such states; the corresponding matrix elements are of the type:

$$\langle \Phi_B | \, \widehat{H} \, | \Phi'_B \rangle \tag{74}$$

where $|\Phi_B\rangle$ and $|\Phi'_B\rangle$ are any two kets of $\mathcal{E}_B(N_0)$. In the computation of this matrix element, the same simplifications as in § 1 will occur: any annihilation operator a_0 on the right can be replaced by the number α_0, any creation operator $a_0^\dagger$ on the left by the complex conjugate α_0^*. We will further simplify the problem by assuming, as in (12), that the total population N_e of the $\mathbf{k} \neq \mathbf{0}$ individual levels is much smaller than $N_0 = |\alpha_0|^2$, and keeping only certain terms among the Hamiltonian interaction terms.

First, we study the forward scattering terms, which are the terms ($\mathbf{k}'' = \mathbf{k}$ and $\mathbf{k}''' = \mathbf{k}'$) in relation (20) of Complement $\mathrm{B_{XIII}}$. Their expression is:

$$\frac{V_0}{2} \sum_{\mathbf{k},\mathbf{k}'} a_{\mathbf{k}}^\dagger a_{\mathbf{k}'}^\dagger a_{\mathbf{k}'} a_{\mathbf{k}} = \frac{V_0}{2} \sum_{\mathbf{k},\mathbf{k}'} \left[a_{\mathbf{k}}^\dagger a_{\mathbf{k}} a_{\mathbf{k}'}^\dagger a_{\mathbf{k}'} - \delta_{\mathbf{k},\mathbf{k}'} \; a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \right] = \frac{V_0}{2} \widehat{N} \left(\widehat{N} - 1 \right) \tag{75}$$

In this equality, $\widehat{N}$ is the total number of particles operator:

$$\widehat{N} = \widehat{N}_0 + \widehat{N}_e \tag{76}$$

where $N_0 = a_0^\dagger a_0$ is the operator associated with the population of the individual state $\mathbf{k} = \mathbf{0}$, and $\widehat{N}_e$ that associated with the total number of particles in the states $\mathbf{k} \neq \mathbf{0}$:

$$\widehat{N}_e = \sum_{\mathbf{k} \neq \mathbf{0}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \tag{77}$$

As for all the other interactions terms, we shall proceed as in § 1-c and will only keep the terms that contain N_0 – the others correspond to interactions between particles in the $\mathbf{k} \neq \mathbf{0}$ individual states, assumed to be negligible when inequality (12) is satisfied. In all the terms we keep, there are either four or two creation or annihilation operators concerning the $\mathbf{k} = \mathbf{0}$ state.

Those containing the product $a_0^\dagger a_0^\dagger a_0 a_0$, or one of the two products $a_{\mathbf{k}}^\dagger a_0^\dagger a_0 a_{\mathbf{k}}$ and $a_0^\dagger a_{\mathbf{k}}^\dagger a_{\mathbf{k}} a_0$, are already taken into account in the mean field term (75). We simply have to add:

– the terms containing the products $a_{\mathbf{k}}^\dagger a_0^\dagger a_{\mathbf{k}} a_0$ or $a_0^\dagger a_{\mathbf{k}}^\dagger a_0 a_{\mathbf{k}}$, i.e. the exchange terms of § 4-b-β in Complement $\mathrm{B_{XVII}}$; they yield a contribution:

$$N_0 V_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}} \tag{78}$$

– the terms in $a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} a_0 a_0$, corresponding to the pair creation from the condensate, or the terms in $a^\dagger_0 a^\dagger_0 a_{\mathbf{k}} a_{-\mathbf{k}}$ corresponding to the annihilation of pairs into the condensate; their contribution is:

$$\frac{N_0 V_{\mathbf{k}}}{2}\left[\mathrm{e}^{2i\zeta_0}\ a^\dagger_{-\mathbf{k}} a^\dagger_{\mathbf{k}} + \mathrm{e}^{-2i\zeta_0}\ a_{\mathbf{k}} a_{-\mathbf{k}}\right] \tag{79}$$

With the above conditions, we get a simplified version of the Hamiltonian $\widehat{H}$, which becomes a reduced Hamiltonian $\widehat{H}_R$:

$$\begin{aligned}\widehat{H}_R =& \frac{V_0 \widehat{N}(\widehat{N}-1)}{2} \\ &+ \sum_{\mathbf{k}\neq\mathbf{0}} \left\{ e_k a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + N_0 V_{\mathbf{k}} \left[a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + \frac{1}{2}\left(\mathrm{e}^{2i\zeta_0}\ a^\dagger_{-\mathbf{k}} a^\dagger_{\mathbf{k}} + \mathrm{e}^{-2i\zeta_0}\ a_{\mathbf{k}} a_{-\mathbf{k}} \right) \right] \right\}\end{aligned} \tag{80}$$

If the number of particles is fixed, the first term in the right hand side (mean field) introduces only the same displacement of all energies, without physical consequence. In the Bogolobov approximation, where the condition $N_0 \gg N_e$ is assumed, one often merely replaces $\widehat{N}$ by $\widehat{N}_0$ in this term, which amounts to restricting the sum of (75) to the terms $\mathbf{k} = \mathbf{k}' = \mathbf{0}$. Since the kets (73) are eigenvectors of a_0 with eigenvalue $\sqrt{N_0}$, we may then replace the mean field operator (75) by the number $V_0 N_0^2/2$. If, moreover, one assumes that $\zeta_0 = 0$, one obtains the simpler expression:

$$\widehat{H}_R = \frac{V_0 N_0^2}{2} + \sum_{\mathbf{k}\neq\mathbf{0}} e_k a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + N_0 V_{\mathbf{k}} \left[a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + \frac{a^\dagger_{-\mathbf{k}} a^\dagger_{\mathbf{k}} + a_{\mathbf{k}} a_{-\mathbf{k}}}{2} \right] \tag{81}$$

Either (80) or (81) can be used as the Hamiltonian within the Bogolubov approximation.

Neither of these operators conserves the total particle number, because of its terms proportional to the product of two creation or two annihilation operators. Complement B_{XVII} explained how such "anomalous" terms can, nevertheless, account for the interaction effects within the framework of certain approximations. We are now going to show that this expression can be put in the form of a Hamiltonian of independent particles, provided the operators undergo the transformation introduced in § E of Chapter XVII.

4-b. Bogolubov Hamiltonian

We obtained in Chapter XVII the expression (E-29) of the Hamiltonian operator $\widehat{H}_B$:

$$\widehat{H}_B = \sum_{\mathbf{k}\in D} \hbar\omega_k \left[b^\dagger_{\mathbf{k}} b_{\mathbf{k}} + b^\dagger_{-\mathbf{k}} b_{-\mathbf{k}} \right] \tag{82}$$

which includes the Bogolubov operators for bosons:

$$\begin{aligned} b_{\mathbf{k}} &= u_{\mathbf{k}} a_{\mathbf{k}} + v_{\mathbf{k}} a^\dagger_{-\mathbf{k}} \\ b_{-\mathbf{k}} &= u_{\mathbf{k}} a_{-\mathbf{k}} + v_{\mathbf{k}} a^\dagger_{\mathbf{k}} \end{aligned} \tag{83}$$

Remember that D is half the momentum space, avoiding double counting of the same pairs of states in (82). Relation (E-15) of Chapter XVII expresses the $u_{\mathbf{k}}$ and $v_{\mathbf{k}}$ in terms of the two parameters $\theta_{\mathbf{k}}$ and $\zeta_{\mathbf{k}}$:

$$\begin{aligned} u_{\mathbf{k}} &= \cosh\theta_{\mathbf{k}}\ e^{-i\zeta_{\mathbf{k}}} \\ v_{\mathbf{k}} &= \sinh\theta_{\mathbf{k}}\ e^{i\zeta_{\mathbf{k}}} \end{aligned} \tag{84}$$

As for the value of the parameter ω_k, it will be fixed later.

We then have:

$$b^\dagger_{\mathbf{k}} b_{\mathbf{k}} = |u_{\mathbf{k}}|^2\ \ a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + |v_{\mathbf{k}}|^2\ \ a_{-\mathbf{k}} a^\dagger_{-\mathbf{k}} + \ \ \left(u^*_{\mathbf{k}} v_{\mathbf{k}} a^\dagger_{\mathbf{k}} a^\dagger_{-\mathbf{k}} + u_{\mathbf{k}} v^*_{\mathbf{k}} a_{-\mathbf{k}} a_{\mathbf{k}}\right) \tag{85}$$

and:

$$b^\dagger_{-\mathbf{k}} b_{-\mathbf{k}} = |u_{\mathbf{k}}|^2\ \ a^\dagger_{-\mathbf{k}} a_{-\mathbf{k}} + |v_{\mathbf{k}}|^2\ \ a_{\mathbf{k}} a^\dagger_{\mathbf{k}} + \ \ \left(u^*_{\mathbf{k}} v_{\mathbf{k}} a^\dagger_{-\mathbf{k}} a^\dagger_{\mathbf{k}} + u_{\mathbf{k}} v^*_{\mathbf{k}} a_{\mathbf{k}} a_{-\mathbf{k}}\right) \tag{86}$$

The operators in these equalities can be rearranged in normal order, using the proper commutation; adding them both, we get:

$$\begin{aligned} b^\dagger_{\mathbf{k}} b_{\mathbf{k}} + b^\dagger_{-\mathbf{k}} b_{-\mathbf{k}} = &\left(|u_{\mathbf{k}}|^2 + |v_{\mathbf{k}}|^2\right)\left(a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + a^\dagger_{-\mathbf{k}} a_{-\mathbf{k}}\right) \\ &+ 2\,|v_{\mathbf{k}}|^2 + 2u^*_{\mathbf{k}} v_{\mathbf{k}} a^\dagger_{-\mathbf{k}} a^\dagger_{\mathbf{k}} + 2u_{\mathbf{k}} v^*_{\mathbf{k}} a_{\mathbf{k}} a_{-\mathbf{k}} \end{aligned} \tag{87}$$

that is, taking (84) into account:

$$\begin{aligned} b^\dagger_{\mathbf{k}} b_{\mathbf{k}} + b^\dagger_{-\mathbf{k}} b_{-\mathbf{k}} = &\cosh 2\theta_{\mathbf{k}}\ \ \left[a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + a^\dagger_{-\mathbf{k}} a_{-\mathbf{k}}\right] \\ &+ 2\sinh^2\theta_{\mathbf{k}} + \sinh 2\theta_{\mathbf{k}} \left[a^\dagger_{-\mathbf{k}} a^\dagger_{\mathbf{k}}\ e^{2i\zeta_{\mathbf{k}}} + a_{\mathbf{k}} a_{-\mathbf{k}}\ e^{-2i\zeta_{\mathbf{k}}}\right] \end{aligned} \tag{88}$$

Operator $\widehat{H}_B$ can therefore be written:

$$\begin{aligned} \widehat{H}_B = \sum_{\mathbf{k}\in D} \hbar\omega_k \Big\{ &\cosh 2\theta_{\mathbf{k}}\ \ \left[a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + a^\dagger_{-\mathbf{k}} a_{-\mathbf{k}}\right] \\ &+ 2\sinh^2\theta_{\mathbf{k}} + \sinh 2\theta_{\mathbf{k}} \left[a^\dagger_{-\mathbf{k}} a^\dagger_{\mathbf{k}}\ e^{2i\zeta_{\mathbf{k}}} + a_{\mathbf{k}} a_{-\mathbf{k}}\ e^{-2i\zeta_{\mathbf{k}}}\right] \Big\} \end{aligned} \tag{89}$$

Now this Hamiltonian may be identified with the approximate Hamiltonian (80). To see this, we replace $\cosh 2\theta_{\mathbf{k}}$ by expression (30), $\sinh 2\theta_{\mathbf{k}}$ by the double of (47), and $\sinh^2\theta_{\mathbf{k}}$ by (32); we are still in the simplified model where $\delta\mu = 0$, and hence the $\widetilde{e}_k$ are replaced by the e_k. Finally we choose for ω_k the value:

$$\hbar\omega_k = \sqrt{e_k\,(e_k + 2N_0 V_{\mathbf{k}})} \tag{90}$$

and we assume that all the $\zeta_{\mathbf{k}}$ are zero. We then get:

$$\begin{aligned} \widehat{H}_B = \sum_{\mathbf{k}\in D} \Big\{ &(e_k + N_0 V_{\mathbf{k}}) \left[a^\dagger_{\mathbf{k}} a_{\mathbf{k}} + a^\dagger_{-\mathbf{k}} a_{-\mathbf{k}}\right] \\ &+ \frac{N_0 V_{\mathbf{k}}}{2} \left[a^\dagger_{-\mathbf{k}} a^\dagger_{\mathbf{k}}\ + a_{\mathbf{k}} a_{-\mathbf{k}}\right] \Big\} - \Delta E_{\text{fond}} \end{aligned} \tag{91}$$

with, again using the value (32) for $\sinh^2\theta_{\mathbf{k}}$:

$$\Delta E_{\text{ground}} = -2\sum_{\mathbf{k}\in D}\hbar\omega_k\,\sinh^2\theta_{\mathbf{k}} = -\sum_{\mathbf{k}\neq\mathbf{0}}\hbar\omega_k\,\sinh^2\theta_{\mathbf{k}}$$
$$= \frac{1}{2}\sum_{\mathbf{k}\neq\mathbf{0}}\left[\sqrt{e_k\,(e_k+2N_0V_{\mathbf{k}})}-e_k-N_0V_{\mathbf{k}}\right] \tag{92}$$

Comparison with relations (52) and (54) shows that ΔE_{ground} is none other than the energy E_{ground} already obtained, shifted by the mean field value:

$$\Delta E_{\text{ground}} = E_{\text{ground}} - \frac{V_0N^2}{2} \tag{93}$$

Finally, taking (80) into account, if the $\zeta_{\mathbf{k}}$ are all chosen equal to zero (phase locking condition), we simply have:

$$\widehat{H}_R = \widehat{H}_B + E_{\text{ground}} \tag{94}$$

4-c. Constructing a basis of excited states, quasi-particles

As E_{ground} is a number, it introduces a simple energy shift in the eigenvalues of $\widehat{H}_R$ compared to those of $\widehat{H}_B$, with no effect on the eigenvectors. Now we saw in § E-3 of Chapter XVII that the eigenvalues of $\widehat{H}_B$ are known, and can be written as:

$$E_B = \sum_{\mathbf{k}\in D}\left[n(b_{\mathbf{k}})+n(b_{-\mathbf{k}})\right]\hbar\omega_k \tag{95}$$

where $n(b_{\mathbf{k}})$ and $n(b_{-\mathbf{k}})$ are any positive or zero integers. As for the eigenstates associated with these energies, they can be simply obtained by the action on the ground state of the following product of creation operators:

$$\prod_{\mathbf{k}\in D}\left[b^\dagger_{\mathbf{k}}\right]^{n(b_{\mathbf{k}})}\left[b^\dagger_{-\mathbf{k}}\right]^{n(b_{-\mathbf{k}})} \tag{96}$$

All things considered, the operator $\widehat{H}_B$ shares a lot of properties with the Hamiltonian of an ensemble of non-interacting particles. Just as the usual creation operators permit adding particles in a system of free identical particles, the $b^\dagger_{\mathbf{k}}$ and $b^\dagger_{-\mathbf{k}}$ creation operators can be considered as adding an extra "quasi-particle" to the physical system. When acting on the ground state, the operator $b^\dagger_{\mathbf{k}}$ yields a ket where both the energy and the momentum are well-defined: the energy is increased by the amount $\hbar\omega_k$ specified in (90), the momentum is increased by $\hbar\mathbf{k}$ with respect to the zero momentum of the ground state. This exact change of momentum occurs because the action of $b^\dagger_{\mathbf{k}}$ on any ket creates two components: one component where one particle with momentum $\hbar\mathbf{k}$ is added, the other component where one particle with momentum $-\hbar\mathbf{k}$ is suppressed. In both cases, the total momentum has increased[5] by the same amount $\hbar\mathbf{k}$. The operator $b^\dagger_{\mathbf{k}}$

[5]This result can also be verified by calculating the commutator $\left[\mathbf{P},b^\dagger_{\mathbf{k}}\right]$ of the total momentum $\mathbf{P}=\sum_{\mathbf{k}}\hbar\mathbf{k}\;a^\dagger_{\mathbf{k}}a_{\mathbf{k}}$ with $b^\dagger_{\mathbf{k}}$. One obtains $\left[\mathbf{P},b^\dagger_{\mathbf{k}}\right]=u^*_{\mathbf{k}}\left[\hbar\mathbf{k}a^\dagger_{\mathbf{k}}a_{\mathbf{k}},a^\dagger_{\mathbf{k}}\right]+v^*_{\mathbf{k}}\left[-\hbar\mathbf{k}a^\dagger_{-\mathbf{k}}a_{-\mathbf{k}},a_{-\mathbf{k}}\right]$, or: $\left[\mathbf{P},b^\dagger_{\mathbf{k}}\right]=\hbar\mathbf{k}\left(u^*_{\mathbf{k}}a^\dagger_{\mathbf{k}}+v^*_{\mathbf{k}}a_{-\mathbf{k}}\right)=\hbar\mathbf{k}\;b^\dagger_{\mathbf{k}}$. As a consequence, the effect of $b^\dagger_{\mathbf{k}}$ on any eigenstate of $\mathbf{P}$ is to increase its eigenvalue by $\hbar\mathbf{k}$.

therefore creates a quasi-particle of well-defined energy and momentum, and $b_{\mathbf{k}}$ annihilates it; of course, $b^{\dagger}_{-\mathbf{k}}$ and $b_{-\mathbf{k}}$ have the same properties for the quasi-particle of opposite momentum. These quasi-particles do not coincide with particles of a system without interactions, as can be seen from the expression of those creation operators. They yield, however, a basis of states that permits reasoning as if there were no interactions; this provides a very powerful point of view in many fields of physics.

We can assume, as in Complement $\mathrm{D_{XV}}$, that the interaction potential has a zero range:

$$W_2(\mathbf{r}, \mathbf{r}') = g\ \delta(\mathbf{r} - \mathbf{r}') \tag{97}$$

Relation (11) then becomes:

$$V_{\mathbf{k}} = \frac{1}{L^3} \int \mathrm{d}^3 r\ e^{-i\mathbf{q}\cdot\mathbf{r}}\ W_2(\mathbf{r}) = \frac{g}{L^3} \tag{98}$$

and equality (90) is written as:

$$\hbar\omega(\mathbf{k}) = \sqrt{e_k\,(e_k + 2gn_0)} = \frac{\hbar^2}{2m}\sqrt{k^2\,(k^2 + k_0^2)} \tag{99}$$

with:

$$k_0 = \frac{2}{\hbar}\sqrt{mgn_0} = \frac{\sqrt{2}}{\xi} \tag{100}$$

In this last equation, ξ is the healing length defined in (36). This equation is the same as relation (34) of Complement $\mathrm{D_{XV}}$, whose Figure 1 represents the quasi-particle spectrum. When the modulus k of the wave vector $\mathbf{k}$ is smaller than the wave vector k_0, we get a linear spectrum whose slope corresponds to the sound velocity in the boson system; for values larger than k_0, the spectrum becomes quadratic, as for free particles.

Conclusion

The calculations presented in this complement illustrate the analogy between the pairing phenomena for attractive fermions and for repulsive bosons. In both cases, binary position correlations are introduced by the dynamic interactions, resulting in a decrease of the interaction potential energy of the physical system; the paired states are a valuable tool for understanding this effect. In both cases, a relative phase locking phenomenon occurs, but the precise nature of that locking is, however, different.

For fermions, the energy gain is due to a collective effect, involving the pair-pair interactions and the relative phase of every pair of states $(\mathbf{k}, -\mathbf{k})$; each contributes to the value of the gap Δ which, in turn, has an effect on all the others – this is translated mathematically by the presence of a double summation over $\mathbf{k}$ and $\mathbf{k}'$ in the energy. This is reminiscent of a ferromagnetic system, where each spin contributes to the collective exchange field that act on all its neighbors. As the interactions are supposed to be attractive, the phase locking to zero maximizes the pair-pair interactions, and hence minimizes the energy.

For bosons, the major role is played by the relative phase of the pairs with respect to that of the reservoir composed of all the particles in the $\mathbf{k} = \mathbf{0}$ state (condensate).

The physical process involved is illustrated in Figure 7 of Complement B_{XVII}, where two particles emerge from the condensate to form a pair, or vice-versa – mathematically, the energy term contains only one summation over **k**. The relative phase locking it introduces will minimize the repulsion between these pairs and the condensate, and hence the total energy. Compared to the fermion case, the presence of a condensate independent of the pairs radically changes the nature of the phase locking.

Chapter XVIII

Review of classical electrodynamics

A Classical electrodynamics 1959
A-1 Basic equations and relations 1959
A-2 Description in the reciprocal space 1960
A-3 Elimination of the longitudinal fields from the expression of the physical quantities 1965
B Describing the transverse field as an ensemble of harmonic oscillators 1968
B-1 Brief review of the one-dimensional harmonic oscillator . . . 1968
B-2 Normal variables for the transverse field 1969
B-3 Discrete modes in a box 1974
B-4 Generalization of the mode concept 1975

Introduction

In the three previous chapters, we studied ensembles of identical particles, which allowed us to introduce the concept of quantum field operators. We now begin a new series of three chapters where this quantum field concept is applied to an important particular case: the electromagnetic field, made of identical bosons called "photons". We start by noting that, in classical electromagnetism, the dynamics of the different field modes is exactly similar to that of a series of harmonic oscillators. Each of these modes may be quantized by the same method as that used for an elementary harmonic oscillator, for a single particle; this method has the great advantage of simplicity. It requires, however, establishing beforehand the equivalence between modes of the classical electrodynamic field and harmonic oscillators; this is the main purpose of the present chapter.

For the presentation to be self-contained, we first review a certain number of properties of classical electromagnetism. One complement is also devoted to a synthetic

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.
© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

presentation of the Lagrangian formalism applied to this case. The reader already familiar with those aspects of classical electrodynamics may wish to go directly to the quantum treatment presented in Chapter XIX.

We start in § A with the equations of Maxwell-Lorentz describing the coupled evolution of the electric field $\boldsymbol{E}(\boldsymbol{r},t)$, the magnetic field $\boldsymbol{B}(\boldsymbol{r},t)$ and the coordinates and speeds of the particles acting as source for this electromagnetic field[1]. We shall give the expressions for a certain number of constants of motion, such as the energy, or the linear and angular momenta of the global system "field + particles". The vector potential $\boldsymbol{A}(\boldsymbol{r},t)$ and scalar potential $U(\boldsymbol{r},t)$ will also be introduced, as well as the gauge transformations that can be performed on these potentials.

We shall then show that it is useful to take the spatial Fourier transforms of these fields, since in the reciprocal space, Maxwell's equations have a simpler form. For a free electromagnetic field (in the absence of charged particles), they are no longer partial differential equations, as in ordinary space, but ordinary time-dependent differential equations. Furthermore, the concept of longitudinal or transverse field vectors has a clear geometrical significance in the reciprocal space[2]. A field vector $\tilde{\boldsymbol{V}}(\boldsymbol{k},t)$ is longitudinal if $\tilde{\boldsymbol{V}}(\boldsymbol{k},t)$ is parallel to $\boldsymbol{k}$ at every point $\boldsymbol{k}$ of the reciprocal space, transverse if $\tilde{\boldsymbol{V}}(\boldsymbol{k},t)$ is perpendicular to $\boldsymbol{k}$ at every point $\boldsymbol{k}$. We will show that two of the four Maxwell's equations yield the value of the longitudinal electrical and magnetic fields, whereas the other two describe the evolution of the transverse fields. It will become clear that the longitudinal electric field is simply the Coulomb electrostatic field created by the charged particles. Consequently, it is not an independent field variable since it only depends on the coordinates of the particles[3]. Furthermore, choosing the Coulomb gauge amounts to choosing the longitudinal potential vector equal to zero; this permits eliminating the longitudinal fields from the expressions for all the physical quantities.

In § B, we establish the equivalence between the radiation field and an ensemble of one-dimensional harmonic oscillators. Maxwell's equations for transverse fields enable introducing linear combinations of the vector potentials and transverse electric fields, whose time evolution, in the absence of particles, is of the form $e^{-i\omega t}$ where $\omega = ck$. These variables, called *normal variables*, thus describe the eigenmodes of the free field vibrations. The dynamics of each of these eigenmodes is similar to that of a one-dimensional harmonic oscillator. The normal mode variable is the equivalent of the linear combination of the position and velocity of the associated operator, and becomes, in the quantization process, the annihilation operator, fundamental in the quantum theory of the harmonic oscillator. Replacing the normal variables and their complex conjugates by annihilation and creation operators will yield, in Chapter XIX, the expressions for the various operators of the quantum theory.

[1] We assume that the speeds of the particles are small compared to the speed of light, so as to use a non-relativistic description.

[2] We shall note $\tilde{G}(\boldsymbol{k})$ the spatial Fourier transform of $G(\boldsymbol{r})$, the symbol "tilde" allowing a clear distinction between the functions in ordinary and reciprocal space.

[3] As for the longitudinal magnetic field, it is simply zero.

A. Classical electrodynamics

A-1. Basic equations and relations

A-1-a. Maxwell's equations

There are four Maxwell's equations in vacuum, and in the presence of sources:

$$\boldsymbol{\nabla} \cdot \boldsymbol{E}(\boldsymbol{r},t) = \frac{1}{\varepsilon_0}\rho(\boldsymbol{r},t) \tag{A-1a}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{B}(\boldsymbol{r},t) = 0 \tag{A-1b}$$

$$\boldsymbol{\nabla} \times \boldsymbol{E}(\boldsymbol{r},t) = -\frac{\partial}{\partial t}\boldsymbol{B}(\boldsymbol{r},t) \tag{A-1c}$$

$$\boldsymbol{\nabla} \times \boldsymbol{B}(\boldsymbol{r},t) = \frac{1}{c^2}\frac{\partial}{\partial t}\boldsymbol{E}(\boldsymbol{r},t) + \frac{1}{\varepsilon_0 c^2}\boldsymbol{j}(\boldsymbol{r},t) \tag{A-1d}$$

where c is the velocity of light in vacuum and ε_0 the vacuum permittivity. These equations yield the divergence and the curl of the electric field $\boldsymbol{E}(\boldsymbol{r},t)$ and the magnetic field $\boldsymbol{B}(\boldsymbol{r},t)$. The charge density $\rho(\boldsymbol{r},t)$ and current density $\boldsymbol{j}(\boldsymbol{r},t)$ appearing in those equations can be expressed, in the non-relativistic limit, in terms of the positions $\boldsymbol{r}_a(t)$ and the speeds $\boldsymbol{v}_a(t) = \mathrm{d}\boldsymbol{r}_a(t)/\mathrm{d}t$ of the various particles a of the system, each having a mass m_a and a charge q_a:

$$\rho(\boldsymbol{r},t) = \sum_a q_a \delta\left[\boldsymbol{r} - \boldsymbol{r}_a(t)\right] \tag{A-2a}$$

$$\boldsymbol{j}(\boldsymbol{r},t) = \sum_a q_a \boldsymbol{v}_a(t) \delta\left[\boldsymbol{r} - \boldsymbol{r}_a(t)\right] \tag{A-2b}$$

A-1-b. Lorentz Equations

Lorentz equations describe the dynamics of each particle a submitted to the electric and magnetic forces exerted by the fields $\boldsymbol{E}(\boldsymbol{r},t)$ and $\boldsymbol{B}(\boldsymbol{r},t)$:

$$m_a \frac{\mathrm{d}^2}{\mathrm{d}t^2}\boldsymbol{r}_a(t) = q_a\left[\boldsymbol{E}\left(\boldsymbol{r}_a(t),t\right) + \boldsymbol{v}_a(t) \times \boldsymbol{B}\left(\boldsymbol{r}_a(t),t\right)\right] \tag{A-3}$$

The particle and field evolutions are coupled: the particles move under the effect of the forces the fields exert on them, but they also act as sources for the evolution of those fields.

A-1-c. Constants of motion

Definitions (A-2a) of $\rho(\boldsymbol{r},t)$ and (A-2b) of $\boldsymbol{j}(\boldsymbol{r},t)$ lead to the continuity equation:

$$\frac{\partial}{\partial t}\rho(\boldsymbol{r},t) + \boldsymbol{\nabla} \cdot \boldsymbol{j}(\boldsymbol{r},t) = 0 \tag{A-4}$$

which implies the time invariance of the total charge of the particle system:

$$Q = \int \mathrm{d}^3 r \rho(\boldsymbol{r},t) = \sum_a q_a \tag{A-5}$$

Other constants of motion exist: the total energy H, the total momentum $\boldsymbol{P}$ and the total angular momentum $\boldsymbol{J}$ of the system field + particles. They are respectively given by:

$$H = \sum_a \frac{1}{2} m_a \boldsymbol{v}_a^2(t) + \frac{\varepsilon_0}{2} \int \mathrm{d}^3 r \left[\boldsymbol{E}^2(\boldsymbol{r},t) + c^2 \boldsymbol{B}^2(\boldsymbol{r},t) \right] \tag{A-6a}$$

$$\boldsymbol{P} = \sum_a m_a \boldsymbol{v}_a(t) + \varepsilon_0 \int \mathrm{d}^3 r \; \boldsymbol{E}(\boldsymbol{r},t) \times \boldsymbol{B}(\boldsymbol{r},t) \tag{A-6b}$$

$$\boldsymbol{J} = \sum_a \boldsymbol{r}_a(t) \times m_a \boldsymbol{v}_a(t) + \varepsilon_0 \int \mathrm{d}^3 r \; \boldsymbol{r} \times [\boldsymbol{E}(\boldsymbol{r},t) \times \boldsymbol{B}(\boldsymbol{r},t)] \tag{A-6c}$$

Using (A-1) and (A-3), we can verify that the derivatives with respect to time of H, $\boldsymbol{P}$ and $\boldsymbol{J}$ are indeed zero (for H and $\boldsymbol{P}$, see for example exercise 1 in Complement $\mathrm{C_I}$ of [16] and its correction).

A-1-d. Scalar and vector potentials: gauge transformations

As we already saw in Complement $\mathrm{H_{III}}$, the fields $\boldsymbol{E}(\boldsymbol{r},t)$ and $\boldsymbol{B}(\boldsymbol{r},t)$ can always be written in the form:

$$\boldsymbol{E}(\boldsymbol{r},t) = -\boldsymbol{\nabla} U(\boldsymbol{r},t) - \frac{\partial}{\partial t} \boldsymbol{A}(\boldsymbol{r},t) \tag{A-7a}$$

$$\boldsymbol{B}(\boldsymbol{r},t) = \boldsymbol{\nabla} \times \boldsymbol{A}(\boldsymbol{r},t) \tag{A-7b}$$

where $\boldsymbol{A}(\boldsymbol{r},t)$ and $U(\boldsymbol{r},t)$ are the vector and scalar potentials defining a *gauge*. For any function $\chi(\boldsymbol{r},t)$ of $\boldsymbol{r}$ and of t, the transformation of these potentials obeying the relations:

$$\boldsymbol{A}(\boldsymbol{r},t) \Rightarrow \boldsymbol{A}'(\boldsymbol{r},t) = \boldsymbol{A}(\boldsymbol{r},t) + \boldsymbol{\nabla} \chi(\boldsymbol{r},t) \tag{A-8a}$$

$$U(\boldsymbol{r},t) \Rightarrow U'(\boldsymbol{r},t) = U(\boldsymbol{r},t) - \frac{\partial}{\partial t} \chi(\boldsymbol{r},t) \tag{A-8b}$$

leads to the same expression for $\boldsymbol{E}(\boldsymbol{r},t)$ and $\boldsymbol{B}(\boldsymbol{r},t)$; the same physical fields can therefore be represented by several different potentials $\boldsymbol{A}(\boldsymbol{r},t)$ and $U(\boldsymbol{r},t)$. The transformation (A-8) associated with the function $\chi(\boldsymbol{r},t)$ is called a *gauge transformation*.

Relations (A-8) allow a flexibility on the choice of the gauge $\{\boldsymbol{A}, U\}$, which allows introducing an additional condition. The Coulomb gauge, which we will use in this chapter and the following, is defined by the condition:

$$\boldsymbol{\nabla} \cdot \boldsymbol{A}(\boldsymbol{r},t) = 0 \tag{A-9}$$

A geometrical interpretation of condition (A-9) in the reciprocal space will be given later.

A-2. Description in the reciprocal space

Using Fourier transforms, the equations of electrodynamics can be put in a form that simplifies calculations.

A-2-a. Spatial Fourier transforms

Let us introduce the Fourier transform of the electric field $\boldsymbol{E}(\boldsymbol{r},t)$:

$$\tilde{\boldsymbol{E}}(\boldsymbol{k},t) = \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3 r\, \boldsymbol{E}(\boldsymbol{r},t) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{A-10}$$

which enables us to write $\boldsymbol{E}(\boldsymbol{r},t)$ as:

$$\boldsymbol{E}(\boldsymbol{r},t) = \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3 k\, \tilde{\boldsymbol{E}}(\boldsymbol{k},t) e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{A-11}$$

Analogous expressions can be written for all the physical quantities we just introduced: magnetic field, charge and current densities, scalar and vector potentials.

It will be useful in what follows to recall the Parseval-Plancherel relation (Appendix I, § 2-c) showing the identity of the scalar products of two functions expressed in position space or in reciprocal space[4]:

$$\int \mathrm{d}^3 r\, F^*(\boldsymbol{r}) G(\boldsymbol{r}) = \int \mathrm{d}^3 k\, \tilde{F}^*(\boldsymbol{k}) \tilde{G}(\boldsymbol{k}) \tag{A-12}$$

and the fact that the product of two functions in reciprocal space, is the Fourier transform of their convolution in position space:

$$\tilde{F}(\boldsymbol{k})\tilde{G}(\boldsymbol{k}) \underset{\mathrm{FT}}{\Leftrightarrow} \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3 r'\, F(\boldsymbol{r}') G(\boldsymbol{r}-\boldsymbol{r}') \tag{A-13}$$

A-2-b. Maxwell's equations in reciprocal space

Maxwell's equations take on a simpler form in the reciprocal space, clearly showing the differences between the longitudinal and transverse components of the various fields. Any vector field $\tilde{\boldsymbol{V}}(\boldsymbol{k},t)$ can be decomposed into a longitudinal field $\tilde{\boldsymbol{V}}_{\parallel}(\boldsymbol{k},t)$, parallel at any point $\boldsymbol{k}$ to the vector $\boldsymbol{k}$, and a transverse field $\tilde{\boldsymbol{V}}_{\perp}(\boldsymbol{k},t)$ perpendicular to $\boldsymbol{k}$:

$$\tilde{\boldsymbol{V}}(\boldsymbol{k},t) = \tilde{\boldsymbol{V}}_{\parallel}(\boldsymbol{k},t) + \tilde{\boldsymbol{V}}_{\perp}(\boldsymbol{k},t) \tag{A-14}$$

with:

$$\tilde{\boldsymbol{V}}_{\parallel}(\boldsymbol{k},t) = \boldsymbol{\kappa}\, \left(\boldsymbol{\kappa}\cdot\tilde{\boldsymbol{V}}(\boldsymbol{k},t)\right) = \boldsymbol{k}\, \left(\boldsymbol{k}\cdot\tilde{\boldsymbol{V}}(\boldsymbol{k},t)\right)/k^2 \tag{A-15a}$$

$$\tilde{\boldsymbol{V}}_{\perp}(\boldsymbol{k},t) = \tilde{\boldsymbol{V}}(\boldsymbol{k},t) - \tilde{\boldsymbol{V}}_{\parallel}(\boldsymbol{k},t) \tag{A-15b}$$

where

$$\boldsymbol{\kappa} = \boldsymbol{k}/k \tag{A-16}$$

is the unit vector along $\boldsymbol{k}$.

[4]The space of the vectors $\boldsymbol{r}$ (ordinary space) is called "position space" whereas "reciprocal space" is the space of the wave vectors $\boldsymbol{k}$.

As the operator $\boldsymbol{\nabla}$ in position space corresponds to the operator $i\boldsymbol{k}$ in reciprocal space, Maxwell's equations (A-1) become in reciprocal space:

$$i\boldsymbol{k}\cdot\tilde{\boldsymbol{E}}(\boldsymbol{k},t)=\frac{1}{\varepsilon_0}\tilde{\rho}(\boldsymbol{k},t) \tag{A-17a}$$

$$i\boldsymbol{k}\cdot\tilde{\boldsymbol{B}}(\boldsymbol{k},t)=0 \tag{A-17b}$$

$$i\boldsymbol{k}\times\tilde{\boldsymbol{E}}(\boldsymbol{k},t)=-\frac{\partial}{\partial t}\tilde{\boldsymbol{B}}(\boldsymbol{k},t) \tag{A-17c}$$

$$i\boldsymbol{k}\times\tilde{\boldsymbol{B}}(\boldsymbol{k},t)=\frac{1}{c^2}\frac{\partial}{\partial t}\tilde{\boldsymbol{E}}(\boldsymbol{k},t)+\frac{1}{\varepsilon_0 c^2}\tilde{\boldsymbol{j}}(\boldsymbol{k},t) \tag{A-17d}$$

Taking into account definitions (A-15) for the longitudinal and transverse components of a vector field, the first two equations (A-17a) and (A-17b) determine the longitudinal parts, projections of the fields $\tilde{\boldsymbol{E}}(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{B}}(\boldsymbol{k},t)$ onto $\boldsymbol{k}$:

$$\tilde{\boldsymbol{E}}_{\parallel}(\boldsymbol{k},t)=-\frac{i}{\varepsilon_0}\tilde{\rho}(\boldsymbol{k},t)\frac{\boldsymbol{k}}{k^2} \tag{A-18a}$$

$$\tilde{\boldsymbol{B}}_{\parallel}(\boldsymbol{k},t)=\boldsymbol{0} \tag{A-18b}$$

The last two equations (A-17c) and (A-17d) yield the rate of change $\partial\tilde{\boldsymbol{E}}(\boldsymbol{k},t)/\partial t$ and $\partial\tilde{\boldsymbol{B}}(\boldsymbol{k},t)/\partial t$ of the fields $\tilde{\boldsymbol{E}}(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{B}}(\boldsymbol{k},t)$, and are the equations of motion of these fields. In the absence of sources ($\tilde{\boldsymbol{j}}(\boldsymbol{k},t)=\boldsymbol{0}$), i.e. for what we will call a "free" field, they are time-dependent differential equations, and no longer partial derivative equations as is the case in position space.

A-2-c. Longitudinal electric and magnetic fields

Equation (A-18b) shows the longitudinal magnetic field $\tilde{\boldsymbol{B}}_{\parallel}(\boldsymbol{k},t)$ is zero. Equation (A-18a) expresses $\tilde{\boldsymbol{E}}_{\parallel}(\boldsymbol{k},t)$ as a product of two functions of $\boldsymbol{k}$, $\tilde{\rho}(\boldsymbol{k},t)$ and $-i\boldsymbol{k}/\varepsilon_0 k^2$ whose Fourier transforms are written (relation (63) of Appendix I):

$$\tilde{\rho}(\boldsymbol{k},t)\ \underset{\text{FT}}{\leftrightarrow}\ \rho(\boldsymbol{r},t) \tag{A-19a}$$

$$-\frac{i}{\varepsilon_0}\frac{\boldsymbol{k}}{k^2}\ \underset{\text{FT}}{\leftrightarrow}\ \frac{(2\pi)^{3/2}}{4\pi\varepsilon_0}\frac{\boldsymbol{r}}{r^3} \tag{A-19b}$$

Using relation (A-13) then leads to:

$$\begin{aligned}\boldsymbol{E}_{\parallel}(\boldsymbol{r},t)&=\frac{1}{4\pi\varepsilon_0}\int \mathrm{d}^3r'\rho(\boldsymbol{r}',t)\frac{\boldsymbol{r}-\boldsymbol{r}'}{|\boldsymbol{r}-\boldsymbol{r}'|^3}\\&=\frac{1}{4\pi\varepsilon_0}\sum_a q_a\frac{\boldsymbol{r}-\boldsymbol{r}_a(t)}{|\boldsymbol{r}-\boldsymbol{r}_a(t)|^3}\end{aligned} \tag{A-20}$$

This means that at time t, the longitudinal electric field coincides with the Coulomb field produced by the charge distribution $\rho(\boldsymbol{r},t)$, computed as if this distribution were static and fixed at that instant t.

Comment

The fact that the longitudinal electric field instantaneously follows the evolution of the charge distribution $\rho(\boldsymbol{r},t)$ should not lead us to believe in an *action at a distance* propagating at an infinite speed. The contribution of the transverse field must also be taken into account, as only the total electric field $\boldsymbol{E}=\boldsymbol{E}_{\parallel}+\boldsymbol{E}_{\perp}$ has a real physical meaning. It can be shown that the transverse electric field also has an instantaneous component, which balances exactly the longitudinal component so that the total field is always retarded (to $t-|r-r'|/c$), as the electromagnetic interactions propagate at the speed of light c (see exercise 3 and its correction in Complement C_I of reference [16]).

The previous results show that the longitudinal fields are not independent quantities: they are either zero (in the case of the longitudinal magnetic field), or simply related to the particle coordinates $\boldsymbol{r}_a(t)$ (in the case of the longitudinal electric field, whose expression is given by (A-20)).

A-2-d. Time evolution of the transverse fields

Now that we showed that the first two Maxwell's equations determine the longitudinal part of the fields, let us consider the last two equations (A-17c) and (A-17d) and focus on their transverse components. Since $\boldsymbol{k}\times\boldsymbol{E}=\boldsymbol{k}\times\boldsymbol{E}_{\perp}$, they can be rewritten as:

$$\frac{\partial}{\partial t}\tilde{\boldsymbol{B}}(\boldsymbol{k},t)=-i\boldsymbol{k}\times\tilde{\boldsymbol{E}}_{\perp}(\boldsymbol{k},t) \tag{A-21a}$$

$$\frac{\partial}{\partial t}\tilde{\boldsymbol{E}}_{\perp}(\boldsymbol{k},t)=ic^2\boldsymbol{k}\times\tilde{\boldsymbol{B}}(\boldsymbol{k},t)-\frac{1}{\varepsilon_0}\tilde{\boldsymbol{j}}_{\perp}(\boldsymbol{k},t) \tag{A-21b}$$

which yield the time evolution of the transverse fields $\tilde{\boldsymbol{E}}_{\perp}(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{B}}(\boldsymbol{k},t)$.

Comment

One can also study the longitudinal projections of the two Maxwell's equations (A-17c) and (A-17d). The result is trivial for the first one: as both sides of the equation are transverse, their longitudinal projections are zero. As for the second equation, (A-17d), it leads to:

$$\frac{\partial}{\partial t}\tilde{\boldsymbol{E}}_{\parallel}(\boldsymbol{k},t)+\frac{1}{\varepsilon_0}\tilde{\boldsymbol{j}}_{\parallel}(\boldsymbol{k},t)=0. \tag{A-22}$$

Taking the scalar product of $\boldsymbol{k}$ with each side of this equation, using (A-18a) and the fact that $\boldsymbol{k}\cdot\tilde{\boldsymbol{j}}=\boldsymbol{k}\cdot\tilde{\boldsymbol{j}}_{\parallel}$, we find:

$$\frac{\partial}{\partial t}\tilde{\rho}(\boldsymbol{k},t)+i\boldsymbol{k}\cdot\tilde{\boldsymbol{j}}(\boldsymbol{k},t)=0 \tag{A-23}$$

which is simply the continuity equation (A-4) in the reciprocal space, and does not provide any new information.

A-2-e. Potentials

In the reciprocal space, relations (A-7a) and (A-7b) between fields and potentials become:

$$\tilde{\boldsymbol{E}}(\boldsymbol{k},t)=-i\boldsymbol{k}\tilde{U}(\boldsymbol{k},t)-\frac{\partial}{\partial t}\tilde{\boldsymbol{A}}(\boldsymbol{k},t) \tag{A-24a}$$

$$\tilde{\boldsymbol{B}}(\boldsymbol{k},t)=i\boldsymbol{k}\times\tilde{\boldsymbol{A}}(\boldsymbol{k},t) \tag{A-24b}$$

and the gauge transformations relations (A-8a) and (A-8b) are written:

$$\tilde{\boldsymbol{A}}(\boldsymbol{k},t) \to \tilde{\boldsymbol{A}}'(\boldsymbol{k},t) = \tilde{\boldsymbol{A}}(\boldsymbol{k},t) + i\boldsymbol{k}\tilde{\chi}(\boldsymbol{k},t) \tag{A-25a}$$

$$\tilde{U}(\boldsymbol{k},t) \to \tilde{U}'(\boldsymbol{k},t) = \tilde{U}(\boldsymbol{k},t) - \frac{\partial}{\partial t}\tilde{\chi}(\boldsymbol{k},t) \tag{A-25b}$$

where $\tilde{\chi}(\boldsymbol{k},t)$ is the Fourier transform of $\chi(\boldsymbol{r},t)$.

Since the last term in (A-25a) is a longitudinal vector, it is clear that a gauge transformation does not change the transverse part $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$, which thus defines a gauge invariant physical field:

$$\tilde{\boldsymbol{A}}'_\perp(\boldsymbol{k},t) = \tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) \tag{A-26}$$

Since $\boldsymbol{k} \times \tilde{\boldsymbol{A}}_\parallel = \boldsymbol{0}$, the transverse projections of relations (A-24a) and (A-24b) yield the equations:

$$\tilde{\boldsymbol{E}}_\perp(\boldsymbol{k},t) = -\frac{\partial}{\partial t}\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) \tag{A-27a}$$

$$\tilde{\boldsymbol{B}}(\boldsymbol{k},t) = i\boldsymbol{k} \times \tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) \tag{A-27b}$$

Note that equation (A-27b) allows expressing $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$ as a function of $\tilde{\boldsymbol{B}}(\boldsymbol{k},t)$, as we now show. Taking the vector product of $\boldsymbol{k}$ with each side of this equation, and using the identity:

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\boldsymbol{c}, \tag{A-28}$$

and the fact that $\boldsymbol{k} \cdot \tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) = 0$, we get:

$$\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) = \frac{i}{k^2}\left(\boldsymbol{k} \times \tilde{\boldsymbol{B}}(\boldsymbol{k},t)\right) \tag{A-29}$$

This equation, together with equation (A-27a), allow rewriting the two time evolution equations (A-21a) and (A-21b) for the transverse fields in a form only involving $\tilde{\boldsymbol{E}}_\perp(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$:

$$\frac{\partial}{\partial t}\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) = -\tilde{\boldsymbol{E}}_\perp(\boldsymbol{k},t) \tag{A-30a}$$

$$\frac{\partial}{\partial t}\tilde{\boldsymbol{E}}_\perp(\boldsymbol{k},t) = c^2k^2\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) - \frac{1}{\varepsilon_0}\tilde{\boldsymbol{j}}_\perp(\boldsymbol{k},t) \tag{A-30b}$$

In the absence of sources ($\tilde{\boldsymbol{j}}_\perp(\boldsymbol{k},t) = \boldsymbol{0}$), we get two coupled time evolution equations for the transverse fields $\tilde{\boldsymbol{E}}_\perp(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$. They will be useful later on for introducing the field normal variables, and for the demonstration of the equivalence between the transverse field and an ensemble of harmonic oscillators.

Time evolution equation for the transverse potential vector

The time evolution equation for $\tilde{\boldsymbol{A}}_\perp$ can be obtained by replacing $\tilde{\boldsymbol{E}}_\perp$ in (A-30b) by $-\partial\tilde{\boldsymbol{A}}_\perp/\partial t$. We obtain:

$$\left[\frac{\partial^2}{\partial t^2} + c^2k^2\right]\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) = \frac{1}{\varepsilon_0}\tilde{\boldsymbol{j}}_\perp(\boldsymbol{k},t) \tag{A-31}$$

which is written, in the position space:

$$\left[\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \Delta\right]\boldsymbol{A}_\perp(\boldsymbol{r},t) = \frac{1}{\varepsilon_0 c^2}\boldsymbol{j}_\perp(\boldsymbol{r},t) \tag{A-32}$$

A-2-f. Coulomb gauge

Condition $\boldsymbol{\nabla} \cdot \boldsymbol{A}(\boldsymbol{r},t) = 0$, which defined in (A-9) the Coulomb gauge, becomes in the reciprocal space:

$$i\boldsymbol{k} \cdot \tilde{\boldsymbol{A}}(\boldsymbol{k},t) = 0 \quad \longleftrightarrow \quad \tilde{\boldsymbol{A}}_{\parallel}(\boldsymbol{k},t) = 0 \tag{A-33}$$

In the Coulomb gauge, the longitudinal vector potential is therefore equal to zero; there only remains the transverse vector potential, which, as mentioned above, is a physical field.

What can be said about the scalar potential U in the Coulomb gauge? Let us consider the longitudinal part of each side of equation (A-24a). As the last term on the right-hand side is transverse in the Coulomb gauge, we get $\tilde{\boldsymbol{E}}_{\parallel}(\boldsymbol{k},t) = -i\boldsymbol{k}\tilde{U}(\boldsymbol{k},t)$, which reads, in position space, $\boldsymbol{E}_{\parallel}(\boldsymbol{r},t) = -\boldsymbol{\nabla}U(\boldsymbol{r},t)$. The scalar potential is the potential whose gradient yields the longitudinal electric field. Equation (A-20) then shows that, to within a constant, $U(\boldsymbol{r},t)$ is equal to:

$$U(\boldsymbol{r},t) = \frac{1}{4\pi\varepsilon_0} \sum_a q_a \frac{1}{|\boldsymbol{r} - \boldsymbol{r}_a(t)|} \tag{A-34}$$

which is the Coulomb potential created by the charge distribution.

Lorenz gauge

In the present chapter and the next one, we shall mainly use the Coulomb gauge. Another gauge often used, in particular in the clearly covariant formulations of electrodynamics, is the *Lorentz gauge*[5] defined by the condition:

$$\boldsymbol{\nabla} \cdot \boldsymbol{A}(\boldsymbol{r},t) + \frac{1}{c^2}\frac{\partial}{\partial t} U(\boldsymbol{r},t) = 0 \tag{A-35}$$

which can be written, using covariant notation:

$$\sum_{\mu} \partial_\mu A^\mu = 0 \tag{A-36}$$

The condition defining the Lorenz gauge thus keeps the same form in every Lorentz reference frame, which is not the case for the Coulomb gauge (since in relativity, a transverse field of zero divergence in one reference frame is no longer necessarily transverse in another frame). Nevertheless, an advantage of the Coulomb gauge is that it allows the immediate identification, in a given reference frame, of the field variables that are really independent.

A-3. Elimination of the longitudinal fields from the expression of the physical quantities

It will be useful for the following discussion to eliminate the longitudinal fields from the expressions of the total energy H and the total momentum given by equations (A-6a) and (A-6b). We shall express these physical quantities only in terms of the truly independent variables, such as particle coordinates and speeds, and transverse fields.

[5]The danish physicist Ludwig Lorenz is often confused with the dutch physicist Hendrik Lorentz.

A-3-a. Total energy

We start by eliminating the longitudinal electric field from the last term in expression (A-6a). Using the Parseval-Plancherel equality (A-12) and the fact that $\tilde{\boldsymbol{E}}^*_{\parallel}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{E}}_{\perp}(\boldsymbol{k},t)=0$, we can rewrite this term as:

$$\frac{\varepsilon_0}{2}\int \mathrm{d}^3r\left[\boldsymbol{E}^2(\boldsymbol{r},t)+c^2\boldsymbol{B}^2(\boldsymbol{r},t)\right]=H_{\mathrm{long}}+H_{\mathrm{trans}} \tag{A-37}$$

where:

$$H_{\mathrm{long}}=\frac{\varepsilon_0}{2}\int \mathrm{d}^3k\;\tilde{\boldsymbol{E}}^*_{\parallel}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{E}}_{\parallel}(\boldsymbol{k},t) \tag{A-38a}$$

$$H_{\mathrm{trans}}=\frac{\varepsilon_0}{2}\int \mathrm{d}^3k\left[\tilde{\boldsymbol{E}}^*_{\perp}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{E}}_{\perp}(\boldsymbol{k},t)+c^2\tilde{\boldsymbol{B}}^*(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{B}}(\boldsymbol{k},t)\right] \tag{A-38b}$$

In (A-38a), we replace $\tilde{\boldsymbol{E}}_{\parallel}(\boldsymbol{k},t)$ by expression (A-18a). We get, taking (A-12) and (A-13) into account:

$$\begin{aligned}H_{\mathrm{long}}&=\frac{1}{2\varepsilon_0}\int \mathrm{d}^3k\;\frac{\tilde{\rho}^*(\boldsymbol{k},t)\tilde{\rho}(\boldsymbol{k},t)}{k^2}\\&=\frac{1}{8\pi\varepsilon_0}\int\int \mathrm{d}^3r\,\mathrm{d}^3r'\;\frac{\tilde{\rho}(\boldsymbol{r},t)\tilde{\rho}(\boldsymbol{r}',t)}{|\boldsymbol{r}-\boldsymbol{r}'|}\\&=\sum_a h^a_{\mathrm{Coul}}+\frac{1}{8\pi\epsilon_0}\sum_{a\neq b}\frac{q_aq_b}{|\boldsymbol{r}_a-\boldsymbol{r}_b|}=V_{\mathrm{Coul}}\end{aligned} \tag{A-39}$$

The longitudinal field energy is thus equal to the Coulomb electrostatic energy V_{Coul} of the charge distribution $\rho(\boldsymbol{r},t)$. In addition to the Coulomb interaction energy between different particles a and b, V_{Coul} also contains the energy h^a_{Coul} of the Coulomb field of each particle a, which diverges for point particles.

Expression (A-38b) for H_{trans} can be rewritten as a function of the variables $\tilde{\boldsymbol{E}}_{\perp}(\boldsymbol{k},t)=-\dot{\tilde{\boldsymbol{A}}}_{\perp}(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)$ introduced above for the transverse field:

$$H_{\mathrm{trans}}=\frac{\varepsilon_0}{2}\int \mathrm{d}^3k\;\left[\dot{\tilde{\boldsymbol{A}}}^*_{\perp}(\boldsymbol{k},t)\cdot\dot{\tilde{\boldsymbol{A}}}_{\perp}(\boldsymbol{k},t)+\omega^2\tilde{\boldsymbol{A}}^*_{\perp}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)\right] \tag{A-40}$$

Finally, the energy of the global system field + particles can be expressed in the form:

$$H=\frac{1}{2}\sum_a m_a\dot{\boldsymbol{r}}^2_a(t)+V_{\mathrm{Coul}}+H_{\mathrm{trans}} \tag{A-41}$$

where we used the simplified notation $\dot{\boldsymbol{r}}_a(t)=\mathrm{d}\boldsymbol{r}_a(t)/\mathrm{d}t=\boldsymbol{v}_a(t)$. It is the sum of the kinetic energy of the particles, of their Coulomb energy, and of the energy of the transverse field.

A-3-b. Total momentum

Similar computations can be carried out for the total momentum $\boldsymbol{P}$. The field contribution contained in the last term of (A-6b) can be written as:

$$\varepsilon_0 \int \mathrm{d}^3k\ \tilde{\boldsymbol{E}}^*(\boldsymbol{k},t)\times\tilde{\boldsymbol{B}}(\boldsymbol{k},t) = \underbrace{\varepsilon_0 \int \mathrm{d}^3k\ \tilde{\boldsymbol{E}}^*_{\parallel}(\boldsymbol{k},t)\times\tilde{\boldsymbol{B}}(\boldsymbol{k},t)}_{\boldsymbol{P}_{\mathrm{long}}} + \underbrace{\varepsilon_0 \int \mathrm{d}^3k\ \tilde{\boldsymbol{E}}^*_{\perp}(\boldsymbol{k},t)\times\tilde{\boldsymbol{B}}(\boldsymbol{k},t)}_{\boldsymbol{P}_{\mathrm{trans}}} \tag{A-42}$$

where we have separated the contributions to $\boldsymbol{P}$ coming from the longitudinal and transverse components of the electric field[6]. Using (A-18a) and (A-27b), taking into account identity (A-28) and the fact that $\boldsymbol{k}\cdot\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)=0$, we get:

$$\begin{aligned}\boldsymbol{P}_{\mathrm{long}} &= \varepsilon_0 \int \mathrm{d}^3k\ \frac{i\tilde{\rho}^*(\boldsymbol{k},t)}{\varepsilon_0}\frac{\boldsymbol{k}}{k^2}\times\big(i\boldsymbol{k}\times\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)\big)\\ &= \int \mathrm{d}^3k\ \tilde{\rho}^*(\boldsymbol{k},t)\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)\end{aligned} \tag{A-43}$$

We then have:

$$\begin{aligned}\boldsymbol{P}_{\mathrm{long}} &= \int \mathrm{d}^3r\ \rho(\boldsymbol{r},t)\boldsymbol{A}_{\perp}(\boldsymbol{r},t)\\ &= \sum_a q_a\boldsymbol{A}_{\perp}(\boldsymbol{r}_a,t).\end{aligned} \tag{A-44}$$

As we did above for (A-40), we can rewrite the expression of $\boldsymbol{P}_{\mathrm{trans}}$ as a function of the variables $\tilde{\boldsymbol{E}}_{\perp}(\boldsymbol{k},t)=-\dot{\tilde{\boldsymbol{A}}}_{\perp}(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)$ of the transverse field:

$$\begin{aligned}\boldsymbol{P}_{\mathrm{trans}} &= -\varepsilon_0\int \mathrm{d}^3k\ \dot{\tilde{\boldsymbol{A}}}^*_{\perp}(\boldsymbol{k},t)\times\big[i\boldsymbol{k}\times\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)\big]\\ &= -i\varepsilon_0\int \mathrm{d}^3k\ \ \boldsymbol{k}\ \ \big[\dot{\tilde{\boldsymbol{A}}}^*_{\perp}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)\big]\end{aligned} \tag{A-45}$$

The momentum of the global system field + particles can be written in the form:

$$\boldsymbol{P} = \sum_a \left[m_a\dot{\boldsymbol{r}}_a(t) + q_a\boldsymbol{A}_{\perp}(\boldsymbol{r}_a,t)\right] + \boldsymbol{P}_{\mathrm{trans}} \tag{A-46}$$

Let us finally introduce the quantity:

$$\boldsymbol{p}_a(t) = m_a\dot{\boldsymbol{r}}_a(t) + q_a\boldsymbol{A}_{\perp}(\boldsymbol{r}_a,t) \tag{A-47}$$

We shall see later that, in the Coulomb gauge electrodynamics, $\boldsymbol{p}_a(t)$ is the conjugate momentum of $\boldsymbol{r}_a(t)$, hence different from the mechanical momentum $m_a\dot{\boldsymbol{r}}_a(t)$. Expressed

[6]The notation $\boldsymbol{P}_{\mathrm{long}}$ should not lead us to believe that $\boldsymbol{P}_{\mathrm{long}}$ is a longitudinal field vector itself: it is actually the vector yielding the longitudinal electric field contribution to the momentum vector; the same comment applies to $\boldsymbol{P}_{\mathrm{trans}}$.

as a function of $\boldsymbol{p}_a(t)$, the total energy (A-41) and total momentum (A-46) are written as:

$$H = \frac{1}{2m_a}\sum_a \left[\boldsymbol{p}_a - q_a \boldsymbol{A}_\perp(\boldsymbol{r}_a, t)\right]^2 + V_{\text{Coul}} + H_{\text{trans}} \tag{A-48}$$

$$\boldsymbol{P} = \sum_a \boldsymbol{p}_a(t) + \boldsymbol{P}_{\text{trans}} \tag{A-49}$$

where H_{trans} and $\boldsymbol{P}_{\text{trans}}$ were introduced in equations (A-38b) and (A-42). We shall see that H actually coincides with the Hamiltonian in the Coulomb gauge of the global system field + particles.

A-3-c. Total angular momentum

Calculations similar to ones just presented, but that will not be detailed here[7], show that the contribution of the longitudinal electric field to the total angular momentum is equal to:

$$\boldsymbol{J}_{\text{long}} = \varepsilon_0 \int \mathrm{d}^3 r\, \boldsymbol{r} \times (\boldsymbol{E}_\parallel \times \boldsymbol{B}) = \sum_a q_a \boldsymbol{r}_a \times \boldsymbol{A}_\perp(\boldsymbol{r}_a). \tag{A-50}$$

Adding $\boldsymbol{J}_{\text{long}}$ to the particles' angular momenta, we get, taking (A-47) into account:

$$\sum_a \boldsymbol{r}_a \times m_a \dot{\boldsymbol{r}}_a + \boldsymbol{J}_{\text{long}} = \sum_a \boldsymbol{r}_a \times \boldsymbol{p}_a \tag{A-51}$$

so that we can finally write:

$$\boldsymbol{J} = \sum_a \boldsymbol{r}_a \times \boldsymbol{p}_a + \boldsymbol{J}_{\text{trans}} \tag{A-52}$$

where:

$$\boldsymbol{J}_{\text{trans}} = \varepsilon_0 \int \mathrm{d}^3 r\, [\boldsymbol{r} \times (\boldsymbol{E}_\perp \times \boldsymbol{B})] \tag{A-53}$$

B. Describing the transverse field as an ensemble of harmonic oscillators

B-1. Brief review of the one-dimensional harmonic oscillator

The energy of a harmonic oscillator is given by:

$$E = \frac{1}{2} m \dot{x}^2 + \frac{1}{2} m \omega^2 x^2 \tag{B-1}$$

where $\omega/2\pi$ is the oscillation frequency, and $\dot{x}$ the oscillator velocity:

$$\frac{\mathrm{d}}{\mathrm{d}t}\, x = \dot{x} \tag{B-2}$$

[7] These calculations can be found in § 1 of Complement $\mathrm{B_I}$ in [16].

This velocity obeys:

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\dot{x} = -\omega^2 x \tag{B-3}$$

so that the equation of motion of x is:

$$\ddot{x} + \omega^2 x = 0 \tag{B-4}$$

Consequently, the time evolution of $x(t)$ is given by a (real) linear combination of $\cos(\omega t)$ and $\sin(\omega t)$.

The dynamic state of the classical harmonic oscillator is defined at each instant by two real variables $x(t)$ and $\dot{x}(t)$. It is often useful to combine them into a single complex variable $\alpha(t)$ by setting:

$$\alpha(t) = C\left[x(t) + i\,\frac{\dot{x}(t)}{\omega}\right] \tag{B-5}$$

where C is an arbitrary (time-independent) constant. Relations (B-2) and (B-3) show that $\alpha(t)$ obeys the first order differential equation:

$$\dot{\alpha} = C(\dot{x} - i\omega x) = -i\omega C\left(x + i\frac{\dot{x}}{\omega}\right) = -i\omega\alpha \tag{B-6}$$

The time dependence of the new variable $\alpha(t)$ is therefore simply $e^{-i\omega t}$.

One can invert the system formed by equation (B-5) and its complex conjugate yielding α^*, and compute x and $\dot{x}$ as a function of α and α^*. Inserting the expressions thus obtained in equation (B-1) for the energy E, we obtain by a simple calculation[8]:

$$E = \frac{m\omega^2}{4C^2}(\alpha^*\alpha + \alpha\alpha^*) \tag{B-7}$$

The constant C can be chosen so that:

$$\frac{m\omega^2}{4C^2} = \frac{\hbar\omega}{2} \tag{B-8}$$

This leads, after quantization, to the Hamiltonian operator:

$$\hat{H} = \frac{\hbar\omega}{2}(\hat{a}^\dagger\hat{a} + \hat{a}\hat{a}^\dagger) \tag{B-9}$$

which is the Hamiltonian of a harmonic oscillator[9].

B-2. Normal variables for the transverse field

B-2-a. Vibration eigenmodes of the free transverse field

In the reciprocal space, expression (A-40) for the free transverse field energy H_{trans} is a sum of quadratic functions of $\dot{\tilde{\boldsymbol{A}}}_\perp(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$. For each value of $\boldsymbol{k}$, we get a harmonic oscillator Hamiltonian. The evolution introduces no coupling between the various spatial Fourier components of the transverse field. We see the advantage of

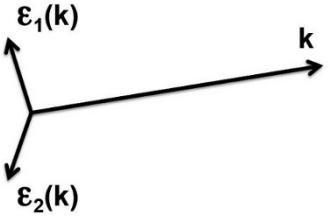


Figure 1: For each vector $\boldsymbol{k}$, the transverse fields can have two polarizations characterized by unit vectors $\boldsymbol{\varepsilon}_1(\boldsymbol{k})$ and $\boldsymbol{\varepsilon}_2(\boldsymbol{k})$ perpendicular both to each other and to $\boldsymbol{k}$.

working in the reciprocal space: it enables us to identify the eigenmodes of the field vibrations, in the absence of sources.

Actually, for each $\boldsymbol{k}$, the transverse field can have two different polarizations [10] characterized by unit vectors $\boldsymbol{\varepsilon}_1(\boldsymbol{k})$ and $\boldsymbol{\varepsilon}_2(\boldsymbol{k})$, both perpendicular to $\boldsymbol{k}$ and to each other, so that we can write for $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$, as an example:

$$\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) = \tilde{A}_{\perp\boldsymbol{\varepsilon}_1(\boldsymbol{k})}(\boldsymbol{k},t)\,\boldsymbol{\varepsilon}_1(\boldsymbol{k}) + \tilde{A}_{\perp\boldsymbol{\varepsilon}_2(\boldsymbol{k})}(\boldsymbol{k},t)\,\boldsymbol{\varepsilon}_2(\boldsymbol{k}) = \sum_{\boldsymbol{\varepsilon}_i(\boldsymbol{k})} \tilde{A}_{\perp\boldsymbol{\varepsilon}_i(\boldsymbol{k})}(\boldsymbol{k},t)\,\boldsymbol{\varepsilon}_i(\boldsymbol{k}) \quad \text{(B-10)}$$

with:

$$\tilde{A}_{\perp\boldsymbol{\varepsilon}_i(\boldsymbol{k})}(\boldsymbol{k},t) = \boldsymbol{\varepsilon}_i(\boldsymbol{k}) \cdot \tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) \quad \text{(B-11)}$$

The set $\{\boldsymbol{k}, \boldsymbol{\varepsilon}_i(\boldsymbol{k})\}$ defines what we shall call in this chapter a free field *mode*; they are the eigenmodes of the free field vibration, with a frequency:

$$\omega = ck \quad \text{(B-12)}$$

To simplify the notation, we shall write the last summation in (B-10) in a more compact form:

$$\sum_{\boldsymbol{\varepsilon}_i(\boldsymbol{k})} \tilde{A}_{\perp\boldsymbol{\varepsilon}_i(\boldsymbol{k})}(\boldsymbol{k},t)\,\boldsymbol{\varepsilon}_i(\boldsymbol{k}) \equiv \sum_{\boldsymbol{\varepsilon}} \tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\boldsymbol{\varepsilon} \quad \text{(B-13)}$$

Let us rewrite expression (A-40) for $\tilde{H}_{\text{trans}}$ expliciting the components of the fields $\boldsymbol{A}_\perp(\boldsymbol{k},t)$ and $\dot{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$ on the polarization vectors. We get:

$$H_{\text{trans}} = \frac{\varepsilon_0}{2} \int \mathrm{d}^3k \sum_{\boldsymbol{\varepsilon}} \left[\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) + \omega^2 \tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \right] \quad \text{(B-14)}$$

[8] In view of the quantization where α and α^* will be replaced by non-commuting operators $\hat{a}$ and $\hat{a}^\dagger$, we keep the sequence of α and α^* as they appear in the computations.

[9] If $\hat{x}$ and $\hat{p}$ obey the canonical commutation relation $[\hat{x},\hat{p}] = i\hbar$, relation (B-8) for the choice of C also leads to the commutation relation $[\hat{a},\hat{a}^\dagger] = 1$.

[10] We choose real vectors $\boldsymbol{\varepsilon}_1(\boldsymbol{k})$ and $\boldsymbol{\varepsilon}_2(\boldsymbol{k})$ corresponding to linear polarizations, but the choice of these two polarizations is arbitrary, since they can always be rotated by any angle around $\boldsymbol{k}$. It is also possible to perform a more general change of basis with complex vectors defining elliptical polarizations, for instance the right and left circular polarizations $\boldsymbol{\varepsilon}_\pm = (\boldsymbol{\varepsilon}_1 \pm i\boldsymbol{\varepsilon}_2)/\sqrt{2}$. Circular polarizations are useful when discussing electromagnetic spin — see § 3 of Complement $\mathrm{B}_{\mathrm{XIX}}$. If complex (orthonormal) polarizations are used, $\boldsymbol{\varepsilon}_i(\boldsymbol{k})$ should be replaced by $\boldsymbol{\varepsilon}^*_i(\boldsymbol{k})$ in the right side hand of relation (B-11).

Note that the components on the two polarizations $\boldsymbol{\varepsilon}$ are truly independent dynamic variables (generalized coordinates and velocities). This is not the case for the Cartesian components $\tilde{A}_{\perp i}(\boldsymbol{k},t)$ and $\dot{\tilde{A}}_{\perp i}(\boldsymbol{k},t)$ (with $i=x,y,z$), because of the transversality condition. For example, the components $\tilde{A}_{\perp i}(\boldsymbol{k},t)$ must obey $\sum_i k_i \tilde{A}_{\perp i}=0$.

Constraints on the dynamic variables in the reciprocal space

Since the fields are real in real space, we have the condition $\tilde{\boldsymbol{A}}^*_\perp(\boldsymbol{k},t) = \tilde{\boldsymbol{A}}_\perp(-\boldsymbol{k},t)$. In half the reciprocal space, the variables $\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ and $\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ can be considered as independent .

B-2-b. Definition of the normal variables, free field case

Let us first assume that we are in the free field case ($\tilde{\boldsymbol{j}}_\perp = \boldsymbol{0}$), and we can replace the field $\tilde{\boldsymbol{E}}_\perp(\boldsymbol{k},t)$ by $-\dot{\tilde{\boldsymbol{A}}}_\perp(\boldsymbol{k},t)$ in equations (A-30a) and (A-30b). As $\omega = kc$, we get two equations exactly similar to those of a harmonic oscillator (B-2) and (B-3), with $\boldsymbol{A}_\perp(\boldsymbol{k},t)$ instead of $x(t)$. This analogy suggests introducing, as in (B-5), a new transverse variable:

$$\begin{aligned}\boldsymbol{\alpha}(\boldsymbol{k},t) &= \mathcal{N}(k)\left[\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) + \frac{i}{\omega}\dot{\tilde{\boldsymbol{A}}}_\perp(\boldsymbol{k},t)\right] \\ &= i\mathcal{N}(k)\left[\frac{\boldsymbol{k}}{k^2}\times\tilde{\boldsymbol{B}}(\boldsymbol{k},t) - \frac{1}{\omega}\tilde{\boldsymbol{E}}_\perp(\boldsymbol{k},t)\right]\end{aligned} \tag{B-15}$$

where $\mathcal{N}(k)$ is a real constant, not yet specified, which can depend on k (its value will be chosen at the beginning of the next chapter). This definition, together with (A-30b), yields the equation of motion for $\boldsymbol{\alpha}(\boldsymbol{k},t)$:

$$\dot{\boldsymbol{\alpha}}(\boldsymbol{k},t) + i\omega\boldsymbol{\alpha}(\boldsymbol{k},t) = 0 \tag{B-16}$$

As opposed to $\boldsymbol{A}_\perp(\boldsymbol{k},t)$ that, according to (A-31), obeys a second order equation, this new variable $\boldsymbol{\alpha}(\boldsymbol{k},t)$ obeys a first order equation. It is a complex variable whose time evolution is proportional to $e^{-i\omega t}$, and not, as is the case for the variable $\boldsymbol{A}_\perp(\boldsymbol{k},t)$, to a linear superposition of $e^{-i\omega t}$ and $e^{+i\omega t}$. It will be useful in what follows to consider the complex conjugate of equation (B-15):

$$\begin{aligned}\boldsymbol{\alpha}^*(\boldsymbol{k},t) &= \mathcal{N}(k)\left[\tilde{\boldsymbol{A}}^*_\perp(\boldsymbol{k},t) - \frac{i}{\omega}\dot{\tilde{\boldsymbol{A}}}^*_\perp(\boldsymbol{k},t)\right] \\ &= \mathcal{N}(k)\left[\tilde{\boldsymbol{A}}_\perp(-\boldsymbol{k},t) - \frac{i}{\omega}\dot{\tilde{\boldsymbol{A}}}_\perp(-\boldsymbol{k},t)\right]\end{aligned} \tag{B-17}$$

To go from the first to the second line of (B-17), we used the fact that $\boldsymbol{A}_\perp$ is real in the real space, which leads to:

$$\tilde{\boldsymbol{A}}^*_\perp(\boldsymbol{k},t) = \tilde{\boldsymbol{A}}_\perp(-\boldsymbol{k},t) \tag{B-18}$$

A similar relation exists for $\dot{\tilde{\boldsymbol{A}}}_\perp$. The transverse variables $\boldsymbol{\alpha}(\boldsymbol{k},t)$ and $\boldsymbol{\alpha}^*(\boldsymbol{k},t)$ are called the transverse field *normal variables*. We will see in the next chapter that the quantization process will transform these variables into annihilation and creation operators of photons.

B-2-c. Equation of motion for the normal variables in the presence of sources

In the presence of sources, $\tilde{\boldsymbol{j}}_\perp$ is no longer zero. We can still define the normal variables $\boldsymbol{\alpha}(\boldsymbol{k},t)$ by relations (B-15), but we must now keep the term in $\tilde{\boldsymbol{j}}_\perp(\boldsymbol{k},t)$ on the right-hand side of equation (A-30b). The same transformation that led us from equations (A-30a) and (A-30b) to (B-16) now yields a new equation of motion in the presence of sources:

$$\dot{\boldsymbol{\alpha}}(\boldsymbol{k},t) + i\omega\boldsymbol{\alpha}(\boldsymbol{k},t) = \frac{i\mathcal{N}(k)}{\varepsilon_0\ \omega}\tilde{\boldsymbol{j}}_\perp(\boldsymbol{k},t) \tag{B-19}$$

This equation is strictly equivalent to Maxwell's equations for the transverse fields. One can see this by taking the time derivative of equations (B-22a) and (B-22b) given below, and using (B-19) to get the time-dependent evolution equations (A-30a) and (A-30b) of these fields.

Independence of the normal variables

Another interest of the normal variables is that they are independent: there is no relation between $\boldsymbol{\alpha}(\boldsymbol{k},t)$ and $\boldsymbol{\alpha}^*(-\boldsymbol{k},t)$ such as the one that exists between $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$ and $\tilde{\boldsymbol{A}}^*_\perp(-\boldsymbol{k},t)$. This is because the real and imaginary parts of $\boldsymbol{\alpha}(\boldsymbol{k},t)$ depend on two independent degrees of freedom, $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$ and its time derivative. It is easy to check, by changing the sign of $\boldsymbol{k}$ in (B-15) and by using (B-18) that:

$$\boldsymbol{\alpha}(-\boldsymbol{k},t) = \mathcal{N}(k)\left[\tilde{\boldsymbol{A}}^*_\perp(\boldsymbol{k},t) + \frac{i}{\omega}\dot{\tilde{\boldsymbol{A}}}^*_\perp(\boldsymbol{k},t)\right] \neq \boldsymbol{\alpha}^*(\boldsymbol{k},t) \tag{B-20}$$

The knowledge of the $\boldsymbol{\alpha}(\boldsymbol{k},t)$ in the entire reciprocal space does not entail the knowledge of the $\boldsymbol{\alpha}^*(\boldsymbol{k},t)$. Consequently, the integrals over $\boldsymbol{k}$ of the normal variables must be taken over the entire space, and not be limited to half the reciprocal space.

B-2-d. Expression of the physical quantities in terms of the normal variables

We are going to show that all the physical quantities can be expressed in terms of the normal variables.

α. Transverse fields in the reciprocal space

Replacing $\boldsymbol{k}$ by $-\boldsymbol{k}$, we can rewrite equation (B-17) as:

$$\boldsymbol{\alpha}^*(-\boldsymbol{k},t) = \mathcal{N}(k)\left[\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) - \frac{i}{\omega}\dot{\tilde{\boldsymbol{A}}}_\perp(\boldsymbol{k},t)\right] \tag{B-21}$$

Using (B-15) and (B-21), we can now express $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$ and $\dot{\tilde{\boldsymbol{A}}}_\perp(\boldsymbol{k},t)$ as a function of $\boldsymbol{\alpha}(\boldsymbol{k},t)$ and $\boldsymbol{\alpha}^*(-\boldsymbol{k},t)$. We get:

$$\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) = \frac{1}{2\mathcal{N}(k)}\left[\boldsymbol{\alpha}(\boldsymbol{k},t) + \boldsymbol{\alpha}^*(-\boldsymbol{k},t)\right] \tag{B-22a}$$

$$\dot{\tilde{\boldsymbol{A}}}_\perp(\boldsymbol{k},t) = -i\frac{\omega}{2\mathcal{N}(k)}\left[\boldsymbol{\alpha}(\boldsymbol{k},t) - \boldsymbol{\alpha}^*(-\boldsymbol{k},t)\right] \tag{B-22b}$$

β. Energy and momentum of the transverse field

We insert relations (B-22a) and (B-22b) for $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$ and $\dot{\tilde{A}}_\perp(\boldsymbol{k},t)$ in the expression (A-40) for H_{trans}, using the more compact notation:

$$\boldsymbol{\alpha}=\boldsymbol{\alpha}(\boldsymbol{k},t) \quad ; \quad \boldsymbol{\alpha}_-=\boldsymbol{\alpha}(-\boldsymbol{k},t) \tag{B-23}$$

We get:

$$\begin{aligned} H_{\text{trans}} &= \frac{\varepsilon_0}{2}\int \mathrm{d}^3k\ \frac{\omega^2}{4\mathcal{N}^2(k)}\left[(\boldsymbol{\alpha}^*-\boldsymbol{\alpha}_-)\cdot(\boldsymbol{\alpha}-\boldsymbol{\alpha}_-^*)+(\boldsymbol{\alpha}^*+\boldsymbol{\alpha}_-)\cdot(\boldsymbol{\alpha}+\boldsymbol{\alpha}_-^*)\right] \\ &= \frac{\varepsilon_0}{2}\int \mathrm{d}^3k\ \frac{\omega^2}{4\mathcal{N}^2(k)}\left[2\boldsymbol{\alpha}^*\cdot\boldsymbol{\alpha}+2\boldsymbol{\alpha}_-\cdot\boldsymbol{\alpha}_-^*\right] \end{aligned} \tag{B-24}$$

(in these equations, we keep the ordered sequence of $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^*$ as they appear in the computations, even though $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^*$ are commuting numbers; the reason is that similar computations can be carried out in the quantum theory where $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^*$ will be replaced by non-commuting operators). A change of variable $\boldsymbol{k}\to-\boldsymbol{k}$ in the integral of the terms in $\boldsymbol{\alpha}_-\cdot\boldsymbol{\alpha}_-^*$ yields an integral of $\boldsymbol{\alpha}\cdot\boldsymbol{\alpha}^*$. We then get:

$$H_{\text{trans}}=\varepsilon_0\int \mathrm{d}^3k\ \frac{\omega^2}{4\mathcal{N}^2(k)}\left[\boldsymbol{\alpha}^*\cdot\boldsymbol{\alpha}+\boldsymbol{\alpha}\cdot\boldsymbol{\alpha}^*\right] \tag{B-25}$$

Expliciting the components of $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^*$ on the two polarization vectors $\boldsymbol{\varepsilon}$ perpendicular to $\boldsymbol{k}$, and using the simplified notation (B-13), we finally get:

$$H_{\text{trans}}=\varepsilon_0\int \mathrm{d}^3k\ \sum_{\boldsymbol{\varepsilon}}\frac{\omega^2}{4\mathcal{N}^2(k)}\left[\alpha_{\boldsymbol{\varepsilon}}^*(\boldsymbol{k},t)\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)+\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\alpha_{\boldsymbol{\varepsilon}}^*(\boldsymbol{k},t)\right] \tag{B-26}$$

This expression looks a lot like a sum of harmonic oscillator Hamiltonians; a suitable choice for the constant $\mathcal{N}$ will be made in the next chapter.

Similar calculations can be carried out for the transverse field momentum $\boldsymbol{P}_{\text{trans}}$[11]. Using equations (A-45), (B-22a) and (B-22b), we get:

$$\boldsymbol{P}_{\text{trans}}=\varepsilon_0\int \mathrm{d}^3k\ \sum_{\boldsymbol{\varepsilon}}\frac{\omega}{4\mathcal{N}^2(k)}\ \boldsymbol{k}\ \left[\alpha_{\boldsymbol{\varepsilon}}^*(\boldsymbol{k},t)\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)+\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\alpha_{\boldsymbol{\varepsilon}}^*(\boldsymbol{k},t)\right] \tag{B-27}$$

γ. Transverse fields in real space

Let us consider first the transverse potential vector $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$, whose expression in terms of the normal variables is given by (B-22a). To get its expression in real space, one must, taking (A-11) into account, multiply (B-22a) by $(2\pi)^{-3/2}e^{i\boldsymbol{k}\cdot\boldsymbol{r}}$ and integrate over $\boldsymbol{k}$. Making the change of variable $\boldsymbol{k}\to-\boldsymbol{k}$ in the integral containing $\boldsymbol{\alpha}^*(-\boldsymbol{k},t)$, we finally get:

$$\boldsymbol{A}_\perp(\boldsymbol{r},t)=\frac{1}{(2\pi)^{3/2}}\int \mathrm{d}^3k\ \sum_{\boldsymbol{\varepsilon}}\frac{1}{2\mathcal{N}(k)}\left[\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\,\boldsymbol{\varepsilon}\ e^{i\boldsymbol{k}\cdot\boldsymbol{r}}+\alpha_{\boldsymbol{\varepsilon}}^*(\boldsymbol{k},t)\,\boldsymbol{\varepsilon}^*\ e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right] \tag{B-28}$$

[11] The expression of the angular momentum $\boldsymbol{J}_{\text{trans}}$ of the transverse field, in terms of the normal variables, will be computed in Complement $\mathrm{B}_{\mathrm{XIX}}$.

This relation (as well as the next two equations) is written in the general case where the polarizations may be complex, elliptical or circular (*cf.* note 10). This is why the term in $\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ contains a complex conjugate polarization $\boldsymbol{\varepsilon}^*$.

Similar calculations can be carried out for the transverse electric field as well as for the magnetic field. They yield:

$$\boldsymbol{E}_\perp(\boldsymbol{r},t) = \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3k \sum_{\boldsymbol{\varepsilon}} \frac{i\omega}{2\mathcal{N}(k)} \left[\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\, \boldsymbol{\varepsilon}\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - \alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\, \boldsymbol{\varepsilon}^*\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right] \tag{B-29}$$

$$\boldsymbol{B}(\boldsymbol{r},t) = \frac{1}{(2\pi)^{3/2}} \int \mathrm{d}^3k \sum_{\boldsymbol{\varepsilon}} \frac{ik}{2\mathcal{N}(k)} \left[\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\, \boldsymbol{\kappa}\times\boldsymbol{\varepsilon}\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - \alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\, \boldsymbol{\kappa}\times\boldsymbol{\varepsilon}^*\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right] \tag{B-30}$$

where $\boldsymbol{\kappa}$ has been defined in (A-16) as the unit vector parallel to $\boldsymbol{k}$.

B-3. Discrete modes in a box

So far, we have considered radiation propagating in an infinite space and used continuous Fourier transforms; in relation (A-11), the electric field is expanded on a continuous basis of normalized plane waves $e^{i\boldsymbol{k}\cdot\boldsymbol{r}}/(2\pi)^{3/2}$. It is often useful, however, to use a discrete basis, assuming the radiation to be contained in a box of finite volume, generally defined as a cube of edge length L; this will frequently occur in the next two chapters when dealing with the quantized radiation. The components of each wave vector must obey the boundary conditions in the box[12], and hence take on discrete values:

$$k_{x,y,z} = 2\pi\ n_{x,y,z}/L \tag{B-31}$$

At the end of the computation, nothing prevents us from choosing a very large value of L in order to check that the final result does not depend on L.

Instead of continuous spatial Fourier transforms, one must now introduce discrete Fourier series where each physical quantity is expanded in terms of normalized plane waves $e^{i\mathbf{k}\cdot\mathbf{r}}/L^{3/2}$. The expansion (A-11) of the electric field then becomes:

$$\boldsymbol{E}(\boldsymbol{r},t) = \frac{1}{L^{3/2}} \sum_{\mathbf{k}} \tilde{\boldsymbol{E}}_{\boldsymbol{k}}(t) e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{B-32}$$

with[13]:

$$\tilde{\boldsymbol{E}}_{\boldsymbol{k}}(t) = \frac{1}{L^{3/2}} \int_{\mathcal{V}} \mathrm{d}^3r\, \boldsymbol{E}(\boldsymbol{r},t) e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{B-33}$$

The summation in (B-32) is discrete, and the integral in (B-33) is now limited to the volume $\mathcal{V}$ of the box.

[12] One can choose to impose the field being zero on the walls, but it is generally easier to enforce periodic boundary conditions (B-31), which lead to the same k density of states.

[13] In Appendix I, we used a slightly different definition for the Fourier series, with which the factor $1/L^{3/2}$ would be missing from (B-32), but where (B-33) would contain a factor $1/L^3$. The definition we use here is chosen to directly yield an expansion of $\boldsymbol{E}(\boldsymbol{r},t)$ on plane waves normalized in the cube.

Note that if the field is zero outside the box, it is obviously possible to use the continuous Fourier transform (A-10) to get the field component $\tilde{\boldsymbol{E}}(\boldsymbol{k},t)$; however, this latter component is different from the discrete component $\tilde{\boldsymbol{E}}_{\boldsymbol{k}}(t)$, because of the coefficients introduced in the definitions. The two components are related by:

$$\tilde{\boldsymbol{E}}_{\boldsymbol{k}}(t) = \left(\frac{2\pi}{L}\right)^{3/2} \tilde{\boldsymbol{E}}(\boldsymbol{k},t) \tag{B-34}$$

The same changes can be made on the Fourier transforms of all the other physical quantities such as the magnetic field, the vector potential, as well as the charge and current densities. The equations in the reciprocal space such as equations (A-17), (A-21), (A-25) and the following, remain valid if we replace the continuous variables $\mathbf{k}$ by discrete variables, since each side of those equations are multiplied by the same factor $(2\pi)^{3/2}\, L^{-3/2}$. In the case of a zero field outside the box, the $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ are also replaced by :

$$\alpha_{\boldsymbol{k},\boldsymbol{\varepsilon}}(t) = \left(\frac{2\pi}{L}\right)^{3/2} \alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \tag{B-35}$$

Coming back to ordinary space $(\boldsymbol{r},t)$ via the inverse Fourier transform, we must use relations of the type (B-32) instead of (A-11). Consequently, once we replace in the integral over d^3k the $\tilde{\boldsymbol{E}}_(\boldsymbol{k},t)$ by the $\tilde{\boldsymbol{E}}_{\boldsymbol{k}}(t)$, we must also introduce a multiplicative factor[14]:

$$\int \mathrm{d}^3k \implies \left(\frac{2\pi}{L}\right)^{3/2} \sum_{\mathbf{k}} \tag{B-36}$$

B-4. Generalization of the mode concept

In the absence of sources, the solution of the equation of motion (B-16) for the normal variable $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ is very simple, since it is an exponential with an angular frequency $\omega = ck$:

$$\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t) = \alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},0)e^{-i\omega t} \tag{B-37}$$

Inserting (B-37) in the expressions we just obtained for the transverse fields and the other physical quantities, we see that the fields are linear superpositions of progressive plane waves, propagating independently of each other. The free field energy and momentum are the sum of the squared moduli of the various normal variables, each being time-independent and proportional to $|\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},0)|^2$.

The modes $\{\boldsymbol{k},\boldsymbol{\varepsilon}\}$ introduced in this chapter permit expanding the free transverse fields on progressive plane waves. Nevertheless, other expansions on monochromatic waves that are not necessarily plane waves are also possible; they involve other families of modes, as we are now show, coming back to equation (A-31). In the absence of sources, any monochromatic solution of this equation, of the form $\boldsymbol{A}_{\perp}^{(+)}(\boldsymbol{r})e^{-i\omega t}$, necessarily obeys equation:

$$(\Delta + k^2)\boldsymbol{A}_{\perp}^{(+)}(\boldsymbol{r}) = 0 \tag{B-38}$$

[14]The product of the multiplicative factor of (B-34) and that of (B-35) yields the usual factor $(2\pi/L)^3$, obtained directly from (B-31).

(which is simply the Helmholtz equation) with $k = \omega/c$. The plane waves $e^{\pm i\boldsymbol{k}\cdot\boldsymbol{r}}$ are a possible basis of eigenfunctions for this eigenvalue equation, but not the only one. There exists other bases, such as the basis of stationary waves $\cos \boldsymbol{k}\cdot\boldsymbol{r}$ and $\sin \boldsymbol{k}\cdot\boldsymbol{r}$, the basis of multipolar waves (radiation modes with a specific angular momentum, whereas plane waves have a specific linear momentum), or the basis corresponding to Gaussian modes. More generally, any linear combination of plane waves with the same modulus $\boldsymbol{k}$ can become a mode. Whatever basis is chosen, the transverse field energy will be a sum of the squared moduli of normal variables introduced in the expansions of the transverse fields on the eigenfunctions of that basis. The expression of the other physical quantities, however, will only have a simple form in a particular basis. As an example, the momentum of the transverse field is a sum of squared moduli only in the basis of progressive plane waves, whereas the field angular momentum has a simple form only in the basis of multipolar waves.

Note finally that the field can be contained in a cavity with well defined boundary conditions. Finding the eigenfunctions of equation (B-38) obeying these boundary conditions is a way to determine the eigenmodes of this cavity.

To conclude this chapter, we can say that the free radiation field is equivalent to an ensemble of one-dimensional harmonic oscillators associated with the modes $\{\boldsymbol{k}, \boldsymbol{\varepsilon}\}$ labeled by their wave vector and their transverse polarization. Each mode is associated with a field normal variable, similar to the classical variable of the corresponding classical oscillator, and which will become, in the quantization process, the oscillator annihilation operator. The results established in this chapter will be the simple starting point for the radiation quantization explained in the next chapter.

COMPLEMENT OF CHAPTER XVIII, READER'S GUIDE

A_{XVIII} : LAGRANGIAN FORMULATION OF ELECTRODYNAMICS

The dynamic equations for the electrodynamic field (Maxwell's equations) can be obtained from the Lagrangian formalism based on a principle of least action. This enables introducing expressions for the conjugate momenta of the various field variables, as well as for the field Hamiltonian when coupled to charged particles. The results of this complement are not indispensable for reading the other chapters and complements. They offer, however, an overview of a more general approach to quantum electrodynamics, which is essential for a relativistic treatment of these problems and for the use of path integrals (Appendix IV).

Complement A_{XVIII}

Lagrangian formulation of electrodynamics

1 Lagrangian with several types of variables 1980
1-a Lagrangian formalism with discrete and real variables 1980
1-b Extension to complex variables 1982
1-c Lagrangian with continuous variables 1984
2 Application to the free radiation field 1986
2-a Lagrangian densities in real and reciprocal spaces 1986
2-b Lagrange's equations 1987
2-c Conjugate momentum of the transverse potential vector . . . 1987
2-d Hamiltonian; Hamilton-Jacobi equations 1988
2-e Field commutation relations 1989
2-f Creation and annihilation operators 1990
2-g Discrete momentum variables 1991
3 Lagrangian of the global system field + interacting particles 1992
3-a Choice for the Lagrangian 1992
3-b Lagrange's equations 1993
3-c Conjugate momenta 1995
3-d Hamiltonian 1995
3-e Commutation relations 1996

Introduction

As shown in Appendix III, the dynamics of a system of point particles in an external potential can be described either by Newton's equations, or by a Lagrangian with the principle of least action leading to Lagrange's equations, equivalent to Newton's equations. An advantage of the Lagrangian formalism is that it facilitates the quantization of the theory: it directly leads to the definition of the conjugate momenta of the particles' coordinates, and of the system's Hamiltonian, which is a function of the coordinates and the conjugate momenta. It then naturally introduces the canonical commutation relations, fundamental for the quantum description of the system. This complement will show, in a succinct way, how the Maxwell-Lorentz equations, studied in this chapter and the next, can be deduced from a Lagrangian and a principle of least action. This will give a more general justification for the expression of the Hamiltonian of the system "field + particles" postulated in Chapter XIX and for the commutation relations also postulated in that chapter[1]. Another advantage of the Lagrangian formalism, that we shall not exploit here, is that it is well suited to a relativistic description of the system "field + particles" which is why it is used in the quantum theory of relativistic fields.

[1]The relations postulated in Chapter XIX are justified a posteriori by the fact that they lead to the correct Heisenberg equations for the quantum operators associated with the particles and the fields.

We start in § 1 by extending the computations of Appendix III to the case where the system coordinates are complex and not real, even though the Lagrangian remains a real quantity. We will also show that the principle of least action and Lagrange's equations can be generalized to the field case, that is to a case where the system coordinates no longer depend on a discrete but on a continuous index, such as the point $\boldsymbol{r}$ in real space.

The discussion will be illustrated in § 2, which studies the Lagrangian of the free radiation field in the absence of sources. The field will be described by its components in reciprocal space, which are complex quantities. The Lagrangian then depends only on the field components and their time derivatives, hence making the computations easier than if the fields were described by the real components in real space (this is because the Lagrangian in real space depends not only on the field components and their time derivatives, but also on their spatial derivatives). In this study, we shall establish the expression for the field Hamiltonian, and the canonical commutation relations of the components of that field.

Finally, we give in § 3 the expression for the electrodynamic Lagrangian in the Coulomb gauge in the presence of sources; we show how Lagrange's equations deduced from this Lagrangian coincide with the Maxwell-Lorentz equations studied in Chapter XVIII. Several important relations for the quantization of the theory will be established: expression for the conjugate momenta of the particles and fields; expression for the Hamiltonian of the global system field + particles; canonical commutation relations. The results obtained in this complement give a base for the quantization process more general than the simplified approach of Chapitre XIX. The interested reader can find a more detailed description of the electrodynamic Lagrangian and Hamiltonian formalism in Chapter II of reference [16] and its complements.

1. Lagrangian with several types of variables

1-a. Lagrangian formalism with discrete and real variables

The Lagrangian L is a real function of dynamical variables composed of "generalized coordinates" $x_i(t)$ labeled by a discrete index i and of the corresponding "generalized velocities" $\dot{x}_i(t) = \mathrm{d}x_i(t)/\mathrm{d}t$. L is written:

$$L\left[x_1(t), x_2(t), ...x_N(t); \dot{x}_1(t), \dot{x}_2(t), ...\dot{x}_N(t)\right] \tag{1}$$

Consider a possible motion of the system where the coordinates $x_i(t)$ follow a certain "path" Γ between an initial time t_{in} and a final time t_2. The integral of L along the path Γ is, by definition, *the action* S_Γ associated with this path:

$$S_\Gamma = \int_{t_1}^{t_2} \mathrm{d}t \; L\left[x_1(t), x_2(t), ...x_N(t); \dot{x}_1(t), \dot{x}_2(t), ...\dot{x}_N(t)\right] \tag{2}$$

The *principle of least action* postulates that, among all the possible paths starting from the same initial conditions described by $x_i(t_{\text{in}})$ and arriving at the same final conditions described by $x_i(t_2)$, the system will follow the one for which S_Γ presents an extremum (if the path varies, S_Γ is stationary). Consider an infinitesimal variation $\delta x_i(t)$ and $\delta \dot{x}_i(t)$ of the dynamical variables around this path of extremum action, which does not change the initial and final values of the coordinates, i.e. such that:

$$\delta x_i(t_{\text{in}}) = \delta x_i(t_2) = 0 \tag{3}$$

The corresponding variation of the action

$$\delta S = \int_{t_1}^{t_2} \sum_i \left[\delta x_i(t) \frac{\partial L}{\partial x_i} + \delta \dot{x}_i(t) \frac{\partial L}{\partial \dot{x}_i} \right] \mathrm{d}t \tag{4}$$

must be zero to first order in $\delta x_i(t)$ and $\delta \dot{x}_i(t)$. We replace, in the last term of (4), $\delta \dot{x}_i(t)$ by:

$$\delta \dot{x}_i(t) = \frac{\mathrm{d}}{\mathrm{d}t} \delta x_i(t) \tag{5}$$

and integrate by parts the corresponding term. The integrated part is zero because of (3). We then get:

$$\delta S = \int_{t_1}^{t_2} \sum_i \delta x_i(t) \left[\frac{\partial L}{\partial x_i} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{x}_i} \right] \mathrm{d}t \tag{6}$$

As δS must be zero for any variation $\delta x_i(t)$, the path actually followed by the system must obey the N equations:

$$\frac{\partial L}{\partial x_i} - \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L}{\partial \dot{x}_i} = 0 \tag{7}$$

These relations are called *Lagrange's equations*; they can be shown to be equivalent to Newton's equations (Appendix III).

The next step of the Lagrangian formalism is to introduce the *conjugate momenta* p_i of the coordinates x_i, defined by the equations:

$$p_i = \frac{\partial L}{\partial \dot{x}_i} \tag{8}$$

as well as the *Hamiltonian* H equal to:

$$H = \sum_i \dot{x}_i \, p_i - L \tag{9}$$

Let us take the differential of H

$$\begin{aligned} \mathrm{d}H &= \sum_i \left[p_i \mathrm{d}\dot{x}_1 + \dot{x}_i \mathrm{d}p_i - \frac{\partial L}{\partial x_i} \mathrm{d}x_i - \frac{\partial L}{\partial \dot{x}_1} \mathrm{d}\dot{x}_i \right] \\ &= \sum_i \left[\dot{x}_i \mathrm{d}p_i - \dot{p}_i \mathrm{d}x_i \right] \end{aligned} \tag{10}$$

To go from the first to the second line of (10), we used (7) and (8) to replace $\partial L/\partial x_i$ by $\dot{p}_i$ and $\partial L/\partial \dot{x}_i$ by p_i. We assume the $\dot{x}_i$ can be expressed as a function of the x_i and the p_i. The Hamiltonian H is then a function of the coordinates x_i and the conjugate momenta p_i, whose evolution obeys, taking (10) into account, the $2N$ equations:

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \frac{\partial H}{\partial p_i} \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H}{\partial x_i} \tag{11}$$

called the *Hamilton-Jacobi equations*.

Let us finally recall the canonical quantization process. One associates with the coordinates x_i and the conjugate momenta p_i the operators $\hat{x}_i$ and $\hat{p}_i$ obeying the commutation relations:

$$[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij} \tag{12}$$

all the other commutators being equal to zero. These results are valid only if the coordinates x_i are Cartesian components (see comment in § B-5 of Chapter III).

1-b. Extension to complex variables

Let us assume, to keep things simple, that the index i takes only $N = 2$ values. With the two real coordinates $x_1(t)$ and $x_2(t)$ we build the complex variables:

$$X(t) = \frac{1}{\sqrt{2}}[x_1(t) + ix_2(t)] \qquad X^*(t) = \frac{1}{\sqrt{2}}[x_1(t) - ix_2(t)] \tag{13}$$

whose real and imaginary parts are, within a factor $1/\sqrt{2}$, equal to $x_1(t)$ and $x_2(t)$ for $X(t)$, $x_1(t)$ and $-x_2(t)$ for $X^*(t)$. Equations (13) can be inverted and yield:

$$x_1(t) = \frac{1}{\sqrt{2}}[X(t) + X^*(t)] \qquad x_2(t) = -\frac{i}{\sqrt{2}}[X(t) - X^*(t)] \tag{14}$$

Analogous equations can be written, relating $\dot{X}(t)$ and $\dot{X}^*(t)$ to $\dot{x}_1(t)$ and $\dot{x}_2(t)$ and vice versa:

$$\dot{X}(t) = \frac{1}{\sqrt{2}}[\dot{x}_1(t) + i\dot{x}_2(t)] \qquad \dot{X}^*(t) = \frac{1}{\sqrt{2}}[\dot{x}_1(t) - i\dot{x}_2(t)] \tag{15}$$

$$\dot{x}_1(t) = \frac{1}{\sqrt{2}}\left[\dot{X}(t) + \dot{X}^*(t)\right] \qquad \dot{x}_2(t) = -\frac{i}{\sqrt{2}}\left[\dot{X}(t) - \dot{X}^*(t)\right] \tag{16}$$

Inserting in the Lagrangian (1) expressions (14) and (16) for the variables, we get a Lagrangian of the form $L\left[X(t), X^*(t), \dot{X}(t), \dot{X}^*(t)\right]$ that depends on complex variables. Note however that just as in (1), this Lagrangian, even though it depends on complex variables, is still a real quantity since its time integral along a path Γ is an action (which is a real quantity). We now study what becomes of all the results established earlier using (1) when they are expressed as a function of $X, X^*, \dot{X}$ and $\dot{X}^*$.

α. *Lagrange's equations*

It is important, for what follows, to relate $\partial L/\partial X$ and $\partial L/\partial \dot{X}$ to $\partial L/\partial x_1$, $\partial L/\partial x_2$, $\partial L/\partial \dot{x}_1$ and $\partial L/\partial \dot{x}_2$. Using (14) and (16), we can write:

$$\frac{\partial L}{\partial X} = \frac{\partial L}{\partial x_1}\frac{\partial x_1}{\partial X} + \frac{\partial L}{\partial x_2}\frac{\partial x_2}{\partial X} = \frac{1}{\sqrt{2}}\left[\frac{\partial L}{\partial x_1} - i\frac{\partial L}{\partial x_2}\right] \tag{17}$$

$$\frac{\partial L}{\partial \dot{X}} = \frac{\partial L}{\partial \dot{x}_1}\frac{\partial \dot{x}_1}{\partial \dot{X}} + \frac{\partial L}{\partial \dot{x}_2}\frac{\partial \dot{x}_2}{\partial \dot{X}} = \frac{1}{\sqrt{2}}\left[\frac{\partial L}{\partial \dot{x}_1} - i\frac{\partial L}{\partial \dot{x}_2}\right] \tag{18}$$

Subtracting from equation (17) the time derivative of equation (18), we get:

$$\frac{\partial L}{\partial X} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{X}} = \frac{1}{\sqrt{2}}\left(\frac{\partial L}{\partial x_1} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}_1}\right) - \frac{i}{\sqrt{2}}\left(\frac{\partial L}{\partial x_2} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}_2}\right) \tag{19}$$

The two parentheses on the right-hand side of this equation are zero since x_1 and x_2 obey Lagrange's equation (7). It then follows that the left-hand side is also zero, as is its complex conjugate[2] :

$$\frac{\partial L}{\partial X} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{X}} = 0 \qquad \frac{\partial L}{\partial X^*} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{X}^*} = 0 \tag{20}$$

which proves that X and X^* also obey Lagrange's equations[3].

β. Conjugate momenta

In (18) we replace $\partial L/\partial \dot{x}_1$ and $\partial L/\partial \dot{x}_2$ by p_1 and p_2 (see Eq. (8)). We get:

$$\frac{\partial L}{\partial \dot{X}} = \frac{1}{\sqrt{2}}(p_1 - ip_2) \tag{21}$$

The complex conjugate of equation (21) is written:

$$\frac{\partial L}{\partial \dot{X}^*} = \frac{1}{\sqrt{2}}(p_1 + ip_2) \tag{22}$$

To choose the definition of the conjugate momentum P of the complex variable X, it is useful to compare the way the conjugate momentum P and the velocity $\dot{X}$ are transformed upon the change of dynamical variables $x_1, x_2 \longrightarrow X, X^*$. We compare the first equation (15) and the two equations (21) and (22). The velocity $\dot{X}$ becomes $\dot{x}_1 + i\dot{x}_2$; a wise choice would be to define the associate momentum P in such a way that its transformation yields $p_1 + ip_2$. Equations (21) and (22) then clearly show that P must not be defined as $\partial L/\partial \dot{X}$, but rather as $\partial L/\partial \dot{X}^*$; the complex conjugate P^* is then equal to $\partial L/\partial \dot{X}$:

$$P = \frac{\partial L}{\partial \dot{X}^*} \qquad P^* = \frac{\partial L}{\partial \dot{X}} \tag{23}$$

This is the definition we shall use in the rest of this complement.

γ. Hamiltonian

The quantity $\dot{x}_1 p_1 + \dot{x}_2 p_2$ appears in the definition (9) of the Hamiltonian. This quantity can be rewritten by replacing $\dot{x}_1$ and $\dot{x}_2$ by their expressions (16) as a function of $\dot{X}$ and $\dot{X}^*$, as well as p_1 and p_2 by analogous expressions as a function of P and P^*:

$$\begin{aligned} \dot{x}_1 p_1 + \dot{x}_2 p_2 &= \frac{1}{\sqrt{2}}(\dot{X} + \dot{X}^*)\frac{1}{\sqrt{2}}(P + P^*) + \frac{-i}{\sqrt{2}}(\dot{X} - \dot{X}^*)\frac{-i}{\sqrt{2}}(P - P^*) \\ &= \dot{X}P^* + \dot{X}^*P \end{aligned} \tag{24}$$

[2]Since L is real, we have $(\partial L/\partial X)^* = \partial L/\partial X^*$ and an analogous equation for $\left(\partial L/\partial \dot{X}\right)^*$.

[3]These results could have been obtained directly by the variational calculation leading to (6), considering δX and δX^* to be independent variables.

We can then write:

$$H = \dot{x}_1 p_1 + \dot{x}_2 p_2 - L = \dot{X} P^* + \dot{X}^* P - L \tag{25}$$

We now take the differential of H:

$$\mathrm{d}H = \dot{X}\mathrm{d}P^* + P^*\mathrm{d}\dot{X} + \dot{X}^*\mathrm{d}P + P\mathrm{d}\dot{X}^* - \frac{\partial L}{\partial X}\mathrm{d}X - \frac{\partial L}{\partial X^*}\mathrm{d}X^* - \frac{\partial L}{\partial \dot{X}}\mathrm{d}\dot{X} - \frac{\partial L}{\partial \dot{X}^*}\mathrm{d}\dot{X}^* \tag{26}$$

Using (20) and (23), we get:

$$\mathrm{d}H = \dot{X}^*\mathrm{d}P - \dot{P}^*\mathrm{d}X + \dot{X}\mathrm{d}P^* - \dot{P}\mathrm{d}X^* \tag{27}$$

If $\dot{X}$ and $\dot{X}^*$ can be expressed in terms of the variables X, X^*, P, P^*, the Hamiltonian H only depends on those variables and we deduce from (27) the Hamilton-Jacobi equations:

$$\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{\partial H}{\partial P^*} \qquad\qquad \frac{\mathrm{d}P}{\mathrm{d}t} = -\frac{\partial H}{\partial X^*} \tag{28}$$

and the complex conjugate equations for X^* and P^*. Note that it is the partial derivative with respect to P^* (and not with respect to P) that is equal to the total derivative of X.

δ. Canonical commutation relations

Upon quantization, the different variables become operators $\hat{x}_1$, $\hat{x}_2$, $\hat{p}_1$, $\hat{p}_2$, $\hat{X}$, $\hat{P}$, $\hat{X}^\dagger$, $\hat{P}^\dagger$. The commutation relations between the operators $\hat{X}$, $\hat{P}$, $\hat{X}^\dagger$, $\hat{P}^\dagger$ are obtained by expressing those operators in terms of $\hat{x}_1$, $\hat{x}_2$, $\hat{p}_1$, $\hat{p}_2$ and using the commutation relations (12). We can easily check that the only non-zero commutators are $[\hat{X}, \hat{P}^\dagger]$ and $[\hat{P}, \hat{X}^\dagger] = -[\hat{X}, \hat{P}^\dagger]$. Using (21), (22) and (23), we obtain:

$$\begin{aligned}[\hat{X}, \hat{P}^\dagger] &= \frac{1}{2}\left[\hat{x}_1 + i\hat{x}_2, \hat{p}_1 - i\hat{p}_2\right] \\ &= \frac{1}{2}\left([\hat{x}_1, \hat{p}_1] + [\hat{x}_2, \hat{p}_2]\right) = i\hbar\end{aligned} \tag{29}$$

1-c. Lagrangian with continuous variables

We now assume that the dynamical variables of the system depend on a continuous index, such as the point $\boldsymbol{r}$ in real space (or the point $\boldsymbol{k}$ in the reciprocal space when the real space is infinite). In other words, they constitute a field $\phi_j(\boldsymbol{r})$, where the discrete index j labels the component of the field if we are dealing with a vector field; in the reciprocal space, this field becomes $\tilde{\phi}_j(\boldsymbol{k})$. We shall only establish here Lagrange's equations for a real field. In the upcoming § 2, we shall study the radiation field described by its complex components in the reciprocal space. We will then generalize to a complex field all the results established earlier for discrete and complex variables. This will yield the expression for the free field Hamiltonian and the commutation relations for the free field, which are essential for the quantum description of the field.

The Lagrangian L of a real field is now the integral in real space of a *Lagrangian density* $\mathscr{L}$:

$$L = \int \mathrm{d}^3 r\, \mathscr{L}(\boldsymbol{r}, t) \tag{30}$$

The Lagrangian density is a function of the field $\phi_j(\boldsymbol{r})$ and of its partial derivatives with respect to t and to the components of $\boldsymbol{r}$:

$$\mathscr{L}(\boldsymbol{r},t) = \mathscr{L}\left(\phi_j(\boldsymbol{r},t), \dot{\phi}_j(\boldsymbol{r},t), \partial_i\phi_j(\boldsymbol{r},t)\right) \tag{31}$$

with the notation ($i = x, y, z$):

$$\dot{\phi}_j(\boldsymbol{r},t) = \partial\phi_j(\boldsymbol{r},t)/\partial t \tag{32}$$

$$\partial_i\phi_j(\boldsymbol{r},t) = \partial\phi_j(\boldsymbol{r},t)/\partial r_i \tag{33}$$

Consider a possible path Γ for the field, going from the value $\phi_j(\boldsymbol{r}, t_{\text{in}})$ at an initial time t_{in} to the final value $\phi_j(\boldsymbol{r}, t_2)$ at a final time t_2. The action S_Γ associated with this path is, by definition:

$$S_\Gamma = \int_{t_1}^{t_2} \mathrm{d}t \int \mathrm{d}^3r\, \mathscr{L}\left(\phi_j(\boldsymbol{r},t), \dot{\phi}_j(\boldsymbol{r},t), \partial_i\phi_j(\boldsymbol{r},t)\right) \tag{34}$$

The principle of least action postulates that among all the possible paths starting from the same initial state and ending at the same final state[4], the path(s) actually followed by the system is the one (or are those) for which S_Γ presents an extremum. Let us compute the variation δS of the action for an infinitesimal variation of the path, characterized by the infinitesimal variations $\delta\phi_j(\boldsymbol{r},t)$, $\delta\left(\partial\phi_j(\boldsymbol{r},t)/\partial t\right)$ and $\delta\left(\partial\phi_j(\boldsymbol{r},t)/\partial r_i\right)$.

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \int \mathrm{d}^3r \sum_j \left[\delta\phi_j(\boldsymbol{r},t)\frac{\partial\mathscr{L}}{\partial\phi_j} + \delta\dot{\phi}_j(\boldsymbol{r},t)\frac{\partial\mathscr{L}}{\partial\dot{\phi}_j} + \sum_i \delta\left(\partial_i\phi_j(\boldsymbol{r},t)\right)\frac{\partial\mathscr{L}}{\partial(\partial_i\phi_j)}\right] \tag{35}$$

Using:

$$\begin{aligned} \delta\dot{\phi}_j(\boldsymbol{r},t) &= \partial\left(\delta\phi_j(\boldsymbol{r},t)\right)/\partial t \\ \delta\left(\partial_i\phi_j(\boldsymbol{r},t)\right) &= \partial_i\left(\delta\phi_j(\boldsymbol{r},t)\right) \end{aligned} \tag{36}$$

and performing an integration by parts of the terms proportional to $\partial\left(\delta\phi_j(\boldsymbol{r},t)\right)/\partial t$ and $\partial_i\left(\delta\phi_j(\boldsymbol{r},t)\right)$, we find that the integrated terms are zero because of the boundary conditions for $\delta\phi_j(\boldsymbol{r},t)$ at the initial and final times, and for $\boldsymbol{r} \to \infty$. The remaining terms are therefore all proportional to $\delta\phi_j(\boldsymbol{r},t)$. Grouping them all, we find[5]:

$$\delta S = \int_{t_1}^{t_2} \mathrm{d}t \int \mathrm{d}^3r \sum_{ij} \delta\phi_j(\boldsymbol{r},t)\left[\frac{\partial\mathscr{L}}{\partial\phi_j} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathscr{L}}{\partial\dot{\phi}_j} - \sum_i \partial_i\frac{\partial\mathscr{L}}{\partial(\partial_i\phi_j)}\right] \tag{37}$$

As δS must be zero for any time or spatial variations of $\delta\phi_j(\boldsymbol{r},t)$, we can deduce that:

$$\frac{\partial\mathscr{L}}{\partial\phi_j} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathscr{L}}{\partial\dot{\phi}_j} - \sum_i \partial_i\frac{\partial\mathscr{L}}{\partial(\partial_i\phi_j)} = 0 \tag{38}$$

which are the Lagrange equations for the field.

[4] We also assume that the Lagrangian density is zero or tends to zero fast enough when $\boldsymbol{r}$ goes to infinity.

[5] The function $\frac{\partial\mathscr{L}}{\partial\dot{\phi}_j}$ does not directly depend on time t. It nevertheless depends indirectly on t if we replace, as in (31), the fields and their derivatives by their values for a given history of the field. By convention, we then denote by $\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathscr{L}}{\partial\dot{\phi}_j}$ the time derivative of this function at each point of space. It contains the sum of the contributions of the partial derivatives of the function with respect to all the initial variables (the fields and their derivatives).

2. Application to the free radiation field

We now study the free radiation field (in the absence of sources) starting from its Lagrangian density in reciprocal space. We choose the Coulomb gauge so that the longitudinal vector potential $\boldsymbol{A}_{\parallel}$ is zero and $\boldsymbol{A}$ is reduced to $\boldsymbol{A}_{\perp}$; furthermore, the scalar potential is also zero since, in the Coulomb gauge, it would be the potential corresponding to the Coulomb field created by the charges (see relation (A-34) of Chapter XVIII), and we are assuming that there are no charges. The only fields we have to consider are thus the transverse electric field and the magnetic field related to the transverse potential vector by the following equations in real space:

$$\boldsymbol{E}_{\perp}(\boldsymbol{r},t) = -\dot{\boldsymbol{A}}_{\perp}(\boldsymbol{r},t) \qquad \boldsymbol{B}(\boldsymbol{r},t) = \boldsymbol{\nabla} \times \boldsymbol{A}_{\perp}(\boldsymbol{r},t) \tag{39}$$

and in the reciprocal space:

$$\tilde{\boldsymbol{E}}_{\perp}(\boldsymbol{k},t) = -\dot{\tilde{\boldsymbol{A}}}_{\perp}(\boldsymbol{k},t) \qquad \tilde{\boldsymbol{B}}(\boldsymbol{k},t) = i\boldsymbol{k} \times \tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t) \tag{40}$$

2-a. Lagrangian densities in real and reciprocal spaces

The Lagrangian density most commonly used in real space is[6]:

$$\mathscr{L}(\boldsymbol{r},t) = \frac{\varepsilon_0}{2}\left[\boldsymbol{E}_{\perp}^2(\boldsymbol{r},t) - c^2\boldsymbol{B}^2(\boldsymbol{r},t)\right] \tag{41}$$

where c is the speed of light. Using (39), we see that this Lagrangian density depends both on $\boldsymbol{A}_{\perp}(\boldsymbol{r},t)$ and $\dot{\boldsymbol{A}}_{\perp}(\boldsymbol{r},t)$, as well as on the spatial derivatives of $\boldsymbol{A}_{\perp}(\boldsymbol{r},t)$.

We now go to the reciprocal space. The Lagrangian L is then written as:

$$L = \int \mathrm{d}^3k\, \tilde{\mathscr{L}}(\boldsymbol{k},t) \tag{42}$$

where the Lagrangian density $\tilde{\mathscr{L}}(\boldsymbol{k},t)$ in the reciprocal space is obtained from (41), rewriting the fields in the reciprocal space[7]. Let us evaluate the contribution to the Lagrangian of the two terms in the bracket of (41). Using (40) and the Parseval-Plancherel equality – see relation (A-12) of Chapter XVIII – we can write:

$$\begin{aligned}\int \mathrm{d}^3r\, \boldsymbol{E}_{\perp}(\boldsymbol{r},t)\cdot\boldsymbol{E}_{\perp}(\boldsymbol{r},t) &= \int \mathrm{d}^3k\, \dot{\tilde{\boldsymbol{A}}}_{\perp}^{\,*}(\boldsymbol{k},t)\cdot\dot{\tilde{\boldsymbol{A}}}_{\perp}(\boldsymbol{k},t)\\ \int \mathrm{d}^3r\, \boldsymbol{B}(\boldsymbol{r},t)\cdot\boldsymbol{B}(\boldsymbol{r},t) &= \int \mathrm{d}^3k\, \left[(-i\boldsymbol{k}\times\tilde{\boldsymbol{A}}_{\perp}^{*}(\boldsymbol{k},t))\cdot(i\boldsymbol{k}\times\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t))\right]\end{aligned} \tag{43}$$

As the two vectors $\boldsymbol{k}\times\tilde{\boldsymbol{A}}_{\perp}^{*}(\boldsymbol{k},t)$ and $\boldsymbol{k}\times\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)$ are in the plane perpendicular to $\boldsymbol{k}$, we have:

$$\int \mathrm{d}^3k\, \left[(-i\boldsymbol{k}\times\tilde{\boldsymbol{A}}_{\perp}^{*}(\boldsymbol{k},t)\cdot(i\boldsymbol{k}\times\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)\right] = \int \mathrm{d}^3k\, k^2\, \tilde{\boldsymbol{A}}_{\perp}^{*}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t) \tag{44}$$

Finally, the Lagrangian density in the reciprocal space is written as:

$$\tilde{\mathscr{L}}(\boldsymbol{k},t) = \frac{\varepsilon_0}{2}\left[\dot{\tilde{\boldsymbol{A}}}_{\perp}^{*}(\boldsymbol{k},t)\cdot\dot{\tilde{\boldsymbol{A}}}_{\perp}(\boldsymbol{k},t) - c^2k^2\tilde{\boldsymbol{A}}_{\perp}^{*}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_{\perp}(\boldsymbol{k},t)\right] \tag{45}$$

[6] This density has the advantage of being a relativistic invariant (Lorentz scalar).

[7] We use the notation $\tilde{\mathscr{L}}(\boldsymbol{k},t)$ for this Lagrangian density even though this function is not the Fourier transform of $\mathscr{L}(\boldsymbol{r},t)$.

The generalized coordinates of the field can be seen as components of the transverse potential vector, and the generalized velocities as the time derivatives of these coordinates. As opposed to $\mathscr{L}(\boldsymbol{r},t)$, $\tilde{\mathscr{L}}(\boldsymbol{k},t)$ depends only on $\tilde{\boldsymbol{A}}^*_\perp(\boldsymbol{k},t)$ and $\dot{\tilde{\boldsymbol{A}}}_\perp(\boldsymbol{k},t)$, and not on the partial derivatives of $\tilde{\boldsymbol{A}}^*_\perp(\boldsymbol{k},t)$ with respect to the $\boldsymbol{k}$ components. The computations will thus be simpler in the reciprocal space.

It will be useful for what follows to introduce the Cartesian components of $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$ in the reference frame formed by $\boldsymbol{\kappa}=\boldsymbol{k}/k$ and the two polarization unit vectors $\boldsymbol{\varepsilon}_1(\boldsymbol{k})$ and $\boldsymbol{\varepsilon}_2(\boldsymbol{k})$ in the plane perpendicular to $\boldsymbol{k}$. Since $\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)$ is transverse, it does not have any component along $\boldsymbol{\kappa}$, and we can write:

$$\tilde{\mathscr{L}}(\boldsymbol{k},t)=\frac{\varepsilon_0}{2}\sum_{\boldsymbol{\varepsilon}}\left[\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)-c^2k^2\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\right] \tag{46}$$

where $\sum_{\boldsymbol{\varepsilon}}$ is a simplified notation representing the summation over the two transverse polarizations $\boldsymbol{\varepsilon}_1(\boldsymbol{k})$ and $\boldsymbol{\varepsilon}_2(\boldsymbol{k})$ – see relation (B-13) of Chapter XVIII.

2-b. Lagrange's equations

Equation (38) becomes here:

$$\frac{\partial\tilde{\mathscr{L}}(\boldsymbol{k},t)}{\partial\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)}-\frac{\mathrm{d}}{\mathrm{d}t}\,\frac{\partial\tilde{\mathscr{L}}(\boldsymbol{k},t)}{\partial\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)}=0 \tag{47}$$

We then get, taking (46) into account:

$$\ddot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)+c^2k^2\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)=0 \tag{48}$$

This equation coincides with equations (A-30a) and (A-30b) of Chapter XVIII, which give the time evolution of the transverse vector potential of a free field in the absence of sources (we set $\tilde{\boldsymbol{j}}_\perp=\boldsymbol{0}$). We recover, as expected, the predictions of Maxwell's equations in the usual formulation of classical electrodynamics.

2-c. Conjugate momentum of the transverse potential vector

To define the conjugate momentum $\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ of the complex variable $\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$, we use expression (23). Note however that the velocity $\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ appears several times in the integral over $\boldsymbol{k}$ of $\tilde{\mathscr{L}}(\boldsymbol{k},t)$. Consequently, we must add all the corresponding contributions of the partial derivatives of $\tilde{\mathscr{L}}(\boldsymbol{k},t)$ with respect to these various velocities in the definition of the conjugate momentum $\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$. This situation results from the fact that the fields are real in real space. The Fourier transform properties then lead to (*cf.* relation (B-18) of Chapter XVIII):

$$\boldsymbol{A}_\perp(\boldsymbol{r},t)=\boldsymbol{A}^*_\perp(\boldsymbol{r},t)\qquad\Rightarrow\qquad\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)=\tilde{\boldsymbol{A}}^*_\perp(-\boldsymbol{k},t) \tag{49}$$

and to an equivalent relation for the time derivatives of the components of the transverse potential vector. In the integral over $\boldsymbol{k}$ of $\tilde{\mathcal{L}}(\boldsymbol{k},t)$ we get, in addition to the term $\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$, the term $\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(-\boldsymbol{k},t)\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(-\boldsymbol{k},t)$ which, according to (49), is equal to

$\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)^*$ and therefore doubles the first term. If we ignore the terms in $-\boldsymbol{k}$, we must then double the contribution of the terms in $\boldsymbol{k}$, which yields:

$$\begin{aligned}\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) &= 2\frac{\partial\tilde{\mathscr{L}}(\boldsymbol{k},t)}{\partial\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)} \\ &= \varepsilon_0\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) = -\varepsilon_0\tilde{E}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\end{aligned} \tag{50}$$

The conjugate momentum of the transverse potential is seen to be equal, within a factor $-\varepsilon_0$, to the transverse electric field.

Another equivalent way to obtain (50) is to use, in the Lagrangian expression, only independent variables. The reality condition (49) ensures that, if one knows the variables in half the reciprocal space, one knows them in the entire space. One can define L as the integral over only half the reciprocal space (where all the variables are independent) of a Lagrangian density, noted $\bar{\mathscr{L}}(\boldsymbol{k},t)$, equal to twice the initial density. Writing "$\fint$ " the integral over half a space (the bar indicating that the $\boldsymbol{k}$ space is divided into two parts), we get:

$$L = \fint \mathrm{d}^3k\ \bar{\mathscr{L}}(\boldsymbol{k},t) \tag{51}$$

with:

$$\bar{\mathscr{L}}(\boldsymbol{k},t) = \varepsilon_0\sum_{\boldsymbol{\varepsilon}}\left[\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) - \omega^2\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\right] = 2\ \tilde{\mathscr{L}}(\boldsymbol{k},t) \tag{52}$$

so that one can also define the conjugate momentum of the transverse potential vector as:

$$\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) = \frac{\partial\bar{\mathscr{L}}(\boldsymbol{k},t)}{\partial\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)} = \varepsilon_0\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \tag{53}$$

2-d. Hamiltonian; Hamilton-Jacobi equations

The Hamiltonian H of the free radiation field is obtained by generalizing, to continuous variables, expression (25) established for discrete variables. To only include independent variables in the integral over $\boldsymbol{k}$ of the Hamiltonian density $\bar{\mathcal{H}}(\boldsymbol{k},t)$, this integral is taken over only half the reciprocal space:

$$H = \fint \mathrm{d}^3k\ \bar{\mathcal{H}}(\boldsymbol{k},t) \tag{54}$$

where the Hamiltonian density $\bar{\mathcal{H}}(\boldsymbol{k},t)$ is equal to:

$$\bar{\mathcal{H}}(\boldsymbol{k},t) = -\bar{\mathscr{L}}(\boldsymbol{k},t) + \sum_{\boldsymbol{\varepsilon}}\left[\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\tilde{\Pi}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) + \dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\right] \tag{55}$$

which yields for H, taking (53) and (52) into account:

$$H = \varepsilon_0\fint \mathrm{d}^3k\ \sum_{\boldsymbol{\varepsilon}}\left[\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) + c^2k^2\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\right]. \tag{56}$$

If the integral over the half-space in (56) is extended to the entire space, one must replace ε_0 by $\varepsilon_0/2$; we then get the same expression as (B-14) of Chapter XVIII for the energy of the free transverse field. The Hamiltonian obtained with the Lagrangian formalism coincides with the field energy.

We can write expression (55) for $\bar{\mathcal{H}}(\boldsymbol{k},t)$ as a function of only the variables $\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ and $\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$. Using (53), we get:

$$\bar{\mathcal{H}}(\boldsymbol{k},t)=\sum_{\boldsymbol{\varepsilon}}\left[\frac{1}{\varepsilon_0}\tilde{\Pi}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)+\varepsilon_0 c^2k^2\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\right] \tag{57}$$

Equations (28) can then be generalized and yield the Hamilton-Jacobi equations for $\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ and $\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$:

$$\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)=\frac{\partial\bar{\mathcal{H}}(\boldsymbol{k},t)}{\partial\tilde{\Pi}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)}=\frac{1}{\varepsilon_0}\tilde{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \tag{58a}$$

$$\dot{\tilde{\Pi}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)=-\frac{\partial\bar{\mathcal{H}}(\boldsymbol{k},t)}{\partial\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)}=-\varepsilon_0 c^2k^2\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \tag{58b}$$

It is easy to check that the two equations (58a) and (58b) are the same as Maxwell's equations (A-30a) and (A-30b) of Chapter XVIII, which describes the evolution of the transverse fields in the absence of sources. Equation (58a) of the present complement and equations (A-30a) of Chapter XVIII are identical, and define the transverse electric field as a function of the time derivative of the transverse potential vector. As for equation (58b) of this complement and equation (A-30b) of Chapter XVIII, they are the same when $\tilde{\boldsymbol{j}}=\boldsymbol{0}$. They describe the evolution of the transverse electric field.

2-e. Field commutation relations

Generalizing the canonical commutation relation (29) to the case of continuous variables, we get:

$$\left[\hat{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k}),\hat{\tilde{\Pi}}^{\dagger}_{\perp\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right]=i\hbar\,\delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}\delta(\boldsymbol{k}-\boldsymbol{k}') \tag{59}$$

all the other commutators being equal to zero. The Kronecker delta $\delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}$ of the vectors $\boldsymbol{\varepsilon}$ and $\boldsymbol{\varepsilon}'$ is equal to 1 if both these vectors are the same, and to 0 if they are different.

Comment

The canonical commutation relations only apply to independent conjugate variables, which is the case for the components of the various fields along the transverse polarization directions. Now the field components on an arbitrary fixed reference frame$\{\boldsymbol{e}_x,\ \boldsymbol{e}_y,\ \boldsymbol{e}_z\}$, $\tilde{A}_{\perp i}(\boldsymbol{k})$ with $i=x,y,z$, are not independent because of the transversality condition $\sum_i k_i\tilde{A}_{\perp i}(\boldsymbol{k})=0$. Therefore:

$$\left[\hat{\tilde{A}}_{\perp i}(\boldsymbol{k}),\hat{\tilde{\Pi}}^{\dagger}_{\perp j}(\boldsymbol{k}')\right]\neq i\hbar\,\delta_{ij}\,\delta(\boldsymbol{k}-\boldsymbol{k}') \tag{60}$$

To get the correct commutation relation between $\hat{\tilde{A}}_{\perp i}(\boldsymbol{k})$ and $\hat{\tilde{\Pi}}^{\dagger}_{\perp j}(\boldsymbol{k}')$, we must express both quantities as functions of their components along the two polarization vectors $\boldsymbol{\varepsilon}$ and

$\boldsymbol{\varepsilon}'$, which are perpendicular both to each other and to $\boldsymbol{k}$ (here we choose a basis of linear, real, polarizations), and then use (59). As an example:

$$\hat{\tilde{A}}_{\perp i}(\boldsymbol{k}) = \varepsilon_i \hat{\tilde{A}}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}) + \varepsilon'_i \hat{\tilde{A}}_{\perp \boldsymbol{\varepsilon}'}(\boldsymbol{k}) \tag{61}$$

where:

$$\varepsilon_i = \boldsymbol{e}_i \cdot \boldsymbol{\varepsilon} \qquad\qquad \varepsilon'_i = \boldsymbol{e}_i \cdot \boldsymbol{\varepsilon}' \tag{62}$$

We thus get:

$$\left[\hat{\tilde{A}}_{\perp i}(\boldsymbol{k}), \hat{\tilde{\Pi}}^{\dagger}_{\perp j}(\boldsymbol{k}')\right] = i\hbar \left(\varepsilon_i \varepsilon_j + \varepsilon'_i \varepsilon'_j\right) \delta(\boldsymbol{k} - \boldsymbol{k}') \tag{63}$$

This equation can be further transformed by noting that $\boldsymbol{\varepsilon}$, $\boldsymbol{\varepsilon}'$ and $\boldsymbol{k}/k$ form an orthonormal basis, so that:

$$\varepsilon_i \varepsilon_j + \varepsilon'_i \varepsilon'_j + k_i k_j / k^2 = \delta_{ij} \tag{64}$$

Finally, the correct commutation relation between $\hat{\tilde{A}}_{\perp i}(\boldsymbol{k})$ and $\hat{\tilde{\Pi}}^{\dagger}_{\perp j}(\boldsymbol{k}')$ is written:

$$\left[\hat{\tilde{A}}_{\perp i}(\boldsymbol{k}), \hat{\tilde{\Pi}}^{\dagger}_{\perp j}(\boldsymbol{k}')\right] = i\hbar \left[\delta_{ij} - \frac{k_i k_j}{k^2}\right] \delta(\boldsymbol{k} - \boldsymbol{k}') \tag{65}$$

We multiply both sides of (65) by $e^{i\boldsymbol{k}\cdot\boldsymbol{r}} e^{-i\boldsymbol{k}'\cdot\boldsymbol{r}'}/(2\pi)^3$ and integrate over $\boldsymbol{k}$ and $\boldsymbol{k}'$. We then get on the left-hand side the commutator of the fields in real space[8], $[\hat{A}_{\perp i}(\boldsymbol{r}), \hat{\Pi}_{\perp j}(\boldsymbol{r}')]$, and on the right-hand side the Fourier transform of the function $(\delta_{ij} - k_i k_j/k^2)$ which is the *transverse delta function*[9] $\delta^{\perp}_{ij}(\boldsymbol{r} - \boldsymbol{r}')$:

$$\begin{aligned}
&\left[\hat{A}_{\perp i}(\boldsymbol{r}), \hat{\Pi}_{\perp j}(\boldsymbol{r}')\right] \\
&\quad = \frac{i\hbar}{(2\pi)^3} \int \mathrm{d}^3 k \int \mathrm{d}^3 k' \; e^{-i\boldsymbol{k}'\cdot\boldsymbol{r}'} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \left[\delta_{ij} - \frac{k_i k_j}{k^2}\right] \delta(\boldsymbol{k} - \boldsymbol{k}') \\
&\quad = \frac{i\hbar}{(2\pi)^3} \int \mathrm{d}^3 k \; e^{i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')} \left(\delta_{ij} - \frac{k_i k_j}{k^2}\right) \equiv i\hbar \delta^{\perp}_{ij}(\boldsymbol{r} - \boldsymbol{r}')
\end{aligned} \tag{66}$$

2-f. Creation and annihilation operators

The normal variable $\boldsymbol{\alpha}(\boldsymbol{k}, t)$ introduced in equation (B-15) of Chapter XVIII has a time evolution $e^{-i\omega t}$, where $\omega = ck$ for the free field case. According to § B-1 of Chapter XVIII, this variable becomes, upon quantization, the annihilation operator $\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}, t)$ of the harmonic oscillator associated with the mode $\{\boldsymbol{k}, \boldsymbol{\varepsilon}\}$. As for the complex conjugate of this normal variable, it becomes the creation operator $\hat{a}^{\dagger}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}, t)$. These two operators are therefore written as:

$$\begin{aligned}
\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}, t) &= \mathcal{N}(k) \left[\hat{A}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t) + \frac{i}{\varepsilon_0 \omega} \hat{\Pi}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t)\right] \\
\hat{a}^{\dagger}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}, t) &= \mathcal{N}(k) \left[\hat{A}^{\dagger}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t) - \frac{i}{\varepsilon_0 \omega} \hat{\Pi}^{\dagger}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t)\right]
\end{aligned} \tag{67}$$

[8] For the term in $\hat{\tilde{\Pi}}^{\dagger}_{\perp j}(\boldsymbol{k}')$, we use the reality condition $\hat{\tilde{\Pi}}^{\dagger}_{\perp j}(\boldsymbol{k}') = \hat{\tilde{\Pi}}_{\perp j}(-\boldsymbol{k}')$ and change $\boldsymbol{k}'$ into $-\boldsymbol{k}'$ in the integral over $\boldsymbol{k}'$.

[9] The interested reader can find details on the properties of that function in Complement A_I of reference [16].

where we have used (53) to write $\dot{\hat{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k}) = (1/\varepsilon_0)\hat{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k})$. The quantity $\mathcal{N}(k)$ is a normalization constant, arbitrary for now. It can, however, be determined by imposing for the commutator of the two operators (67) a generalization of the well-known relation $[\hat{a}, \hat{a}^\dagger] = 1$ for the harmonic oscillator. We thus use the two equations (67) to compute the commutator $\left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}), \hat{a}^\dagger_{\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right]$ as a function of the commutators of the fields $\hat{A}_\perp$ and $\hat{\Pi}^\dagger_\perp$ and of their adjoints. Since the only non-zero commutators are between $\hat{A}_\perp$ and $\hat{\Pi}^\dagger_\perp$ as well as between $\hat{A}^\dagger_\perp$ and $\hat{\Pi}_\perp$, we get:

$$\begin{aligned}\left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}), \hat{a}^\dagger_{\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right] &= \mathcal{N}^2(k)\frac{-i}{\varepsilon_0\omega}\left\{\left[\hat{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k}), \hat{\Pi}^\dagger_{\perp\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right] - \left[\hat{\Pi}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k}), \hat{A}^\dagger_{\perp\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right]\right\} \\ &= \mathcal{N}^2(k)\frac{-i}{\varepsilon_0\omega}\left\{2i\hbar\delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}\delta(\boldsymbol{k}-\boldsymbol{k}')\right\} \\ &= \mathcal{N}^2(k)\frac{2\hbar}{\varepsilon_0\omega}\,\delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}\delta(\boldsymbol{k}-\boldsymbol{k}') \end{aligned} \tag{68}$$

To go from the first to the second line of (68), we used (59) and its complex conjugate. The constant $\mathcal{N}(k)$ is finally determined by imposing the commutators between the annihilation and creation operators to be equal to $\delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}\delta(\boldsymbol{k}-\boldsymbol{k}')$, which yields:

$$\mathcal{N}(k) = \sqrt{\frac{\varepsilon_0\omega}{2\hbar}} = \sqrt{\frac{\varepsilon_0\ ck}{2\hbar}} \tag{69}$$

Inserting this relation in equality (B-25) of Chapter XVIII, we find that the contribution to the classical energy of the mode $\{\mathbf{k}, \boldsymbol{\varepsilon}\}$ is:

$$\frac{\hbar\omega}{2}\left[\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k}) + \alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\right] \tag{70}$$

We shall see in Chapter XIX that it is indeed the equivalent of the expression for the quantized radiation Hamiltonian.

2-g. Discrete momentum variables

We examined, in § B-3 of Chapter XVIII, the case where the radiation is contained in a box of finite volume L^3, which leads to a discrete summation over the momenta. Relation (59) then become:

$$\left[\hat{\hat{A}}_{\perp\boldsymbol{k},\varepsilon}, \hat{\hat{\Pi}}^{\dagger}_{\perp\boldsymbol{k}',\varepsilon'}\right] = i\hbar\ \delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}\delta_{\boldsymbol{k},\boldsymbol{k}'} \tag{71}$$

Applying the substitution (B-34) or (B-35) of that chapter to both sides of (67), the two coefficients $(2\pi/L)^3$ cancel each other, and these relations remain unchanged (aside from the fact that $\mathbf{k}$ is now a discrete index rather than a continuous variable).

As for the relations (68), they become:

$$\left[\hat{a}_{\boldsymbol{k},\varepsilon}, \hat{a}^\dagger_{\boldsymbol{k}',\varepsilon'}\right] = \mathcal{N}^2(k)\frac{2\hbar}{\varepsilon_0\omega}\ \delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}\delta_{\boldsymbol{k},\boldsymbol{k}'} \tag{72}$$

With the choice (69) for $\mathcal{N}(k)$, we check that the commutator is equal to unity if $\boldsymbol{k} = \boldsymbol{k}'$ and $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}'$.

3. Lagrangian of the global system field + interacting particles

We now study the Lagrangian of the total system, including the interactions between the particles and the electromagnetic field.

3-a. Choice for the Lagrangian

We choose a Lagrangian expressed as:

$$L = L_R + L_P + L_I \tag{73}$$

where L_R depends only on the radiation variables, L_P only on the particle variables, and L_I on both types of variables as it describes the interactions between particles and radiation.

For L_R, we shall take the Lagrangian introduced above for the free field – see relations (51) and (52):

$$L_R = \oint \mathrm{d}^3k\, \bar{\mathscr{L}}_R(\boldsymbol{k},t) = \varepsilon_0 \oint \mathrm{d}^3k \left[\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) - \omega^2 \tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\right]. \tag{74}$$

For the Lagrangian L_P of the particles, labeled by the index a, we shall use the usual Lagrangian for a system of particles, i.e. the difference between their kinetic energy and their potential interaction energy which comes from the Coulomb forces they exert on each others:

$$L_P = \sum_a \frac{1}{2} m_a \left[\dot{\boldsymbol{r}}_a(t)\right]^2 - V_{\mathrm{Coul}} \tag{75}$$

Finally, the interaction Lagrangian will be chosen as:

$$L_I = \int \mathrm{d}^3r\, \boldsymbol{j}(\boldsymbol{r},t) \cdot \boldsymbol{A}_\perp(\boldsymbol{r},t) \tag{76}$$

where $\boldsymbol{j}(\boldsymbol{r},t)$ is the particle current density given by the expression:

$$\boldsymbol{j}(\boldsymbol{r},t) = \sum_a q_a \dot{\boldsymbol{r}}_a(t)\, \delta\left(\boldsymbol{r} - \boldsymbol{r}_a(t)\right) \tag{77}$$

and q_a is the charge of particle a. This expression does not contain terms including $\boldsymbol{A}_\parallel$, the charge density $\rho(\boldsymbol{r},t)$ or the scalar potential $U(\boldsymbol{r}_a,t)$. This is because of our choice of the Coulomb gauge in which $\boldsymbol{A}_\parallel$ is zero, so that the energy of the longitudinal electric field only depends on the particle variables; the same is true for the scalar potential, which is at the origin of the term V_{Coul} included in the particle Lagrangian (75).

For the following computations, it will be useful to give other equivalent expressions for L_I; depending on the problem we focus on, we shall use the most suitable expression. Inserting (77) into (76), we get:

$$L_I = \int \mathrm{d}^3r\, \boldsymbol{j}(\boldsymbol{r},t) \cdot \boldsymbol{A}_\perp(\boldsymbol{r},t) = \sum_a q_a \dot{\boldsymbol{r}}_a(t) \cdot \boldsymbol{A}_\perp(\boldsymbol{r}_a,t) \tag{78}$$

Now, using the Parseval-Plancherel identity, we get:

$$L_I = \int \mathrm{d}^3r\, \boldsymbol{j}(\boldsymbol{r},t) \cdot \boldsymbol{A}_\perp(\boldsymbol{r},t) = \int \mathrm{d}^3k\, \tilde{\boldsymbol{j}}^*(\boldsymbol{k},t) \cdot \tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) \tag{79}$$

In the integral over k in (79) we get the sum:

$$\tilde{\boldsymbol{j}}^*(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)+\tilde{\boldsymbol{j}}^*(-\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_\perp(-\boldsymbol{k},t)=\tilde{\boldsymbol{j}}^*(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)+\tilde{\boldsymbol{j}}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_\perp^*(\boldsymbol{k},t) \quad (80)$$

Limiting the integral to half the reciprocal space, we can write:

$$\begin{aligned} L_I &= \int \mathrm{d}^3k\ \tilde{\boldsymbol{j}}^*(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) \\ &= \fint \mathrm{d}^3k\ \left[\tilde{\boldsymbol{j}}^*(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t)+\tilde{\boldsymbol{j}}(\boldsymbol{k},t)\cdot\tilde{\boldsymbol{A}}_\perp^*(\boldsymbol{k},t)\right] \end{aligned} \quad (81)$$

3-b. Lagrange's equations

We now show that the Lagrange's equations associated with (73) coincide with the Maxwell-Lorentz equations, which will be a justification for the choice of L.

α. Field Lagrange's equations

The time derivative of the transverse potential vector $\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}$ only appears in the Lagrangian density $\bar{\mathscr{L}}_R(\boldsymbol{k},t)$ written in (74). Writing $\bar{\mathscr{L}}$ the Lagrangian density of the total system, we can write:

$$\frac{\partial\bar{\mathscr{L}}}{\partial\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)}=\frac{\partial\bar{\mathscr{L}}_R}{\partial\dot{\tilde{A}}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)}=\varepsilon_0\dot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \quad (82)$$

Now $\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ appears both in $\bar{\mathscr{L}}_R$ and in $\bar{\mathscr{L}}_I$ which is the function to be integrated in the last integral of (81). We get:

$$\frac{\partial\bar{\mathscr{L}}}{\partial\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t))}=\frac{\partial\bar{\mathscr{L}}_R}{\partial\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t))}+\frac{\partial\bar{\mathscr{L}}_I}{\partial\tilde{A}^*_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)}=-\varepsilon_0c^2k^2\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)+\tilde{j}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \quad (83)$$

The field Lagrange's equation is obtained by setting equal the time derivative of (82) and relation (83), which yields:

$$\ddot{\tilde{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)+c^2k^2\tilde{A}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t)=\frac{1}{\varepsilon_0}\tilde{j}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \quad (84)$$

We thus get the time evolution equation of the transverse potential vector in the presence of sources, which we already obtained in (A-31) of Chapter XVIII.

β. Lagrange's equations for the particles

The velocity $\dot{\boldsymbol{r}}_a$ of particle a appears in both L_P and L_I. Noting $\dot{x}_a$ the component of $\dot{\boldsymbol{r}}_a$ on the x axis, we get:

$$\frac{\partial L}{\partial\dot{x}_a}=\frac{\partial L_P}{\partial\dot{x}_a}+\frac{\partial L_I}{\partial\dot{x}_a}=m_a\dot{x}_a+q_aA_{\perp x}\left(\boldsymbol{r}_a(t),t\right) \quad (85)$$

We now compute the time derivative of this expression. The time dependence of the second term of (85) is explicit via the time dependence of the vector potential, implicit

via the time dependence of the point $\boldsymbol{r}_a(t)$ where this potential is evaluated. We therefore have:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}_a} = m_a\ddot{x}_a + q_a\frac{\partial A_{\perp x}(\boldsymbol{r}_a,t)}{\partial t} + q_a\dot{\boldsymbol{r}}_a\cdot\boldsymbol{\nabla}A_{\perp x}(\boldsymbol{r}_a,t) \tag{86}$$

The partial derivative of the transverse vector potential with respect to t leads to the transverse electric field:

$$q_a\frac{\partial A_{\perp x}(\boldsymbol{r}_a,t)}{\partial t} = -q_aE_{\perp x}(\boldsymbol{r}_a,t) \tag{87}$$

As for the last term of (86), it can be written:

$$q_a\dot{\boldsymbol{r}}_a\cdot\boldsymbol{\nabla}A_{\perp x}(\boldsymbol{r}_a,t) = q_a\left[\dot{x}_a\frac{\partial A_{\perp x}}{\partial x_a} + \dot{y}_a\frac{\partial A_{\perp x}}{\partial y_a} + \dot{z}_a\frac{\partial A_{\perp x}}{\partial z_a}\right] \tag{88}$$

We now compute the partial derivative of L with respect to the component x_a of $\boldsymbol{r}_a$, which appears in the term V_{Coul} of L_P as well as in L_I (via the position dependence of $\boldsymbol{A}_\perp$). We obtain:

$$\frac{\partial L}{\partial x_a} = \frac{\partial L_P}{\partial x_a} + \frac{\partial L_I}{\partial x_a} = -\frac{\partial V_{\mathrm{Coul}}}{\partial x_a} + q_a\dot{\boldsymbol{r}}_a\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial x_a} \tag{89}$$

The term $-\partial V_{\mathrm{Coul}}/\partial x_a$ is the Coulomb force exerted on the particle a, i.e. the force due to the electrostatic field created by the charge distribution and acting on the charge q_a at point $\boldsymbol{r}_a$ where the particle a is located. It can also be written as:

$$-\frac{\partial V_{\mathrm{Coul}}}{\partial x_a} = q_aE_{\parallel x}(\boldsymbol{r}_a,t) \tag{90}$$

where $\boldsymbol{E}_\parallel$ is the longitudinal electric field, since, as we saw in Chapter XVIII, the longitudinal electric field is equal to the electrostatic field created by the charge distribution. Finally, let us explicitly write the last term of (89):

$$q_a\dot{\boldsymbol{r}}_a\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial x_a} = q_a\left[\dot{x}_a\frac{\partial A_x}{\partial x_a} + \dot{y}_a\frac{\partial A_y}{\partial x_a} + \dot{z}_a\frac{\partial A_z}{\partial x_a}\right] \tag{91}$$

Lagrange's equation for particle a :

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}_a} = \frac{\partial L}{\partial x_a} \tag{92}$$

can thus be written, using the previous results:

$$m_a\ddot{x}_a = q_aE_x(\boldsymbol{r}_a,t) - q_a\dot{\boldsymbol{r}}_a\cdot\boldsymbol{\nabla}A_{\perp x} + q_a\dot{\boldsymbol{r}}_a\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial x} \tag{93}$$

where the total electric field at point $\boldsymbol{r}_a$ is :

$$\boldsymbol{E}(\boldsymbol{r}_a) = \boldsymbol{E}_\perp(\boldsymbol{r}_a) + \boldsymbol{E}_\parallel(\boldsymbol{r}_a) \tag{94}$$

We can write explicitly the last term of (93) by regrouping equations (88) and (91). We then get:

$$
\begin{aligned}
&- q_a \dot{\boldsymbol{r}}_a \cdot \boldsymbol{\nabla} A_{\perp x} + q_a \dot{\boldsymbol{r}}_a \cdot \frac{\partial \boldsymbol{A}_\perp}{\partial x_a} \\
&\qquad = q_a \left[\dot{y}_a \left(\frac{\partial A_{\perp y}}{\partial x_a} - \frac{\partial A_{\perp x}}{\partial y_a} \right) - \dot{z}_a \left(\frac{\partial A_{\perp x}}{\partial z_a} - \frac{\partial A_{\perp z}}{\partial x_a} \right) \right] \\
&\qquad = q_a \left[\dot{\boldsymbol{r}}_a \times (\boldsymbol{\nabla} \times \boldsymbol{A}_\perp) \right]_x = q_a (\dot{\boldsymbol{r}}_a \times \boldsymbol{B})_x
\end{aligned} \tag{95}
$$

This yields the Lorentz magnetic force exerted by the magnetic field on the particle with velocity $\dot{\boldsymbol{r}}_a$. To sum up, Lagrange's equation for particle a is written:

$$
m_a \ddot{\boldsymbol{r}}_a = q_a \boldsymbol{E}(\boldsymbol{r}_a) + q_a \dot{\boldsymbol{r}}_a \times \boldsymbol{B}(\boldsymbol{r}_a) \tag{96}
$$

and coincides with the Lorentz equation (A-3) of ChapterXVIII. The Lagrangian we chose above therefore leads to the right equations for the field and the particles.

3-c. Conjugate momenta

The equation (82) established above can be used to compute the conjugate momenta $\tilde{\Pi}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t)$ of the field variables $\tilde{A}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t)$:

$$
\tilde{\Pi}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t) = \frac{\partial \bar{\mathscr{L}}}{\partial \dot{\tilde{A}}^*_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t)} = \frac{\partial \bar{\mathscr{L}}_R}{\partial \dot{\tilde{A}}^*_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t)} = \varepsilon_0 \dot{\tilde{A}}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t)) = -\varepsilon_0 \tilde{E}_{\perp \boldsymbol{\varepsilon}}(\boldsymbol{k}, t) \tag{97}
$$

In a similar way, the computation of the conjugate momentum of the coordinate $\boldsymbol{r}_a$ of particle a is the same as the one leading to equation (85):

$$
\boldsymbol{p}_a = \frac{\partial L}{\partial \dot{\boldsymbol{r}}_a} = \frac{\partial L_P}{\partial \dot{\boldsymbol{r}}_a} + \frac{\partial L_I}{\partial \dot{\boldsymbol{r}}_a} = m_a \dot{\boldsymbol{r}}_a + q_a \boldsymbol{A}_\perp(\boldsymbol{r}_a) \tag{98}
$$

3-d. Hamiltonian

Since $\dot{\tilde{A}}_{\perp \boldsymbol{\varepsilon}}$ and $\tilde{\Pi}^*_{\perp \varepsilon} = \varepsilon_0 \dot{\tilde{A}}^*_{\perp \boldsymbol{\varepsilon}}$ appear only in the radiation Lagrangian $\mathscr{L}_R$, which only depends on the radiation variables, the Hamiltonian of the global system must contain a term H_R identical to the expression (56) found for the free field.

To obtain the other terms coming from the particle conjugate momenta and from the subtraction of L_P and L_I, we first compute $\sum_a \dot{\boldsymbol{r}}_a \cdot \boldsymbol{p}_a$. Equation (98) yields:

$$
\sum_a \dot{\boldsymbol{r}}_a \cdot \boldsymbol{p}_a = \sum_a \left[m_a \dot{\boldsymbol{r}}_a^2 + q_a \dot{\boldsymbol{r}}_a \cdot \boldsymbol{A}_\perp(\boldsymbol{r}_a) \right] \tag{99}
$$

We must now subtract from (99) the values of L_P and L_I given by equations (75) and (78). The term coming from L_I cancels the last term of the right-hand side of (99) and we are left with:

$$
V_{\text{Coul}} + \sum_a \frac{1}{2} m_a \dot{\boldsymbol{r}}_a^2 = V_{\text{Coul}} + \sum_a \frac{\left[\boldsymbol{p}_a - q_a \boldsymbol{A}_\perp(\boldsymbol{r}_a) \right]^2}{2 m_a} \tag{100}
$$

Finally, the global system Hamiltonian is given by:

$$H = H_R + V_{\text{Coul}} + \sum_a \frac{[\boldsymbol{p}_a - q_a \boldsymbol{A}_\perp(\boldsymbol{r}_a)]^2}{2m_a} \tag{101}$$

where H_R has the same form as (56) for the free radiation. This result is a justification for the expression of H given in equation (A-41) of Chapter XVIII.

3-e. Commutation relations

Since the radiation variables and their conjugate momenta are the same as for the free field, the commutation relations (59) established for the free field remain valid:

$$\left[\hat{\vec{A}}_{\perp\boldsymbol{\varepsilon}}(\boldsymbol{k}), \hat{\vec{\Pi}}^{\dagger}_{\perp\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right] = i\hbar\ \delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}\ \delta(\boldsymbol{k} - \boldsymbol{k}') \tag{102}$$

all the other commutators being equal to zero. As for the commutation relations for the positions and conjugate momenta of the particles, they are the usual relations:

$$[\hat{r}_{ai}, \hat{p}_{bj}] = i\hbar\ \delta_{ab}\ \delta_{ij} \tag{103}$$

where the indices a, b label the particles and the indices $i, j = x, y, z$ the Cartesian components of $\boldsymbol{r}$ and $\boldsymbol{p}$.

As in § 2-g, one can extend these commutation relations to the case where the momenta are discrete.

Chapter XIX

Quantization of electromagnetic radiation

A Quantization of the radiation in the Coulomb gauge 1999
A-1 Quantization rules 1999
A-2 Radiation contained in a box 2001
A-3 Heisenberg equations 2002
B Photons, elementary excitations of the free quantum field . 2004
B-1 Fock space of the free quantum field 2004
B-2 Corpuscular interpretation of states with fixed total energy and momentum 2005
B-3 Several examples of quantum radiation states 2006
C Description of the interactions 2009
C-1 Interaction Hamiltonian 2009
C-2 Interaction with an atom. External and internal variables . . 2010
C-3 Long wavelength approximation 2010
C-4 Electric dipole Hamiltonian 2011
C-5 Matrix elements of the interaction Hamiltonian; selection rules 2014

Introduction

This chapter presents a quantum description of the electromagnetic field and its interactions with an ensemble of charged particles. Such a description is necessary for interpreting certain physical phenomena such as the spontaneous emission of a photon by an excited atom, which cannot be carried out with the semiclassical treatments we have used previously[1] (classical description for the field, and quantum description for the

[1] See for example in Complement A_{XIII} the study of the interaction between an atom and an electromagnetic wave.

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

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

particles). Imagine, for example, that a monochromatic field with angular frequency ω is described by a classical field $\mathbf{E}_0 \cos \omega t$; its interaction with an atom is then described by the Hamiltonian $\widehat{H}_I = -\widehat{\mathbf{D}} \cdot \mathbf{E}_0 \cos \omega t$, where $\widehat{\mathbf{D}}$ is an operator (the electric dipole moment) whereas $\mathbf{E}_0$ remains a classical quantity[2]. Such a treatment is adequate for understanding how the field can excite the atom from its ground state a with energy E_a towards an excited state b of energy E_b; the processus is resonant if ω is close to the atomic Bohr frequency $\omega_0 = (E_b - E_a)/\hbar$. Imagine now that the atom is initially in the excited state b, in the absence of any incident radiation. The classical field $\mathbf{E}_0$ is then identically zero and, consequently, so is the interaction Hamiltonian $\widehat{H}_I$. The Hamiltonian of the total system is then reduced to the atomic Hamiltonian $\widehat{H}_A$. Since this operator is time-independent, its eigenstates are stationary, including, in particular, the excited state b. The semiclassical theory predicts that an atom, initially excited in a state b in the absence of incident radiation, will remain indefinitely in that state. But this is not what is experimentally observed: after a certain time, the atom spontaneously falls into a lower level a, emitting a photon whose frequency is close to $\omega_0 = (E_b - E_a)/\hbar$. This process is called spontaneous emission and happens after an average time called the *radiative lifetime* of the excited state b. This is a first example of a situation where a radiation quantum treatment is indispensable. It is far from being the only example: numerous experiments, more and more elaborate, have created situations where the quantum description of the electromagnetic field is necessary.

This chapter presents the base of this quantum description, while following an approach that is as simple as possible – a more general presentation of the quantization of the electromagnetic field is possible with the Lagrangian formulation of electrodynamics (Complement $\mathrm{A_{XVIII}}$. In the previous chapter, we underlined the analogy between the eigenmodes of the radiation field vibrations and an ensemble of harmonic oscillators. We shall use this analogy in § A of this chapter, and proceed to a simple quantization of this ensemble of oscillators. With each eigenmode i of the classical field, described by normal variables α_i and α_i^*, we shall associate annihilation $\widehat{a}_i$ and creation $\widehat{a}_i^\dagger$ operators, obeying the well-known commutation relations $\left[\widehat{a}_i, \widehat{a}_i^\dagger\right] = 1$. We shall also propose a plausible form for the quantum Hamiltonian of the system "field + particles", starting from the classical energy of that system established in the previous chapter. We will see that the equations of evolution[3] for these various quantities in the Heisenberg picture (Complement $\mathrm{G_{III}}$) are the transposition of the Maxwell-Lorentz equations to operators describing fields and particles, properly symmetrized. This will yield an a posteriori justification for the simple quantization procedure we used.

Several important properties of the free field (in the absence of sources) are described in § B. The state space of this field has the structure of a tensor product of Fock spaces, analogous to those studied in Chapter XV; the elementary excitations of the field are called photons. A few important states of the field will be described: the photon vacuum, where no photons are present (but where there exists, nonetheless, a fluctuating field throughout the entire space, with a zero average value), the one-photon states, and the quasi-classical states, which reproduce the properties of a given classical field.

Finally, § C studies the interaction Hamiltonian between an electromagnetic field and particles, in particular when those are neutral atoms (such as the Hydrogen atom

[2]For the sake of clarity, we use in the entire chapter and its complements the symbol "hat" to distinguish an operator $\widehat{G}$ from its corresponding classical quantity G.

[3]More concisely, we shall call them Heinsenberg equations.

where the positive and negative charges of the atom's constituents balance each other). It is then possible to distinguish between two types of atomic variables: the center of mass variables (external variables) and the "relative motion" variables in the center of mass frame (internal variables). We shall also study the electric dipole approximation, valid when the radiation wavelength is large compared to the atomic sizes, as well as the selection rules associated with the interaction Hamiltonian.

A. Quantization of the radiation in the Coulomb gauge

A-1. Quantization rules

In the previous chapter, we established in relation (B-26) the following expression for the energy of the classical transverse field:

$$H_{\text{trans}} = \varepsilon_0 \int \mathrm{d}^3k \sum_{\boldsymbol{\varepsilon}} \frac{\omega^2}{4\mathcal{N}^2(k)} \left[\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t) + \alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\right] \tag{A-1}$$

where $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ and $\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ are the normal variables describing the transverse field, $\omega = ck$, and $\mathcal{N}(k)$ a real normalization constant that appeared in the equations defining the normal variables in terms of the transverse potential vector and its time derivative:

$$\begin{aligned}
\boldsymbol{\alpha}(\boldsymbol{k},t) &= \mathcal{N}(k)\left[\tilde{\boldsymbol{A}}_\perp(\boldsymbol{k},t) + \frac{i}{\omega}\dot{\tilde{\boldsymbol{A}}}_\perp(\boldsymbol{k},t)\right] \\
\boldsymbol{\alpha}^*(\boldsymbol{k},t) &= \mathcal{N}(k)\left[\tilde{\boldsymbol{A}}^*_\perp(\boldsymbol{k},t) - \frac{i}{\omega}\dot{\tilde{\boldsymbol{A}}}^*_\perp(\boldsymbol{k},t)\right]
\end{aligned} \tag{A-2}$$

The analogy between the free transverse field and an ensemble of classical harmonic oscillators of frequency ω associated with the modes $\{\boldsymbol{k},\boldsymbol{\varepsilon}\}$ is clearly seen in expression (A-1).

To quantize the field, this analogy suggests replacing the normal variables $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ and $\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ by annihilation and creations operators. We shall use in this § A the Schrödinger picture where these operators are time-independent and where the time dependence only appears in the evolution of the state vector. The quantization procedure will consist in replacing the $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t=0)$ by time-independent annihilation operators $\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$, and of course the $\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t=0)$ by the adjoint creation operators $\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$. Once this operation is performed on (A-1), we obtain a quantum Hamiltonian identical to a sum of standard harmonic oscillator Hamiltonians, provided the factor $\omega^2/4\mathcal{N}^2(k)$ multiplying the bracket on the right-hand side of (A-1) is equal to $\hbar\omega/2\epsilon_0$. We therefore choose for $\mathcal{N}(k)$ the value:

$$\mathcal{N}(k) = \sqrt{\frac{\varepsilon_0\, ck}{2\hbar}} = \sqrt{\frac{\varepsilon_0\,\omega}{2\hbar}} \tag{A-3}$$

This relation is the same as relation (69) of Complement $\mathrm{A_{XVIII}}$, obtained from the commutation relations. We now replace in (A-1) the classical normal variables $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ and $\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ by the operators $\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ and $\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ obeying the commutation relations:

$$\left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}), \hat{a}^\dagger_{\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right] = \delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'}\ \delta(\boldsymbol{k}-\boldsymbol{k}') \tag{A-4a}$$

$$\left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}), \hat{a}_{\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right] = \left[\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k}), \hat{a}^\dagger_{\boldsymbol{\varepsilon}'}(\boldsymbol{k}')\right] = 0 \tag{A-4b}$$

This yields the Hamiltonian operator (as this operator will be frequently used, we simplify the notation and replace $\hat{H}_{\rm trans}$ by $\hat{H}_R$):

$$\hat{H}_R \equiv \hat{H}_{\rm trans} = \int \mathrm{d}^3k \sum_{\boldsymbol{\varepsilon}} \frac{\hbar\omega}{2} \left[\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}) + \hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\right] \tag{A-5}$$

which has the expected form for the quantum Hamiltonian of the transverse field.

Extending this procedure, we now replace the classical normal variables by annihilation and creation operators in all the classical expressions established in the previous chapter for the various physical quantities. The transverse momentum – see equation (B-27) of Chapter XVIII – becomes:

$$\hat{\boldsymbol{P}}_{\rm trans} = \int \mathrm{d}^3k \sum_{\boldsymbol{\varepsilon}} \frac{\hbar\boldsymbol{k}}{2} \left[\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}) + \hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\right] \tag{A-6}$$

As for the transverse fields, written in (B-29), (B-30) and (B-28) of Chapter XVIII, they become linear combinations of creation and annihilation operators:

$$\hat{\boldsymbol{E}}_\perp(\boldsymbol{r}) = i \int \frac{\mathrm{d}^3k}{(2\pi)^{3/2}} \sum_{\boldsymbol{\varepsilon}} \left[\frac{\hbar\omega}{2\varepsilon_0}\right]^{1/2} \left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\, \boldsymbol{\varepsilon}\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - \hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\, \boldsymbol{\varepsilon}^*\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right] \tag{A-7}$$

$$\hat{\boldsymbol{B}}(\boldsymbol{r}) = \frac{i}{c} \int \frac{\mathrm{d}^3k}{(2\pi)^{3/2}} \sum_{\boldsymbol{\varepsilon}} \left[\frac{\hbar\omega}{2\varepsilon_0}\right]^{1/2} \left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\, \boldsymbol{\kappa}\times\boldsymbol{\varepsilon}\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - \hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\, \boldsymbol{\kappa}\times\boldsymbol{\varepsilon}^*\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right] \tag{A-8}$$

$$\hat{\boldsymbol{A}}_\perp(\boldsymbol{r}) = \int \frac{\mathrm{d}^3k}{(2\pi)^{3/2}} \sum_{\boldsymbol{\varepsilon}} \left[\frac{\hbar}{2\varepsilon_0\omega}\right]^{1/2} \left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\, \boldsymbol{\varepsilon}\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} + \hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\, \boldsymbol{\varepsilon}^*\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right] \tag{A-9}$$

Comment:

As in Chapter XVIII, these relations are written in the general case where the polarizations may be complex (elliptical or circular). Complex conjugate $\boldsymbol{\varepsilon}*$ of the polarization vectors are therefore associated with the creation operators. It is of course necessary to check that the quantification procedure is independent of the arbitrary choice of the polarization basis. If a quantization is performed with a given basis of polarizations, by substitution one can calculate the operators multiplying $\boldsymbol{\varepsilon}$ and $\boldsymbol{\varepsilon}^*$ in the new basis, and check that the commutation relations of these operators are indeed those of standard creation and annihilation operators. This ensures the polarization basis independence.

Finally, relation (A-48) of Chapter XVIII for the total energy of the system "particles + fields" becomes:

$$\hat{H} = \frac{1}{2m_a} \sum_a \left[\hat{\boldsymbol{p}}_a - q_a \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a)\right]^2 + \hat{V}_{\rm Coul} + \int \mathrm{d}^3k \sum_{\boldsymbol{\varepsilon}} \frac{\hbar\omega}{2} \left[\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k}) + \hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\right] \tag{A-10}$$

which is a plausible form for the quantum Hamiltonian of the system "particles + fields". The position $\hat{\boldsymbol{r}}_a$ and momentum $\hat{\boldsymbol{p}}_a$ operators defined using equation (A-47) of Chapter XVIII obey the usual commutation relations:

$$[(\hat{\boldsymbol{r}}_a)_i, (\hat{\boldsymbol{p}}_b)_j] = i\hbar\, \delta_{ab}\, \delta_{ij} \tag{A-11a}$$

$$[(\hat{\boldsymbol{r}}_a)_i, (\hat{\boldsymbol{r}}_b)_j] = [(\hat{\boldsymbol{p}}_a)_i, (\hat{\boldsymbol{p}}_b)_j] = 0 \tag{A-11b}$$

The quantization rules we just heuristically introduced have the advantage of simplicity. We are going to show in addition that the Heisenberg equations for the various operators describing the particles and the fields, deduced from the Hamiltonian (A-10) as well as from the commutation relations (A-4), (A-11a) and (A-11b), are indeed the Maxwell-Lorentz equations for operators. This result justifies a posteriori the quantization procedure exposed in this chapter.

A-2. Radiation contained in a box

If the real space is infinite, $\boldsymbol{k}$ is a continuous variable, and there exists a continuous infinity of modes. However, as we mentioned in § B-3 of Chapter XVIII, it is often more convenient to consider the field to be contained in a cube of edge length L with periodic boundary conditions; the variable $\boldsymbol{k}$ is now discrete:

$$k_{x,y,z} = 2\pi \; n_{x,y,z}/L \tag{A-12}$$

where $n_{x,y,z}$ are positive, negative or zero integers. All the physical predictions must be independent of L when it is large enough. In such an approach, we replace the Fourier integrals by Fourier series and the integrals over $\boldsymbol{k}$ by discrete summations. For a classical field, the continuous variables $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k}, t)$ then become discrete variables $\alpha_{\boldsymbol{k},\boldsymbol{\varepsilon}}(t)$. If the field is zero outside the box, relation (B-35) of chapter XVIII indicates the multiplicative factor that must be used to go from one type of variable to the other.

The system is then quantized as we just explained. In the Schrödinger picture, each classical coefficient $\alpha_{\boldsymbol{k},\boldsymbol{\varepsilon}}(t=0)$ in a Fourier series becomes an annihilation operator $\hat{a}_{\boldsymbol{k},\boldsymbol{\varepsilon}}$; each coefficient $\alpha^*_{\boldsymbol{k},\boldsymbol{\varepsilon}}(t=0)$ becomes a creation operator $\hat{a}^\dagger_{\boldsymbol{k},\boldsymbol{\varepsilon}}$. This latter operator creates a quantum in a field mode confined inside the box (instead of spreading over the entire space). The commutation relations (A-4) are then written:

$$\left[\hat{a}_{\boldsymbol{k},\boldsymbol{\varepsilon}}, \hat{a}^\dagger_{\boldsymbol{k},\boldsymbol{\varepsilon}'}\right] = \delta_{\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'} \; \delta_{\boldsymbol{k},\boldsymbol{k}'} \tag{A-13a}$$

$$[\hat{a}_{\boldsymbol{k},\boldsymbol{\varepsilon}}, \hat{a}_{\boldsymbol{k}',\boldsymbol{\varepsilon}'}] = \left[\hat{a}^\dagger_{\boldsymbol{k},\boldsymbol{\varepsilon}}, \hat{a}^\dagger_{\boldsymbol{k}',\boldsymbol{\varepsilon}'}\right] = 0 \tag{A-13b}$$

Relation (B-36) of Chapter XVIII indicates that once the discrete variables have been inserted in the expressions for the fields, the following rule must be applied to go from a continuous to a discrete summation:

$$\int \mathrm{d}^3k \implies \left(\frac{2\pi}{L}\right)^{3/2} \sum_{\mathbf{k}} \tag{A-14}$$

Expressions (A-7) to (A-9) must be modified. As an example, relation (A-7) becomes:

$$\hat{\boldsymbol{E}}_\perp(\boldsymbol{r}) = i \sum_{\boldsymbol{k},\boldsymbol{\varepsilon}} \left[\frac{\hbar\omega}{2\varepsilon_0 L^3}\right]^{1/2} \left[\hat{a}_{\boldsymbol{k},\boldsymbol{\varepsilon}} \; \boldsymbol{\varepsilon} \; e^{i\boldsymbol{k}\cdot\boldsymbol{r}} - \hat{a}^\dagger_{\boldsymbol{k},\boldsymbol{\varepsilon}} \; \boldsymbol{\varepsilon}^* \; e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\right] \tag{A-15}$$

This means that in addition to replacing the integral by a discrete summation, and multiplying by a factor $(2\pi)^{3/2}$, one must divide the field expansion by the square root of the volume L^3. Both relations (A-8) and (A-9) undergo the same changes.

A-3. Heisenberg equations

A-3-a. Heisenberg equations for massive particles

We start with the equation for the evolution of $\hat{\boldsymbol{r}}_a(t)$:

$$\dot{\hat{\boldsymbol{r}}}_a(t) = \frac{1}{i\hbar}\left[\hat{\boldsymbol{r}}_a(t), \hat{H}\right] \tag{A-16}$$

The only term in Hamiltonian (A-10) that does not commute with $\hat{\boldsymbol{r}}_a$ is the first one. Using the commutation relation deduced from (A-11a) and (A-11b) :

$$[(\hat{\boldsymbol{r}}_a)_i, f((\hat{\boldsymbol{p}}_a)_j)] = \delta_{ij}\ \hbar \frac{\partial f}{\partial (\hat{\boldsymbol{p}}_a)_i} \tag{A-17}$$

we get:

$$\begin{aligned}\dot{\hat{\boldsymbol{r}}}_a(t) &= \frac{1}{i\hbar}\left[\hat{\boldsymbol{r}}_a(t), \frac{1}{2m_a}\left(\hat{\boldsymbol{p}}_a(t) - q_a\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a, t)\right)^2\right] \\ &= \frac{1}{m_a}\left[\hat{\boldsymbol{p}}_a(t) - q_a\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a, t)\right]\end{aligned} \tag{A-18}$$

This equality is simply the operator form:

$$\hat{\boldsymbol{p}}_a(t) = m_a\dot{\hat{\boldsymbol{r}}}_a(t) + q_a\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a, t) \tag{A-19}$$

of the classical equation relating the generalized (or canonical) momentum $\boldsymbol{p}_a$ and the mechanical momentum $m_a\dot{\boldsymbol{r}}_a$. We then define the velocity operator $\hat{\boldsymbol{v}}_a$ of particle a by:

$$\hat{\boldsymbol{v}}_a(t) = \frac{1}{m_a}\left[\hat{\boldsymbol{p}}_a(t) - q_a\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a, t)\right] \tag{A-20}$$

Consider now the Heisenberg equation for the evolution of this operator. It yields the equation of motion of that particle:

$$m_a\dot{\hat{\boldsymbol{v}}}_a(t) = m_a\ddot{\hat{\boldsymbol{r}}}_a(t) = \frac{m_a}{i\hbar}\left[\hat{\boldsymbol{v}}_a(t), \hat{H}\right] \tag{A-21}$$

We shall compute below the commutator $\left[\hat{\boldsymbol{v}}_a(t), \hat{H}\right]$; it leads to the quantum equation of motion for particle a:

$$m_a\ddot{\hat{\boldsymbol{r}}}_a = q_a\hat{\boldsymbol{E}}(\hat{\boldsymbol{r}}_a) + \frac{q_a}{2}\left[\hat{\boldsymbol{v}}_a \times \hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a) - \hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a) \times \hat{\boldsymbol{v}}_a\right] \tag{A-22}$$

which is simply the quantum Lorentz equation describing the motion of particles interacting with the magnetic field $\hat{\boldsymbol{B}}$ and the total electric field $\hat{\boldsymbol{E}} = \hat{\boldsymbol{E}}_\parallel + \hat{\boldsymbol{E}}_\perp$. The special form of the magnetic force $q_a\left[\hat{\boldsymbol{v}}_a \times \hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a) - \hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a) \times \hat{\boldsymbol{v}}_a\right]/2$ comes, as shown in the computation below, from using the Heisenberg equations, and from the fact that the operator $\hat{\boldsymbol{v}}_a \times \hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a)$ is not Hermitian. To make that operator Hermitian, we must add its adjoint $\left(\hat{\boldsymbol{v}}_a \times \hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a))\right)^\dagger$, which is simply $-\hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a) \times \hat{\boldsymbol{v}}_a$, and divide the result by 2.

Demonstration of equation (A-22)

To compute the commutator of $m_a\hat{\boldsymbol{v}}_a/i\hbar$ with the first term of $\hat{H}$, it is useful to first calculate the following commutators:

$$\begin{aligned} m_a^2\left[(\hat{\boldsymbol{v}}_a)_j,(\hat{\boldsymbol{v}}_a)_l\right] &= -q_a\left[(\hat{\boldsymbol{p}}_a)_j,(\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a))_l\right]-q_a\left[(\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a))_j,(\hat{\boldsymbol{p}}_a)_l\right]\\ &= i\hbar q_a\left[\partial_j(\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a))_l-\partial_l(\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a))_j\right]\\ &= i\hbar q_a\sum_k\epsilon_{jlk}(\hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a))_k, \end{aligned} \tag{A-23}$$

where ϵ_{jkl} is the completely antisymmetric tensor that allows writing the cross product components of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ in the form $(\boldsymbol{a}\times\boldsymbol{b})_k=\sum_{jl}\epsilon_{kjl}a_jb_l$. We then get:

$$\begin{aligned} \frac{m_a}{i\hbar}\left[(\hat{\boldsymbol{v}}_a)_j,\sum_l m_a(\hat{\boldsymbol{v}}_a)_l^2/2\right] &= \frac{m_a^2}{2i\hbar}\sum_l\left\{(\hat{\boldsymbol{v}}_a)_l\left[(\hat{\boldsymbol{v}}_a)_j,(\hat{\boldsymbol{v}}_a)_l\right]+\left[(\hat{\boldsymbol{v}}_a)_j,(\hat{\boldsymbol{v}}_a)_l\right](\hat{\boldsymbol{v}}_a)_l\right\}\\ &= \frac{q_a}{2}\sum_k\sum_l\epsilon_{jlk}\left\{(\hat{\boldsymbol{v}}_a)_l(\hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a))_k+(\hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a))_k(\hat{\boldsymbol{v}}_a)_l\right\} \end{aligned} \tag{A-24}$$

The last line in (A-24) can be rewritten in the form:

$$\frac{q_a}{2}\left[\hat{\boldsymbol{v}}_a\times\hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a)-\hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a)\times\hat{\boldsymbol{v}}_a\right]_j \tag{A-25}$$

and is thus the component along the j axis of the symmmetrized magnetic force.

The commutator of $m_a\hat{\boldsymbol{v}}_a/i\hbar$ with the second term of $\hat{H}$ is written:

$$\frac{m_a}{i\hbar}\left[(\hat{\boldsymbol{v}}_a)_j,V_{\text{Coul}}\right]=\frac{1}{i\hbar}\left[(\hat{\boldsymbol{p}}_a)_j,V_{\text{Coul}}\right]=-\frac{\partial}{\partial(\hat{\boldsymbol{r}}_a)_j}V_{\text{Coul}}=q_a(\hat{\boldsymbol{E}}_\parallel(\hat{\boldsymbol{r}}_a))_j \tag{A-26}$$

It describes the interaction between particle a and the longitudinal electric field.

We finally have to compute the commutator of $m_a\hat{\boldsymbol{v}}_a/i\hbar$ with the last term of $\hat{H}$. Using the commutation relations (A-4) and expressions (A-9) and (A-7) for $\hat{\boldsymbol{A}}_\perp$ and $\hat{\boldsymbol{E}}_\perp$, we get:

$$\begin{aligned} &\frac{m_a}{i\hbar}\left[(\hat{\boldsymbol{v}}_a)_j,\int \mathrm{d}^3k\sum_{\boldsymbol{\varepsilon}}\hbar\omega\left(\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)+1/2\right)\right]\\ &= iq_a\int \mathrm{d}^3k\sum_{\boldsymbol{\varepsilon}}\omega\left[(\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a))_j,\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)\right]\\ &= q_a(\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{r}}_a))_j \end{aligned} \tag{A-27}$$

This term describes the interaction of particle a with the transverse electric field. Finally grouping (A-25), (A-26) and (A-27) leads to (A-22).

A-3-b. Heisenberg equations for fields

As all the fields are linear combinations of the operators $\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ and $\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$, we simply have to consider the Heisenberg equation for $\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$:

$$\dot{\hat{a}}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)=\frac{1}{i\hbar}\left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t),\hat{H}\right] \tag{A-28}$$

We assume that the polarizations $\boldsymbol{\varepsilon}$ are real (linear polarizations). The commutator with the first term of $\hat{H}$ yields, with the use of(A-4a) and (A-20):

$$\frac{1}{i\hbar}\left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t),\sum_a \frac{m_a\hat{\boldsymbol{v}}_a^2}{2}\right] = \sum_a \frac{-q_a}{2i\hbar}\left\{\hat{\boldsymbol{v}}_a\cdot\frac{\partial \boldsymbol{A}_\perp}{\partial\alpha^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})}+\frac{\partial \boldsymbol{A}_\perp}{\partial\alpha^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})}\cdot\hat{\boldsymbol{v}}_a\right\}$$
$$= \sum_a \frac{iq_a}{2\hbar}\sqrt{\frac{\hbar}{2\varepsilon_0\omega(2\pi)^3}}\,\boldsymbol{\varepsilon}\cdot\left[\hat{\boldsymbol{v}}_a e^{-i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}_a}+e^{-i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}_a}\hat{\boldsymbol{v}}_a\right] \qquad \text{(A-29)}$$

where $\partial\boldsymbol{A}_\perp/\partial a^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)$ denotes the coefficient of $\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ in the integral (A-9) of $\hat{\boldsymbol{A}}_\perp(\boldsymbol{r})$, which is nothing but the coefficient of $\alpha^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t=0)$ in the classical expression of $\boldsymbol{A}_\perp(\boldsymbol{r})$. We introduce the current operator (symmetrized to make it Hermitian):

$$\hat{\boldsymbol{j}}(\boldsymbol{r}) = \frac{1}{2}\sum_a q_a\left[\hat{\boldsymbol{v}}_a\delta(\boldsymbol{r}-\hat{\boldsymbol{r}}_a)+\delta(\boldsymbol{r}-\hat{\boldsymbol{r}}_a)\hat{\boldsymbol{v}}_a\right] \qquad \text{(A-30)}$$

The right hand side term of equation (A-29) can then be rewritten in the form:

$$\sum_a \frac{iq_a/2}{\sqrt{2\varepsilon_0\hbar\omega(2\pi)^3}}\,\boldsymbol{\varepsilon}\cdot\left[\hat{\boldsymbol{v}}_a e^{-i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}_a}+e^{-i\boldsymbol{k}\cdot\hat{\boldsymbol{r}}_a}\hat{\boldsymbol{v}}_a\right] = \frac{i}{\sqrt{2\varepsilon_0\hbar\omega(2\pi)^3}}\int \mathrm{d}^3r\; e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\boldsymbol{\varepsilon}\cdot\hat{\boldsymbol{j}}(\boldsymbol{r})$$
$$= \frac{i}{\sqrt{2\varepsilon_0\hbar\omega}}\,\boldsymbol{\varepsilon}\cdot\hat{\tilde{\boldsymbol{j}}}(\boldsymbol{k}) \qquad \text{(A-31)}$$

The commutator with the second term of $\hat{H}$ is zero, whereas the commutator with the third term yields, using (A-4):

$$\frac{1}{i\hbar}\left[\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t),\int \mathrm{d}^3k'\;\sum_{\boldsymbol{\varepsilon}'}\hbar\omega\left(\hat{a}^\dagger_{\boldsymbol{\varepsilon}'}(\boldsymbol{k}',t)\hat{a}_{\boldsymbol{\varepsilon}'}(\boldsymbol{k}',t)+1/2\right)\right] = -i\omega\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t) \qquad \text{(A-32)}$$

Finally, regrouping (A-31) and (A-32) yields:

$$\dot{\hat{a}}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t)+i\omega\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k},t) = \frac{i}{\sqrt{2\varepsilon_0\hbar\omega}}\,\boldsymbol{\varepsilon}\cdot\hat{\tilde{\boldsymbol{j}}}(\boldsymbol{k},t) \qquad \text{(A-33)}$$

This equation is, for the operator $\hat{a}_\varepsilon(\boldsymbol{k},t)$, an equation of motion of the same form as the equation of motion of the classical normal variables $\boldsymbol{\alpha}(\boldsymbol{k},t)$, which is given by equation (B-19) of Chapter XVIII. As this latter equation is equivalent to Maxwell's equations for the transverse fields, we may conclude that the Heisenberg equations for the quantum transverse fields are simply the usual Maxwell's equations applied to the field operators.

B. Photons, elementary excitations of the free quantum field

We now study a certain number of properties of the electromagnetic field we just quantized, starting with the simplest case: the field in the absence of charged particles.

B-1. Fock space of the free quantum field

The state space of the total system "field + particles" is the tensor product of the particle state space $\mathcal{E}_P$ and the radiation field state space $\mathcal{E}_R$. This latter space is itself

the tensor product of the state spaces of the harmonic oscillators associated with the different modes $\{\boldsymbol{k}, \boldsymbol{\varepsilon}\}$:

$$\mathcal{E}_R = \mathcal{E}_{\mathbf{k}_1,\varepsilon_1} \otimes \mathcal{E}_{\mathbf{k}_2,\varepsilon_2} \otimes .... \otimes \mathcal{E}_{\mathbf{k}_i,\varepsilon_i} \otimes ... \tag{B-1}$$

where $\mathcal{E}_{\mathbf{k}_i,\varepsilon_i}$ is the state space of the harmonic oscillator associated with the mode $\{\boldsymbol{k}_i, \boldsymbol{\varepsilon}_i\}$, with frequency ω_i.

As in § A-2, we assume the radiation to be contained in a box of edge length L. The operators $\hat{a}_\varepsilon(\boldsymbol{k})$ depending on the variables $\boldsymbol{k}$ are then transformed into operators $\hat{a}_{\boldsymbol{k}_i,\boldsymbol{\varepsilon}_i}$ depending only on discrete variables. We can even use a more compact notation $\hat{a}_i$, where the index i labels[4] the whole set of indices $\boldsymbol{k}_i, \boldsymbol{\varepsilon}_i$; the operators $\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ are now simply written $\hat{a}_i$. In this section, it is convenient to use the Heisenberg picture; the time dependence of the $\hat{a}_i$ and $\hat{a}_i^\dagger$ is then particularly simple, since we have:

$$\hat{a}_i(t) = \exp(i\hat{H}_R t/\hbar)\, \hat{a}_i\, \exp(-i\hat{H}_R t/\hbar) = \hat{a}_i\, e^{-i\omega_i t} \tag{B-2}$$

as well as the Hermitian conjugate relation.

Once the discrete variables have been inserted in the continuous expressions of the fields, we must use rule (A-14) to transform the continuous integrals into discrete summations. The expansions of these fields in term of normal variables are then:

$$\hat{\boldsymbol{E}}_\perp(\boldsymbol{r}, t) = i \sum_i \left(\frac{\hbar\omega_i}{2\epsilon_0 L^3}\right)^{1/2} \left[\hat{a}_i \boldsymbol{\varepsilon}_i e^{i(\boldsymbol{k}_i\cdot\boldsymbol{r}-\omega_i t)} - \hat{a}_i^\dagger \boldsymbol{\varepsilon}_i^* e^{-i(\boldsymbol{k}_i\cdot\boldsymbol{r}-\omega_i t)}\right] \tag{B-3}$$

$$\hat{\boldsymbol{B}}(\boldsymbol{r}, t) = i \sum_i \left(\frac{\hbar k_i}{2\epsilon_0 c L^3}\right)^{1/2} \left[\hat{a}_i \boldsymbol{\kappa}_i \times \boldsymbol{\varepsilon}_i e^{i(\boldsymbol{k}_i\cdot\boldsymbol{r}-\omega_i t)} - \hat{a}_i^\dagger \boldsymbol{\kappa}_i \times \boldsymbol{\varepsilon}_i^* e^{-i(\boldsymbol{k}_i\cdot\boldsymbol{r}-\omega_i t)}\right] \tag{B-4}$$

$$\hat{\boldsymbol{A}}_\perp(\boldsymbol{r}, t) = \sum_i \left(\frac{\hbar}{2\epsilon_0 \omega_i L^3}\right)^{1/2} \left[\hat{a}_i \boldsymbol{\varepsilon}_i e^{i(\boldsymbol{k}_i\cdot\boldsymbol{r}-\omega_i t)} + \hat{a}_i^\dagger \boldsymbol{\varepsilon}_i^* e^{-i(\boldsymbol{k}_i\cdot\boldsymbol{r}-\omega_i t)}\right] \tag{B-5}$$

$$\hat{H}_R = \sum_i \frac{\hbar\omega_i}{2} \left[\hat{a}_i^\dagger \hat{a}_i + \hat{a}_i \hat{a}_i^\dagger\right] = \sum_i \hbar\omega_i \left[\hat{a}_i^\dagger \hat{a}_i + \frac{1}{2}\right] \tag{B-6}$$

$$\hat{\boldsymbol{P}}_{\text{trans}} = \sum_i \frac{\hbar\boldsymbol{k}_i}{2} \left[\hat{a}_i^\dagger \hat{a}_i + \hat{a}_i \hat{a}_i^\dagger\right] = \sum_i \hbar\boldsymbol{k}_i\, \hat{a}_i^\dagger \hat{a}_i \tag{B-7}$$

Note that the last term in (B-7) does not contain the factor 1/2, since $\sum_i \boldsymbol{k}_i = \boldsymbol{0}$.

B-2. Corpuscular interpretation of states with fixed total energy and momentum

Consider first the mode i. The eigenvalues of the operator $\hat{a}_i^\dagger \hat{a}_i$ appearing in expressions (B-6) and (B-7) for $\hat{H}_R$ and $\hat{\boldsymbol{P}}_{\text{trans}}$ are all the positive or zero integers n_i :

$$\hat{a}_i^\dagger \hat{a}_i |n_i\rangle = n_i |n_i\rangle, \quad n_i = 0, 1, 2... \tag{B-8}$$

[4] For each $\boldsymbol{k}_i$, there exists two polarization vectors $\boldsymbol{\varepsilon}_{i1}$ and $\boldsymbol{\varepsilon}_{i2}$ perpendicular to $\boldsymbol{k}_i$ and perpendicular to each other. The compact notation $\sum_i$ must be interpreted as a summation over $\boldsymbol{k}_i$ and, for each value of $\boldsymbol{k}_i$, as a sum over $\boldsymbol{\varepsilon}_{i1}$ and $\boldsymbol{\varepsilon}_{i2}$.

Remember the well-known actions of operators $\hat{a}_i^\dagger$ and $\hat{a}_i$ on the states $|n_i\rangle$:

$$\begin{aligned}\hat{a}_i^\dagger|n_i\rangle &= \sqrt{n_i+1}|n_i\rangle \\ \hat{a}_i|n_i\rangle &= \sqrt{n_i}|n_i-1\rangle \\ \hat{a}_i|0_i\rangle &= 0\end{aligned} \tag{B-9}$$

As $\hat{a}_i^\dagger\hat{a}_i$ commutes with $\hat{a}_j^\dagger\hat{a}_j$, the eigenstates of $\hat{H}_R$ and $\hat{\boldsymbol{P}}_{\text{trans}}$ are the tensor products of the eigenstates $|n_1...n_i...\rangle = |n_1\rangle\otimes...\otimes|n_i\rangle...$ of the creation and annihilation operators $\hat{a}_1^\dagger\hat{a}_1$,....$\hat{a}_i^\dagger\hat{a}_i$ ...:

$$\hat{H}_R|n_1...n_i...\rangle = \sum_i \left(n_i+\frac{1}{2}\right)\hbar\omega_i|n_1...n_i...\rangle \tag{B-10a}$$

$$\hat{\boldsymbol{P}}_{\text{trans}}|n_1...n_i...\rangle = \sum_i n_i\hbar\boldsymbol{k}_i|n_1...n_i...\rangle \tag{B-10b}$$

The field's ground state corresponds to all the n_i equal to zero, and will be noted $|0\rangle$:

$$|0\rangle = |0_1...0_i...\rangle \tag{B-11}$$

the states $|n_1...n_i...\rangle$ being obtained by the action of a certain number of creation operators on this $|0\rangle$ state:

$$|n_1...n_i...\rangle = \frac{(\hat{a}_1^\dagger)^{n_1}}{\sqrt{n_1!}}...\frac{(\hat{a}_i^\dagger)^{n_i}}{\sqrt{n_i!}}...|0\rangle \tag{B-12}$$

With respect to the field ground state, the state $|n_1...n_i...\rangle$ has an energy $\sum_i n_i\hbar\omega_i$ and a momentum $\sum_i n_i\hbar\boldsymbol{k}_i$. It can be interpreted as describing an ensemble of n_1 particles of energy $\hbar\omega_1$ and momentum $\hbar\boldsymbol{k}_1$,....., n_i particles of energy $\hbar\omega_i$ and momentum $\hbar\boldsymbol{k}_i$, *etc...* . These particles characterize the elementary excitations of the quantum field and are called *photons*. The quantum number n_i is therefore the number of photons occupying the mode i, so that the ground state $|0\rangle$, corresponding to all the n_i equal to zero, can be called the photon *vacuum*.

Whereas there exists for photons eigenstates of momentum and energy, there are no quantum states of the electromagnetic field where the position can be perfectly known; no position operator is associated with this field. This is a different situation from what we encounter with massive particles, which have both a position and a momentum operator; the wave functions in the two representations are related by a simple Fourier transform. This non-existence of a position operator is linked to the impossibility of building, by a linear superposition of transverse electromagnetic waves, a vector wave perfectly localized at a point in space. The relativistic and transverse character of the electromagnetic field yields commutation relations between its components that involve the transverse delta function (Complement A_{XVIII}, § 2-e) instead of the usual delta function.

B-3. Several examples of quantum radiation states

We now study several examples of states of quantum radiation.

B-3-a. Photon vacuum

The presence of the 1/2 term in the parenthesis on the right-hand side of equation (B-10a) shows that the vacuum state energy is not zero, but equal to $\sum_i \hbar\omega_i/2$; this sum is an infinite quantity. We encounter here a first example of the difficulties linked to the divergences appearing in quantum electrodynamics. They can be resolved by *renormalization* techniques, whose presentation is outside the scope of this book. We shall avoid this difficulty by only considering energy differences with respect to the vacuum.

If we consider a single mode i of the field, the energy $\hbar\omega_i/2$ of the vacuum state for this mode is finite, and reminiscent of the zero-point energy of a harmonic oscillator of frequency ω_i. As you may recall, this zero-point energy is due to the impossibility of having simultaneously zero values for the position x and momentum p of that oscillator, because of the Heisenberg relations. The lowest energy state of the oscillator results from a compromise between the kinetic energy, proportional to p^2, and the potential energy, proportional to x^2 (this problem is discussed in § D-2 of Chapter V). The same arguments can be presented for the contribution, at a given point $\boldsymbol{r}$, of mode i to the electric $\hat{\boldsymbol{E}}_\perp(\boldsymbol{r},t)$ and magnetic $\hat{\boldsymbol{B}}(\boldsymbol{r},t)$ fields; according to (B-3) and (B-4), those fields are represented by two different linear superpositions of operators $\hat{a}_i$ and $\hat{a}_i^\dagger$, which thus do not commute. Consequently, one cannot have simultaneously a zero value for the electric energy proportional to $\hat{\boldsymbol{E}}_\perp^2$, and for the magnetic energy proportional to $\hat{\boldsymbol{B}}^2$.

One can further calculate the average value and variance of the contribution of mode i to the electric field $\hat{\boldsymbol{E}}_\perp(\mathbf{r},t)$ at point $\boldsymbol{r}$. Since $\hat{a}_i$ and $\hat{a}_i^\dagger$ change n_i by ± 1, a simple calculation yields:

$$\langle 0|\hat{\boldsymbol{E}}_\perp(\boldsymbol{r},t)|0\rangle_{\text{mode i}} = 0 \tag{B-13a}$$

$$\langle 0|\hat{\boldsymbol{E}}_\perp^2(\boldsymbol{r},t)|0\rangle_{\text{mode i}} = \frac{\hbar\omega_i}{2\varepsilon_0 L^3} \tag{B-13b}$$

Similar calculations can be done for the magnetic field. They show that in the photon vacuum state, the average value of both the electric and magnetic fields is zero, but not their variance. Since result (B-13b) is proportional to $\hbar$, the non-zero variance of the fields in the vacuum is a quantum effect.

Comments

(i) The summation over all the modes of expressions (B-13) yields, once we have replaced the discrete sum by an integral:

$$\langle 0|\hat{\boldsymbol{E}}_\perp(\boldsymbol{r},t)|0\rangle = 0 \tag{B-14a}$$

$$\langle 0|\hat{\boldsymbol{E}}_\perp^2(\boldsymbol{r},t)|0\rangle = \sum_i \frac{\hbar\omega_i}{2\varepsilon_0 L^3} = \frac{\hbar c}{2\varepsilon_0\pi^2}\int_0^{k_M} k^3 \mathrm{d}k \tag{B-14b}$$

This means that the variance of the electric field diverges as the fourth power of the upper boundary of the integral over k appearing in the summation of the modes of frequency $\omega = ck$. This divergence is the same as that mentioned above.

(ii) To characterize the dynamics of these field fluctuations, it is possible to compute the field correlation functions in vacuum[5]. This calculation shows that the electric and magnetic fields fluctuate very rapidly around their zero average value. These fluctuations are called the *vacuum fluctuations*. Certain radiative corrections, such as the "Lamb shift"

[5] See for example § III-C-3-c and Complement C_{III} of reference [16].

in atoms, can be interpreted from a physical point of view, as resulting from the vibration of the atom's electron caused by its interaction with this fluctuating electric field. This vibration leads the electron to explore the nucleus Coulomb potential over the range of its vibrational motion. The corresponding correction to its binding energy depends on the energy level it occupies; this explains why the degeneracy between the $2s_{1/2}$ and $2p_{1/2}$ states of the hydrogen atom, predicted by the Schrödinger and Dirac equations, can be lifted by the interaction with the vacuum fluctuations [6].

B-3-b. Field quasi-classical states

The state and observables of a classical field are characterized by the normal variables $\{\alpha_i\}$ introduced in § B-2-b of Chapter XVIII. The coherent states of a one-dimensional harmonic oscillator studied in Complement G_V, can be used to build the field quantum states whose properties are closest to those of the classical field $\{\alpha_i\}$.

The coherent state, supposed to be normalized, of a one-dimensional harmonic oscillator is the eigenstate of the annihilation operator $\hat{a}$, with eigenvalue α:

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle \tag{B-15}$$

The eigenvalue α may be a complex number since operator $\hat{a}$ is not Hermitian. Equation (B-15) leads to:

$$\langle\alpha|\hat{a}|\alpha\rangle = \alpha \qquad \langle\alpha|\hat{a}^\dagger|\alpha\rangle = \alpha^* \tag{B-16}$$

More generally, the average value of any function of $\hat{a}$ and $\hat{a}^\dagger$, once put in the *normal* order, i.e. where all the annihilation operators are positioned to the right of the creation operators (Complement B_{XVI}, § 1-a-α), is equal to the expression obtained by replacing operator $\hat{a}$ by α and operator $\hat{a}^\dagger$ by α^*. As an example:

$$\langle\alpha|\hat{a}^\dagger\hat{a}|\alpha\rangle = \alpha^*\alpha \tag{B-17}$$

Consider then the field quantum state:

$$|\alpha_1\rangle \otimes |\alpha_2\rangle \otimes \ldots |\alpha_i\rangle \ldots = |\alpha_1, \alpha_2 \ldots \alpha_i \ldots\rangle \tag{B-18}$$

where each mode i is in the coherent state $|\alpha_i\rangle$ corresponding to the classical normal variable α_i. Using equations (B-16) and (B-17), we can obtain the average values of the various field operators (B-3), (B-4) and (B-5) in the state (B-18); they coincide with the values of these various physical quantities for a classical field described by the normal variables $\{\alpha_i\}$. The same is true for the observables (B-6) and (B-7) corresponding to the energy and momentum of the transverse field. This is why the quantum state (B-18), which yields average values identical to all the properties of a classical field, is called a *quasi-classical state*[7]. We shall see later that the correlation functions of the quantum and classical fields involved in various photodetection signals also coincide when the field state is a quasi-classical state.

[6] See for example [17].

[7] For more details on the properties of the radiation quasi-classical states, see § III-C-4 of reference [16].

B-3-c. Single photon state

Consider the state vector:

$$|\Psi\rangle = \sum_i c_i |1_i\rangle \prod_{\otimes j \neq i} |0_j\rangle \tag{B-19}$$

which is a linear superposition of kets where a mode i contains one photon, whereas all the other modes $j \neq i$ are empty. Such a ket is an eigenket of the operator total number of photons $\hat{N} = \sum_i \hat{a}_i^\dagger \hat{a}_i$ with an eigenvalue equal to 1. It is therefore a single photon state. However, except in special cases, it is not a stationary state since it is not an eigenstate of the field energy $\hat{H}_R$. It describes a single photon propagating in space with velocity c. We shall see later (Complement $\mathrm{D_{XX}}$) that, when the field is in the state (B-19), a photodetector placed in a small region of space yields a signal corresponding to the passage, in that region, of a wave packet.

C. Description of the interactions

C-1. Interaction Hamiltonian

The Hamiltonian $\hat{H}$ of the system "particles + field" has been given above. In its expression (A-10), we now separate the terms that depend only on the particle variables or only on the field variables, and those that depend on both. We can then write $\hat{H} = \hat{H}_P + \hat{H}_R + \hat{H}_I$, where the particle Hamiltonian is:

$$\hat{H}_P = \sum_a \frac{\hat{\boldsymbol{p}}_a^2}{2m_a} + \hat{V}_{\mathrm{Coul}} \tag{C-1}$$

whereas the radiation one is:

$$\hat{H}_R = \sum_i \hbar\omega_i \left(\hat{a}_i^\dagger \hat{a}_i + \frac{1}{2} \right) \tag{C-2}$$

Finally, the interaction Hamiltonian is the sum:

$$\hat{H}_I = \hat{H}_{I1} + \hat{H}_{I2}, \tag{C-3}$$

with:

$$\hat{H}_{I1} = -\sum_a \frac{q_a}{2m_a} \left[\hat{\boldsymbol{p}}_a \cdot \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a) + \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a) \cdot \hat{\boldsymbol{p}}_a \right] \tag{C-4}$$

$$\hat{H}_{I2} = \sum_a \frac{q_a^2}{2m_a} \left[\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_a) \right]^2 \tag{C-5}$$

(we have separated the linear and quadratic terms with respect to the fields).

To that interaction Hamiltonian, we must further add the term:

$$\hat{H}_{I1}^s = -\sum_a \hat{\boldsymbol{M}}_a^s \cdot \hat{\boldsymbol{B}}(\hat{\boldsymbol{r}}_a) \tag{C-6}$$

describing the interaction of the spin magnetic moments of the various particles with the magnetic field of the radiation (Complement A_{XIII}, § 1-d):

$$\hat{\boldsymbol{M}}_a^s = g_a \frac{q_a}{2m_a} \hat{\boldsymbol{S}}_a \tag{C-7}$$

where g_a is the "Landé g-factor" of particle a whose spin is noted $\hat{\boldsymbol{S}}_a$.

Comment

Even with this additional term, all the possible interactions are not contained in that Hamiltonian: missing for example are the electron spin-orbit coupling, the hyperfine interaction between the electron and the nucleus, etc. – see comment (iii) of § C-5. The Hamiltonian we wrote is however sufficient in a great number of cases.

C-2. Interaction with an atom. External and internal variables

Consider the case where the particle system is a single atom, assumed to be neutral, formed by an electron e and a nucleus N which have opposite charges ($q_e = -q_N = q$) and whose masses are noted m_e and m_N. This is the case for example of the hydrogen atom. It is standard practice (see for example § B of Chapter VII) to separate the variables $\hat{\boldsymbol{R}}$ and $\hat{\boldsymbol{P}}$ of the system's center of mass and the variables $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{p}}$ of the relative motion. These two types of variables commute with each other and are given by equations:

$$\begin{cases} \boldsymbol{R} = \dfrac{m_e\, \boldsymbol{r}_e + m_N\, \boldsymbol{r}_N}{M} \\ \boldsymbol{r} = \boldsymbol{r}_e - \boldsymbol{r}_N \end{cases} \qquad \begin{cases} \boldsymbol{P} = \boldsymbol{p}_e + \boldsymbol{p}_N \\ \dfrac{\boldsymbol{p}}{m} = \dfrac{\boldsymbol{p}_e}{m_e} - \dfrac{\boldsymbol{p}_N}{m_N} \end{cases} \tag{C-8}$$

where we have noted M the total mass of the system, and m its reduced mass:

$$M = m_e + m_N \qquad ; \qquad m = \frac{m_e m_N}{M} \tag{C-9}$$

Expressed as a function of these new variables, the particle Hamiltonian is written:

$$\hat{H}_P = \frac{\hat{\boldsymbol{P}}^2}{2M} + \frac{\hat{\boldsymbol{p}}^2}{2m} + \hat{V}_{\text{Coul}}(\hat{r}) \tag{C-10}$$

The center of mass variables, also called *external variables*, describe the global motion of the atom, whereas the variables of the relative motion, also called *internal variables*, describe the motion in the center of mass reference frame.

C-3. Long wavelength approximation

The interaction Hamiltonians (C-4), (C-5) and (C-6) contain fields evaluated at the electron $\boldsymbol{r}_e$ and nucleus $\boldsymbol{r}_N$ positions. These positions can be described with respect to the position of the center of mass and we can write for example:

$$\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_e) = \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}} + \hat{\boldsymbol{r}}_e - \hat{\boldsymbol{R}}) \tag{C-11}$$

In an atom, the distance between the position of the electron or the nucleus and the atom's center of mass is of the order of the atom's size, i.e. just a fraction of a nanometer. Now

the radiation wavelengths that can have a resonant interaction with the atom are of the order of a fraction of a micron, much larger than the atomic dimension. One can thus neglect the variation of the fields over distances of the order of $|\boldsymbol{r}_e - \boldsymbol{R}|$ (or $|\boldsymbol{r}_N - \boldsymbol{R}|$) and write:

$$\begin{aligned} \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_e) &\simeq \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}}) \\ \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_N) &\simeq \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}}) \end{aligned} \tag{C-12}$$

Such an approximation is called the long wavelength approximation (or dipole approximation).

Using this approximation in the interaction Hamiltonian $\hat{H}_{I1}$, yields:

$$\begin{aligned} \hat{H}_{I1} &= -\frac{q_e}{m_e}\hat{\boldsymbol{p}}_e \cdot \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_e) - \frac{q_N}{m_N}\hat{\boldsymbol{p}}_N \cdot \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{r}}_N) \\ &\simeq -q\left[\frac{\hat{\boldsymbol{p}}_e}{m_e} - \frac{\hat{\boldsymbol{p}}_N}{m_N}\right] \cdot \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}}) \\ &= -\frac{q}{m}\hat{\boldsymbol{p}} \cdot \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}}) \end{aligned} \tag{C-13}$$

We used the relation $q_e = -q_N = q$ as well as definition (C-8) for the relative momentum.

As for Hamiltonian $\hat{H}_{I2}$, it becomes with this approximation:

$$\begin{aligned} \hat{H}_{I2} &= \frac{q_e^2}{2m_e}\hat{\boldsymbol{A}}_\perp^2(\hat{\boldsymbol{r}}_e) + \frac{q_N^2}{2m_N}\hat{\boldsymbol{A}}_\perp^2(\hat{\boldsymbol{r}}_N) \\ &\simeq \frac{q^2}{2m}\hat{\boldsymbol{A}}_\perp^2(\hat{\boldsymbol{R}}) \end{aligned} \tag{C-14}$$

Comment

When we include the Hamiltonian describing the spin magnetic coupling $\hat{H}_{I1}^s$ written in (C-6), we also replace all the $\hat{\boldsymbol{r}}_a$ by $\hat{\boldsymbol{R}}$. This is however insufficient: we must add other terms of the same order, obtained by including first order terms in $\boldsymbol{k}\cdot(\hat{\boldsymbol{r}}_e - \hat{\boldsymbol{R}})$ in $\hat{H}_{I1}$ and $\hat{H}_{I2}$, and representing corrections to the long wavelength approximation. This is because a computation analogous to the one in § 1-d of Complement AXIII shows that these corrections yield new interaction terms of the same order as $\hat{H}_{I1}^s$: interaction between the atomic orbital momentum $\boldsymbol{L}$ and the radiation magnetic field; electric quadrupole interaction.

C-4. Electric dipole Hamiltonian

Using the long wavelength approximation, the global Hamiltonian for the system atom + field is written:

$$\hat{H} = \frac{\hat{\boldsymbol{P}}^2}{2M} + \frac{1}{2m}\left[\hat{\boldsymbol{p}} - q\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}})\right]^2 + \hat{V}_{\text{Coul}} + \sum_i \frac{\hbar\omega_i}{2}\left[\hat{a}_i^\dagger \hat{a}_i + \hat{a}_i \hat{a}_i^\dagger\right] \tag{C-15}$$

We are going to perform a unitary transformation on this Hamiltonian, leading to a new interaction Hamiltonian, composed of a single term of the form $-\hat{\boldsymbol{D}} \cdot \hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})$, where $\hat{\boldsymbol{D}}$ is the electric dipole moment of the atom:

$$\hat{\boldsymbol{D}} = q\,\hat{\boldsymbol{r}} \tag{C-16}$$

and $\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})$, the quantum field given by expression (B-3). This new interaction Hamiltonian is called the *electric dipole Hamiltonian*.

To find this unitary transformation, it is useful to start with the simpler case where the radiation field is treated classically.

C-4-a. Electric dipole Hamiltonian for a classical field

When the radiation field is treated classically, as an external field whose dynamic is externally imposed and hence has a fixed time dependence, the last term of relation (C-15) does not exist; operator $\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}})$, which appears in the second term, must be replaced by the external field $\boldsymbol{A}_{\perp e}(\hat{\boldsymbol{R}}, t)$. The system Hamiltonian is then written:

$$\hat{H} = \frac{\hat{\boldsymbol{P}}^2}{2M} + \frac{1}{2m}\left[\hat{\boldsymbol{p}} - q\boldsymbol{A}_{\perp e}(\hat{\boldsymbol{R}}, t)\right]^2 + \hat{V}_{\text{Coul}} \tag{C-17}$$

We are looking for a unitary transformation that performs a translation of $\hat{\boldsymbol{p}}$ by a quantity $q\boldsymbol{A}_{\perp e}(\hat{\boldsymbol{R}}, t)$, so that the second term in (C-17) is reduced to $\hat{\boldsymbol{p}}^2/2m$. Such a transformation reads:

$$\hat{T}(t) = \exp\left[-\frac{i}{\hbar} q\,\hat{\boldsymbol{r}} \cdot \boldsymbol{A}_{\perp e}(\hat{\boldsymbol{R}}, t)\right] \tag{C-18}$$

We can check this since, using $[\hat{\boldsymbol{p}}, f(\hat{\boldsymbol{r}})] = -i\hbar\,\partial f/\partial\hat{\boldsymbol{r}}$ and the fact that the internal variable $\hat{\boldsymbol{r}}$ commutes with the external variable $\hat{\boldsymbol{R}}$, we have:

$$\hat{T}(t)\,\hat{\boldsymbol{p}}\,\hat{T}^\dagger(t) = \hat{\boldsymbol{p}} + q\boldsymbol{A}_{\perp e}(\hat{\boldsymbol{R}}, t) \tag{C-19}$$

As they do not depend on $\hat{\boldsymbol{p}}$, the other terms of (C-17) are unchanged by the transformation. On the other hand, since this transformation has an explicit time dependence via the term $q\boldsymbol{A}_{\perp e}(\hat{\boldsymbol{R}}, t)$, the new Hamiltonian H' that governs the evolution of the new state vector:

$$|\Psi'(t)\rangle = \hat{T}(t)|\Psi(t)\rangle \tag{C-20}$$

is given by:

$$\hat{H}'(t) = \hat{T}(t)\hat{H}(t)\hat{T}^\dagger(t) + i\hbar\left[\frac{\mathrm{d}\hat{T}(t)}{\mathrm{d}t}\right]\hat{T}^\dagger(t) \tag{C-21}$$

As we have in addition:

$$i\hbar\left[\frac{\mathrm{d}\hat{T}(t)}{\mathrm{d}t}\right]\hat{T}^\dagger(t) = q\hat{\boldsymbol{r}} \cdot \frac{\partial \boldsymbol{A}_{\perp e}(\hat{\boldsymbol{R}}, t)}{\partial t} = -\hat{\boldsymbol{D}} \cdot \boldsymbol{E}_{\perp e}(\hat{\boldsymbol{R}}, t) \tag{C-22}$$

where $\hat{\boldsymbol{D}} = q\,\hat{\boldsymbol{r}}$ is the electric dipole moment of the atom, we finally obtain:

$$\hat{H}'(t) = \frac{\hat{\boldsymbol{P}}^2}{2M} + \frac{\hat{\boldsymbol{p}}^2}{2M} + V_{\text{Coul}} - \hat{\boldsymbol{D}} \cdot \boldsymbol{E}_{\perp e}(\hat{\boldsymbol{R}}, t) \tag{C-23}$$

where the last term has the expected form for an electric dipole Hamiltonian.

C-4-b. Electric dipole Hamiltonian for a quantum field

The results we just obtained suggest using the unitary transformation:

$$\hat{T} = \exp\left[-\frac{i}{\hbar} q\, \hat{\boldsymbol{r}} \cdot \hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}}) \right] \tag{C-24}$$

where it is now the operator $\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}})$ that appears in the exponential. One can check that this operator is still a translation operator for $\hat{\boldsymbol{p}}$, so that the second term in (C-15) is now simply of the form $\hat{\boldsymbol{p}}^2/2m$.

As the transformation (C-24) no longer has an explicit time dependence, the term analogous to (C-22) does not exist anymore. On the other hand, we must study the transformation of the last term of (C-15), which represents the energy $\hat{H}_R$ of the transverse quantum field. We therefore rewrite expression (C-24) using the expansion (B-5) of $\hat{\boldsymbol{A}}_\perp(\hat{\boldsymbol{R}})$ as a function of $\hat{a}_i$ and $\hat{a}_i^\dagger$:

$$\hat{T} = \exp\left[\sum_i \left(\lambda_i^* \hat{a}_i - \lambda_i \hat{a}_i^\dagger \right) \right] \tag{C-25}$$

with:

$$\lambda_i = \frac{i}{\sqrt{2\varepsilon_0 \hbar \omega_i L^3}}\, \boldsymbol{\varepsilon}_i \cdot \hat{\boldsymbol{D}}\; e^{-i\boldsymbol{k}_i \cdot \hat{\boldsymbol{R}}} \tag{C-26}$$

In this form, operator $\hat{T}$ does appear as a translation operator (Complement $\mathrm{G_V}$, § 2-d); it obeys the equations:

$$\hat{T}\, \hat{a}_j\, \hat{T}^\dagger = \hat{a}_j + \lambda_j \qquad \hat{T}\, \hat{a}_j^\dagger\, \hat{T}^\dagger = \hat{a}_j^\dagger + \lambda_j^* \tag{C-27}$$

To prove relations (C-27), one can use (Complement $\mathrm{B_{II}}$ $\mathrm{B_{II}}$, § 5-d) the identity:

$$e^{(A+B)} = e^A e^B e^{-[A,B]/2} \tag{C-28}$$

valid if A and B commute with their commutator $[A, B]$, as well as the commutation relation $\left[\hat{a}, f(\hat{a}^\dagger)\right] = \partial f / \partial \hat{a}^\dagger$. The transformation of the last term in (C-15) then yields:

$$\hat{T} \hat{H}_R \hat{T}^\dagger = \sum_i \frac{\hbar\omega_i}{2} \left[(\hat{a}_i + \lambda_i)(\hat{a}_i^\dagger + \lambda_i^*) + (\hat{a}_i^\dagger + \lambda_i^*)(\hat{a}_i + \lambda_i) \right] \tag{C-29}$$

The terms on the right-hand side of (C-29) that are independent of λ_i and λ_i^* yield again $\hat{H}_R$. The terms linear in λ_i and λ_i^* yield:

$$\begin{aligned} \sum_i \hbar\omega_i \left(\lambda_i \hat{a}_i^\dagger + \lambda_i^* \hat{a}_i \right) &= -\sum_i i \sqrt{\frac{\hbar\omega_i}{2\epsilon_0 L^3}} \left(\hat{a}_i \boldsymbol{\varepsilon}_i e^{i\boldsymbol{k}_i \cdot \hat{\boldsymbol{R}}} - \hat{a}_i^\dagger \boldsymbol{\varepsilon}_i^* e^{-i\boldsymbol{k}_i \cdot \hat{\boldsymbol{R}}} \right) \cdot \hat{\boldsymbol{D}} \\ &= -\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}}) \cdot \hat{\boldsymbol{D}} \end{aligned} \tag{C-30}$$

where we have used (B-3). We thus get the expected electric dipole form for the interaction Hamiltonian:

$$\hat{H}_I = -\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}}) \cdot \hat{\boldsymbol{D}} \tag{C-31}$$

Finally, the terms quadratic in λ_i and λ_i^* introduce a term we shall note $\hat{h}_{\text{dip}}$:

$$\hat{h}_{\text{dip}} = \sum_i \hbar\omega_i \lambda_i^* \lambda_i = \sum_i \frac{1}{2\varepsilon_0 L^3} (\boldsymbol{\varepsilon}_i \cdot \hat{\boldsymbol{D}})(\boldsymbol{\varepsilon}_i^* \cdot \hat{\boldsymbol{D}}) \tag{C-32}$$

It represents a dipolar energy intrinsic to the atom.

To sum up, regrouping all the previous terms, we get for the transformed Hamiltonian:

$$\hat{H}' = \frac{\hat{\boldsymbol{P}}^2}{2M} + \frac{\hat{\boldsymbol{p}}^2}{2M} + V_{\text{Coul}} + \hat{H}_R - \hat{\boldsymbol{D}} \cdot \hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}}) + \hat{h}_{\text{dip}} \tag{C-33}$$

This is a form similar to (C-23), with an additional term $\hat{h}_{\text{dip}}$.

Comments

(i) The same mathematical operator does not describe the same physical quantity in two different representations, deduced from one another by a unitary transformation. As an example, the operator $\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})$ appearing in (C-31), does not represent the transverse electric field in the new point of view, which should be $\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})$ transformed by $\hat{T}$, written as $\hat{T}\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})\hat{T}^\dagger$, and hence different from $\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})$. Actually, one can show that the operator $\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})$ represents in the new point of view the physical quantity $\hat{\mathfrak{D}}(\hat{\boldsymbol{R}})/\varepsilon_0$ where $\hat{\mathfrak{D}}(\hat{\boldsymbol{R}})$ is called the *electric displacement field* (see Complement A_{IV} of [16]).

(ii) The intrinsic dipolar energy $\hat{h}_{\text{dip}}$ is given by an integral over k, which diverges at infinity. This integral must however be limited to values of k for which the long wavelength approximation is still valid.

C-5. Matrix elements of the interaction Hamiltonian; selection rules

Consider an initial state where the atom is described by $|\psi_{\text{in}}^{\text{int}}\rangle$ for its internal state, $|\psi_{\text{in}}^{\text{ext}}\rangle$ for its external state, and where the radiation is in the state $|\psi_{\text{in}}^R\rangle$. The interaction Hamiltonian (C-31) couples this initial state to a final state where the atomic internal and external variables, as well as the radiation variables are respectively in the states $|\psi_{\text{fin}}^{\text{int}}\rangle$, $|\psi_{\text{fin}}^{\text{ext}}\rangle$, and $|\psi_{\text{fin}}^R\rangle$. As the operator $\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})$ appearing in (C-31) is a linear superposition of annihilation $\hat{a}_i$ and creation $\hat{a}_i^\dagger$ operators, the matrix element of $\hat{H}_I$ describes two types of processes: the absorption processes associated with operator $\hat{a}_i$ where one photon disappears, and the emission processes associated with operator $\hat{a}_i^\dagger$ where a new photon appears. This matrix element can be factored into a product of three matrix elements concerning the three types of variables; they are written, for the absorption processes:

$$i\sqrt{\frac{\hbar\omega_i}{2\varepsilon_0 L^3}} \left\langle \psi_{\text{fin}}^{\text{int}} \right| \boldsymbol{\varepsilon}_i \cdot \hat{\boldsymbol{D}} \left| \psi_{\text{in}}^{\text{int}} \right\rangle \left\langle \psi_{\text{fin}}^{\text{ext}} \right| \exp(i\,\boldsymbol{k}_i \cdot \hat{\boldsymbol{R}}) \left| \psi_{\text{in}}^{\text{ext}} \right\rangle \left\langle \psi_{\text{fin}}^R \right| \hat{a}_i \left| \psi_{\text{in}}^R \right\rangle \tag{C-34}$$

and for the emission processes:

$$-i\sqrt{\frac{\hbar\omega_i}{2\varepsilon_0 L^3}} \left\langle \psi_{\text{fin}}^{\text{int}} \right| \boldsymbol{\varepsilon}_i^* \cdot \hat{\boldsymbol{D}} \left| \psi_{\text{in}}^{\text{int}} \right\rangle \left\langle \psi_{\text{fin}}^{\text{ext}} \right| \exp(-i\,\boldsymbol{k}_i \cdot \hat{\boldsymbol{R}}) \left| \psi_{\text{in}}^{\text{ext}} \right\rangle \left\langle \psi_{\text{fin}}^R \right| \hat{a}_i^\dagger \left| \psi_{\text{in}}^R \right\rangle \tag{C-35}$$

The central term in these expressions is a matrix element concerning the external atomic variables; it expresses the conservation of the global momentum as we now

show. Operator $\exp(\pm i\,\boldsymbol{k}_i \cdot \hat{\boldsymbol{R}})$ translates the momentum by a quantity $\pm\hbar\boldsymbol{k}_i$. If the atom's center of mass has an initial momentum $\hbar\boldsymbol{K}_{\text{in}}$, once it absorbs a photon, its final momentum will be $\hbar\boldsymbol{K}_{\text{fin}} = \hbar\boldsymbol{K}_{\text{in}} + \hbar\boldsymbol{k}_i$; the momentum $\hbar\boldsymbol{k}_i$ of the absorbed photon is therefore transferred to the atom during the absorption process. In a similar way, one can show that the atom's momentum decreases by the quantity $\hbar\boldsymbol{k}_i$ when a photon is emitted.

In the first matrix element of (C-34), which concerns the internal atomic variables, operator $\hat{\boldsymbol{D}}$ is an odd operator. The matrix element will be different from zero only if the initial and final internal atomic states have opposite parity, as for instance the $1s$ and $2p$ states of the hydrogen atom. We rediscover here a second conservation law, the conservation of parity. In addition, as the operator $\hat{\boldsymbol{D}}$ is a vector operator, it leads to selection rules on the internal angular momentum which will be studied in Complement C_{XIX}.

Comments

(i) The conservation of the total momentum comes from the central matrix elements in expressions (C-34) and (C-35). One may wonder whether this result is only valid for the approximate form (C-31) of the interaction Hamiltonian used to establish those equations. Actually it can be shown, using the commutation relations $[\boldsymbol{p}_a, F(\boldsymbol{r}_a)] = -i\hbar\partial F/\partial \boldsymbol{r}_a$ and $[\hat{a}_i^\dagger \hat{a}_i, \hat{a}_i] = -\hat{a}_i$, that the interaction Hamiltonian $\hat{H}_{I1}$ written is (C-4) (without the long wavelength approximation) commutes with the system total momentum $\sum_a \hat{\boldsymbol{p}}_a + \sum_i \hbar\boldsymbol{k}_i \hat{a}_i^\dagger \hat{a}_i$. The same result is true for all the terms of the interaction Hamiltonian. Consequently, the exact (without approximation) interaction Hamiltonian has non-zero matrix elements only between states having the same total momentum. The fact that the total momentum commutes with all the terms in the Hamiltonian is related to the system invariance with respect to spatial translation. The properties of the system are unchanged upon the translation by the same quantity of the particles and the fields. Similar considerations apply to the rotational invariance and cause the interaction Hamiltonian to only connect states with the same total angular momentum. These results are important for understanding in a simple fashion the exchanges of linear and angular momenta between atoms and photons, which will be discussed in Complements A_{XIX} and C_{XIX}.

(ii) Conservation of total momentum during the absorption process, combined with total energy conservation, shows that the energy of the absorbed photon is different from the energy separating the two internal levels involved in the transition. Two effects account for this difference: the Doppler effect, and the recoil effect (Complement A_{XIX}); they play an important role in laser cooling methods.

(iii) If we continue the calculations beyond the long wavelength approximation, we find additional terms for the interaction Hamiltonian, describing the interaction between the radiation magnetic field and the atomic orbital or spin magnetic moments (Complement A_{XIII}, § 1). Some of these terms have already been written in (C-6). Transitions, called magnetic dipole transitions, may occur between levels having the same parity, as opposed to the electric dipole transitions studied above. Other types of transitions may also be observed at higher orders, such as the quadrupole transitions.

Note finally that, if the initial radiation state already contains n_i photons, the last two matrix elements of (C-34) and (C-35) are equal to $\langle n_i - 1|\hat{a}_i|n_i\rangle = \sqrt{n_i}$ and $\langle n_i + 1|\hat{a}_i^\dagger|n_i\rangle = \sqrt{n_i + 1}$. In the presence of n_i incident photons, the probability of the absorption process is thus proportional to n_i, whereas the emission probability is

proportional to $n_i + 1$. We shall see in Chapter XX that this difference is linked to the existence of two types of emission, the stimulated emission and the spontaneous emission.

With the knowledge of the various Hamiltonians $\hat{H}_A$, $\hat{H}_R$ and $\hat{H}_I$, as well as their matrix elements, we can now solve Schrödinger's equation to compute the transition amplitude between an initial state and a final state of the system "atom + field". This will be done in the next chapter, where we study various processes, such as the absorption or emission of photons for an incident radiation either monochromatic or having a large spectral band, the photoionization phenomenon, multiphoton processes and photon scattering.

COMPLEMENTS OF CHAPTER XIX, READER'S GUIDE

The processes of photon absorption and emission by atoms must obey conservation laws for the total linear or angular momentum. This has implications that are highlighted in the three complements of this chapter. When an atom absorbs (or emits) a photon, it gains (or loses) an energy, a momentum and an angular momentum equal to the energy, momentum and angular momentum of the photon. This allows "manipulating" several properties of the atoms. It is, for example, at the base of optical pumping and laser cooling methods.

A_{XIX} : MOMENTUM EXCHANGE BETWEEN ATOMS AND PHOTONS	Momentum exchange between atoms and photons plays an important role in determining, for example, the Doppler effect and the shape of the spectral lines emitted or absorbed by gases. As an atom continually absorbs and re-emits photons, its momentum can be greatly affected. The atom can be decelerated, which allows slowing down and even bringing to rest an atomic beam over short distances. Other uses of the Doppler effect include the introduction of a friction force that slows down atoms to form ultra-cold gases. When the atoms are confined in a trap, the nature of the Doppler shift may be greatly changed and even completely disappear (Mössbauer effect). Multiphoton processes, in which the total momentum of the absorbed photons is zero, are also discussed.
B_{XIX} : ANGULAR MOMENTUM OF RADIATION	This complement is more technical than the previous one. It shows how the photon can be seen as a spin 1 particle; this particle also has an orbital angular momentum. The photon thus possesses two types of angular momentum. These concepts are useful for reading the next complement.
C_{XIX} : ANGULAR MOMENTUM EXCHANGE BETWEEN ATOMS AND PHOTONS	This complement studies the exchanges between the photon spin angular momentum (related to the light beam polarization) and the internal degrees of freedom of the atoms. These exchanges obey selection rules for the atomic transitions. Such rules are essential for many experimental methods in atomic physics, such as "optical pumping" where a polarized light beam allows accumulating the atoms of a gas in specific Zeeman sublevels. A number of applications of these methods are briefly reviewed in this complement.

Complement A_{XIX}

Momentum exchange between atoms and photons

1 Recoil of a free atom absorbing or emitting a photon 2020
1-a Conservation laws 2020
1-b Doppler effect, Doppler width 2022
1-c Recoil energy 2023
1-d Radiation pressure force in a plane wave 2024
2 Applications of the radiation pressure force: slowing and cooling atoms 2025
2-a Deceleration of an atomic beam 2025
2-b Doppler laser cooling of free atoms 2026
2-c Magneto-optical trap 2034
3 Blocking recoil through spatial confinement 2036
3-a Assumptions concerning the external trapping potential . . . 2036
3-b Intensities of the vibrational lines 2037
3-c Effect of the confinement on the absorption and emission spectra 2038
3-d Case of a one-dimensional harmonic potential 2039
3-e Mössbauer effect 2040
4 Recoil suppression in certain multi-photon processes 2040

We studied, in § C of Chapter XIX, the matrix elements of the interaction Hamiltonian between an atom and a field. This led us to establish selection rules based on the conservation of the total momentum of the system "atom + field" during the absorption or emission of photons by the atom. The study in Chapter XX of the absorption and emission amplitudes will show that the global energy of the system is also conserved during these processes. The goal of this complement is to show how these conservation laws[1] provide an interesting view on many aspects of momentum exchange between atoms and photons.

We start in § 1 with the case of a free atom, whose center of mass is not submitted to any external potential. We shall establish the expressions for the Doppler shift and the recoil energy that appear in the equation yielding the frequency of the absorbed or emitted photons. In a gas containing a large number of atoms with different velocities, the dispersion of these velocities yields a Doppler broadening of the emission and absorption lines. This broadening, as well as the shift related to the recoil energy, introduce perturbations in the lines observed using high resolution spectroscopy, hence limiting its precision. In addition, when the atom constantly absorbs and re-emits lots of photons, its momentum change per unit time can become very large. This results in a force generated

[1] The first study along this line concerned the Compton effect (1922), where the scattering of a photon by an electron is considered as a collision between two particles. Assigning the photon an energy $h\nu$ and a momentum $\hbar\boldsymbol{k}$, with $|\boldsymbol{k}| = 2\pi\nu/c$, one writes the conservation equations for the total momentum and energy during the collision. This yields the change of frequency of the photon as it is scattered in a given direction, in complete agreement with experimental observations.

by the radiation pressure. We shall calculate the order of magnitude of that force and show that it can produce an acceleration or deceleration of the atom a hundred thousand times larger than the one due to gravity.

In § 2, we will show that this force is able to slow down and immobilize a beam of atoms propagating in the direction opposed to that of the light beam. The velocity dependence of the radiation pressure force, due to the Doppler effect, is also very interesting. It allows, using two light beams in opposite directions, but with the same intensity and frequency, to generate a friction force on the atom, provided the light frequency is lower than the atomic frequency: the radiation pressure force is zero when the atomic velocity v is zero, but opposite to v when it is different from zero, therefore producing a damping of that velocity. This is the principle of one of the first laser cooling mechanisms observed experimentally. The very low temperatures obtained, millions of times lower than room temperature, explain the increasing number of application of the ultra-cold atoms thus obtained. We also explain in § 2-c the principle of the magneto-optical trap, which involves a position dependence of the radiation pressure force.

We describe in §§ 3 and 4 of this complement a number of methods developed to suppress or circumvent the effect related to the recoil. Confining the atom in a trap, as studied in § 3, may prevent the atom's recoil if the trapping is strong enough. If the transition is multi-photonic, for example if two photons having the same energy but opposite momenta are absorbed in the transition, no recoil is experienced by the atom, and there is no Doppler shift. An important example of this method is the Doppler-free two-photon spectroscopy (§ 4).

1. Recoil of a free atom absorbing or emitting a photon

Consider first an atom that is not subjected to any external potential (free atom). The Hamiltonian $\hat{H}_{\text{ext}}$ for the external variables is reduced to the kinetic energy term $\hat{\boldsymbol{P}}^2/2M$, where $\hat{\boldsymbol{P}}$ is the momentum of the center of mass and M the total mass of the atom. The eigenstates of $\hat{H}_{\text{ext}}$ can be chosen as states having well defined momentum $\boldsymbol{P}$ and energy $P^2/2M$.

1-a. Conservation laws

Let us express the radiation field as a function of plane waves of wave vector $\boldsymbol{k}$ and frequency $\omega = ck$. The eigenstates of the radiation Hamiltonian $\hat{H}_R$ can be described in terms of photons having an energy $\hbar\omega$ and a momentum $\hbar\boldsymbol{k}$. The interaction Hamiltonian $\hat{H}_I$ studied in § C of Chapter XIX is invariant[2] under spatial translation of the total system "atom + field". Consequently, it commutes with the system's total momentum, and can induce transitions only between states having the same total momentum. Furthermore, the transition amplitude associated with an interaction lasting a time T can only connect states of the total system having the same total energy, within $\hbar/T$ (this point will be further discussed in the next chapter). These two conservation laws can be used to study the influence of the motion of the center of mass on the frequencies of the photons it can absorb or emit.

[2]The interaction Hamiltonian involves field values at points where particles are located. It is therefore invariant when both fields and particles are shifted (by the same quantity).

Consider first an absorption process where the atom goes from an internal state a to another internal state b by absorbing a photon of energy $\hbar\omega$ and momentum $\hbar\boldsymbol{k}$. We shall note:

$$\hbar\omega_0 = E_b - E_a \tag{1}$$

the energy difference between the two internal states. The initial (final) momentum of the center of mass before (after) the photon absorption is noted $\boldsymbol{P}_{\rm in}$ ($\boldsymbol{P}_{\rm fin}$). The conservation of the total momentum leads to:

$$\boldsymbol{P}_{\rm fin} = \boldsymbol{P}_{\rm in} + \hbar\boldsymbol{k} \tag{2}$$

The total energy in the initial state is the sum of the photon energy, the internal energy of the atom and its translation energy $P_{\rm in}^2/2M$:

$$E_{\rm in} = \hbar\omega + E_a + \frac{P_{\rm in}^2}{2M} \tag{3}$$

The total energy in the final state reduces to the atom's energy since the photon has been absorbed:

$$E_{\rm fin} = E_b + \frac{P_{\rm fin}^2}{2M} \tag{4}$$

The conservation of the total energy means that $E_{\rm in} = E_{\rm fin}$. Using (1) and (2), this last relation reads:

$$\hbar\omega = \hbar\omega_0 + \frac{\hbar\boldsymbol{k}\cdot\boldsymbol{P}_{\rm in}}{M} + \frac{\hbar^2k^2}{2M} \tag{5}$$

The last two terms in (5) represent the variation of the external energy of the atom during the transition – i.e. the variation of the atom's center of mass kinetic energy between the final state where it is equal to $(\boldsymbol{P}_{\rm in} + \hbar\boldsymbol{k})^2/2M$, and the initial state, where it is equal to $\boldsymbol{P}_{\rm in}^2/2M$. This equation can be rewritten as:

$$\omega = \omega_0 + \boldsymbol{k}\cdot\boldsymbol{v}_{\rm in} + \frac{E_{\rm rec}}{\hbar} \tag{6}$$

where $\boldsymbol{v}_{\rm in} = \boldsymbol{P}_{\rm in}/M$ is the initial velocity of the center of mass, and where:

$$E_{\rm rec} = \frac{\hbar^2k^2}{2M} = \hbar\omega_{\rm R} \tag{7}$$

is the *recoil energy*; it is the energy an atom, initially at rest, will acquire upon the absorption of a photon having a momentum $\hbar k$.

The same type of calculation holds for the emission process during which an atom, whose center of mass has an initial momentum $\boldsymbol{P}_{\rm in}$, goes from the internal state b to the internal state a by emitting a photon of energy $\hbar\omega$ and momentum $\hbar\boldsymbol{k}$. Equation (6) must now be replaced by:

$$\omega = \omega_0 + \boldsymbol{k}\cdot\boldsymbol{v}_{\rm in} - \frac{E_{\rm rec}}{\hbar} \tag{8}$$

where only one sign is changed with respect to (6).

1-b. Doppler effect, Doppler width

The term $\boldsymbol{k} \cdot \boldsymbol{v}_{\text{in}}$ in equations (6) and (8) is simply the Doppler shift, due to the motion of the atom, of the frequencies it absorbs or emits. Setting:

$$\omega - \omega_0 = 2\pi(\nu - \nu_0) = 2\pi\Delta\nu \qquad \text{and} \qquad k = 2\pi\frac{\nu}{c} \tag{9}$$

we get for the frequency variation $\Delta\nu$ due to the Doppler effect:

$$\frac{\Delta\nu}{\nu} = \frac{\boldsymbol{\kappa} \cdot \boldsymbol{v}_{\text{in}}}{c} \tag{10}$$

where $\boldsymbol{\kappa} = \boldsymbol{k}/k$ is the unit vector defining the propagation direction of the photon. Note that a radiation quantum theory is not needed to account for this frequency shift, which can be predicted by a classical theory. This was to be expected since, among the last two terms of (6) and (8), $\boldsymbol{k} \cdot \boldsymbol{v}_{\text{in}}$ is the only one that does not go to zero when $\hbar$ tends toward zero, as opposed to $E_{\text{rec}}/\hbar = \hbar k^2/2M$ (which is proportional to $\hbar$).

For an ensemble of atoms in a dilute gas at thermal equilibrium at temperature T, the velocities are distributed according to the Maxwell-Boltzmann law, and the velocity dispersion Δv is of the order of $\sqrt{k_B T/M}$, where k_B is the Boltzmann constant. The shifts $\Delta\nu$ of the frequencies emitted or absorbed by the atoms are distributed following a Gaussian curve[3] whose width $\Delta\nu_D$ (standard deviation of the frequency distribution, equal to the square root of the variance), called the *Doppler width*, is given by:

$$\frac{\Delta\nu_D}{\nu_0} = \sqrt{\frac{k_B T}{Mc^2}} \tag{11}$$

In general, in the optical domain and for temperatures around 300 K, the Doppler width $\Delta\nu_D$ is of the order of 1GHz=10^9 Hz, much smaller than the frequency ν_0 (of the order of 10^{15} Hz), but much larger than the natural width Γ, of the order of 10^7 Hz. In this domain, the resolution of spectroscopic measurements of line frequencies emitted by a dilute gas is generally limited by the Doppler broadening of the lines.

Relativistic Doppler effect

The previous calculations are only valid in the non-relativistic limit ($v \ll c$). The Doppler shift expression can be generalized to any value of v by noting that the four quantities $\{k_x, k_y, k_z, \omega/c\}$ are the four components of a four-vector. Let us assume the atom is at rest in a reference frame R and emits a photon of frequency ω along the x axis (we ignore here the recoil energy). An observer, in a reference frame R' moving with velocity v along the x axis, sees the atom moving away at velocity $-v$ and measures a frequency ω' for the photon emitted by the atom. According to the relativistic expressions for the transformations of four-vector components, we have:

$$\omega' = \frac{\omega - kv}{\sqrt{1 - v^2/c^2}} \tag{12}$$

To first order in v/c, replacing $\boldsymbol{k} \cdot \boldsymbol{v}_{\text{in}}$ by $-kv$, we again find the Doppler shift of equation (8) – the relativistic correction being (in relative value) of the order of v^2/c^2 for $v \ll c$.

[3] Actually, this distribution is the convolution of a Gaussian and a curve of width Γ, where Γ is the natural width (due to the spontaneous emission) of the line emitted or absorbed by the atoms. For a gas at room temperature, Γ is much smaller than the Doppler width.

1-c. Recoil energy

Imagine the atom is initially at rest, so that the terms in $\boldsymbol{k} \cdot \boldsymbol{v}_{\text{in}}$ in (6) and (8) are zero. When the atom absorbs a photon, its momentum increases by a quantity $\hbar\boldsymbol{k}$ equal to the momentum of the absorbed photon. Consequently, the atom recoils with a velocity $v_{\text{rec}} = \hbar k/M$ in the direction of the incident photon, and its kinetic energy becomes $E_{\text{fin}} = Mv_{\text{rec}}^2/2 = E_{\text{rec}}$. The energy $\hbar\omega$ of the incident photon is used both to increase the internal energy of the atom by $\hbar\omega_0$ (since the atom goes from a to b), and to increase its kinetic energy from 0 to E_{rec}. We then have $\hbar\omega = \hbar\omega_0 + E_{\text{rec}}$, which is a particular case of (6) for a zero initial velocity. However, for the emission process where the atom goes from b to a by emitting a photon, this relation is modified. As the photon leaves the atom with a momentum $\hbar\boldsymbol{k}$, the atom recoils with the same momentum but in the opposite direction, and acquires a kinetic energy E_{rec}. The loss of internal energy of the atom, equal to $\hbar\omega_0$, must now be used both to increase the radiation energy by $\hbar\omega$ (the energy of the emitted photon), and to increase the kinetic energy of the atom from zero to E_{rec}. We then have $\hbar\omega_0 = \hbar\omega + E_{\text{rec}}$, that is $\hbar\omega = \hbar\omega_0 - E_{\text{rec}}$, which coincides with (8) with the minus sign on the right-hand side.

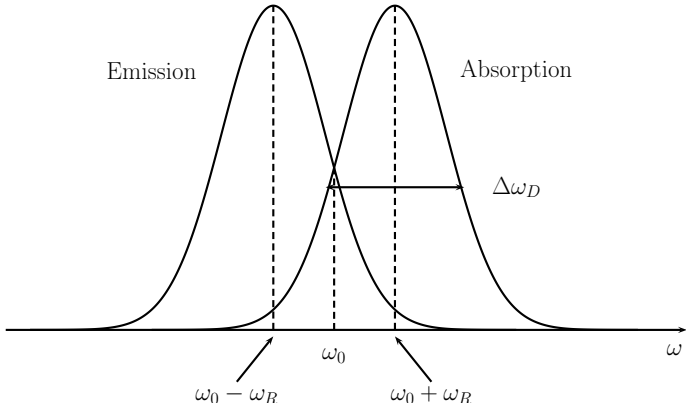


Figure 1: Because of the recoil effect of the atom, the absorption and emission lines do not coincide, but form a doublet called the recoil doublet*; they are centered at $\omega + \omega_R$ for the absorption line and $\omega - \omega_R$ for the emission line. In a gas, their width is the Doppler width $\Delta\omega_D$.*

The recoil of the atom with a kinetic energy E_{rec} causes the centers of the absorption and emission lines to be different (Figure 1); their position is $\omega_0 + \omega_R$ for the absorption line and $\omega_0 - \omega_R$ for the emission line, with $\omega_R = E_{\text{rec}}/\hbar$. For a gas containing a large number of moving atoms, the velocity dispersion around a zero average gives each of these two lines a Doppler width $\Delta\omega_D$. In the optical domain, the recoil frequency ω_R is of the order of a few kHz, much smaller than the Doppler width at room temperature, and also smaller than the natural width. Consequently the recoil doublet shown in Figure 1 is not resolved: the distance between the centers of the two lines is less than their width. However, when one studies the lines emitted by nuclei in the X-ray or γ-ray domains, the recoil frequency (which increases as k^2) becomes comparable to or

even larger than the Doppler width (which only increases as k). The two lines in Figure 1 then only overlap far out on their wings. In this case, the photon emitted by a nucleus in an excited state b has very little chance of being absorbed by another identical nucleus in the lower state a. We shall see later how the recoil of a nucleus can be blocked when the atom having that nucleus is sufficiently strongly bound to other atoms in a crystal (Mössbauer effect).

1-d. Radiation pressure force in a plane wave

Each time an atom absorbs a photon, it gains a momentum $\hbar\mathbf{k}$. If $\dot{N}_{\text{abs}}$ is the number of absorptions per unit time, the atom gains a momentum $\dot{N}_{\text{abs}}\hbar\mathbf{k}$, per unit time. In a steady state, the number $\dot{N}_{\text{abs}}$ of absorptions per unit time is equal to the number of emissions per unit time $\dot{N}_{\text{em}}$. This latter number is equal to $\Gamma\sigma_{bb}$, where Γ is the natural width of the atom's excited state, and where the diagonal element σ_{bb} of the atom's density operator in b is the occupation probability of that state. This gain in momentum per unit time of the atom can be considered as coming from the action of a force, associated with the radiation pressure exerted by the light beam on the atom. One often calls this force the "radiation pressure force". According to what we just saw, it is equal to[4]:

$$\mathbf{F} = \dot{N}_{\text{abs}}\hbar\mathbf{k} = \Gamma\sigma_{bb}\hbar\mathbf{k} \tag{13}$$

To get an idea of the order of magnitude of this force, let us assume the light intensity is very high; the atomic transition is then saturated, meaning the occupation probabilities σ_{bb} and σ_{aa} of the higher level b and the lower level a are both equal to $1/2$. We then have:

$$\mathbf{F} = \hbar\mathbf{k}\frac{\Gamma}{2} \tag{14}$$

Such a force can communicate an acceleration $\mathbf{A}$ equal to $\mathbf{F}/M$ to the atom of mass M. Taking (14) into account, this acceleration is equal to:

$$\mathbf{A} = \frac{\hbar\mathbf{k}}{M}\frac{\Gamma}{2} = \frac{\mathbf{v}_{\text{rec}}}{2\tau} \tag{15}$$

where $\mathbf{v}_{\text{rec}} = \hbar\mathbf{k}/M$ is the recoil velocity of an atom absorbing or emitting a photon, and $\tau = 1/\Gamma$ is the radiative lifetime of the excited state b.

Let us calculate the value of this acceleration for a sodium atom. The recoil velocity is of the order of 3×10^{-2}m/s and the radiative lifetime of the order of 16.2×10^{-9}s, meaning the acceleration A is of the order of 10^6m/s^2, which is 10^5 times larger (!) than the acceleration due to gravity (of the order of 10m/s^2). This high value for the acceleration A arises because the velocity change of the atom v_{rec} for each absorption-emission cycle, though very small by itself, accumulates during the very large number of cycles, $1/2\tau$, occurring in one second.

[4]To compute the momentum change, we only considered here the photon absorption processes. The spontaneous emission processes also change the atom's momentum, as the atom recoils when it emits a photon. However, the spontaneously emitted photons can have any direction in space, and the momentum change of the atom has a zero average. This phenomenon, however, gives rise to a diffusion of momentum, hence increasing the velocity dispersion of the atoms. We shall see in § 2-b-γ that this momentum diffusion must be taken into account when evaluating the limits of laser cooling.

2. Applications of the radiation pressure force: slowing and cooling atoms

We now study three applications of the radiation pressure force using either one laser beam, or two laser beams with the same intensity and same frequency, propagating in opposite directions along the x axis. We shall see in § 2-a how, with just a single laser beam, the radiation pressure force exerted by the beam on the atoms in an atomic beam propagating in the opposite direction can be used to slow down, and even bring to rest, the atoms of the beam. With two laser beams propagating in opposite directions, interesting effects occur if one can introduce an imbalance between the two radiation pressure forces, depending either on the atom's velocity v along the x axis, or on its position x. We study in § 2-b how to get this velocity dependence, and how the sum F of the two forces exerted by the two waves, zero for $v = 0$, becomes, for $v \neq 0$ but sufficiently small, a linear function of v; it can then be expressed as $F \simeq \gamma v$. For a proper detuning of the lasers' frequency, the coefficient γ can be negative, so that the resulting force F acts as a friction force that damps the velocities of all the atoms in the beam, and hence cools them down. This is the principle of laser cooling. In § 2-c, we shall see how a position dependence can be obtained, and how the resulting force, zero for $x = 0$, becomes different from zero if $x \neq 0$ and equal to βx if x is sufficiently small. If β is negative, this force becomes a restoring force that can trap the atoms around $x = 0$. This is the principle of the magneto-optical trap.

2-a. Deceleration of an atomic beam

Imagine that an atomic beam is irradiated by a resonant laser beam propagating in the direction opposite to that of the beam. Due to the radiation pressure force, the atoms of the beam will slow down. Is it possible to bring them to rest? It is important to notice that even if the laser beam is initially resonant, it will not be when the atoms' velocities change, since the Doppler effect takes the atoms out of resonance; this will significantly lower the radiation pressure force, and hence the slowing down effect. For the sake of simplicity, we shall first ignore the Doppler effect following the change in the atoms' velocities; we shall see later how it can actually be circumvented.

We first assume the radiation pressure force does not change in the course of the deceleration process, and that the laser intensity is high enough for the atomic transition to be saturated; we can then use the orders of magnitude calculated in the previous paragraph. We found that for sodium atoms, the deceleration A is of the order of $10^6 \mathrm{m/s^2}$. If the atoms of the beam have an initial velocity of the order of $10^3 \mathrm{m/s}$, their velocity will be zero after a time T of the order of $10^{-3}\mathrm{s}$, after they traveled a distance $AT^2/2$ of the order of 0.5m, which shows that the size of such an experiment is not, a priori, excessive.

To solve the problem of the atoms going out of resonance because of the Doppler effect which changes in the course of the slowing down, an ingenious method was proposed and demonstrated [18]. It is based on the propagation of the atoms in a spatially inhomogeneous magnetic field. More precisely, the atomic beam travels along the axis of a spatially varying solenoid coil (Figure 2). The magnetic field produced by the solenoid is parallel to the beam direction, and its intensity varies along the beam axis. As an atom propagates along this field, it undergoes a variable Zeeman shift of its resonance frequency. One can then adjust the profile of the field so that as the atom is slowed down, the Zeeman shift of the atomic frequency balances the Doppler shift of the laser

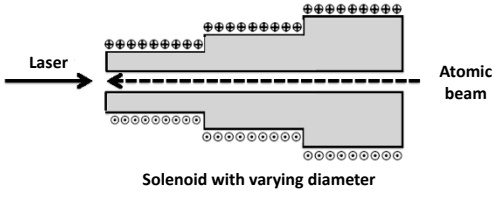


Figure 2: Schematic diagram of a Zeeman slower. The atomic beam is cooled by a laser beam propagating in the opposite direction. It travels along the axis of a solenoid composed of a set of magnetic coils with decreasing diameters, whose cross sections are shown in the figure (the current in the coils flows perpendicularly to the figure). While they propagate, the atoms are submitted to a larger and larger magnetic field. The Zeeman shift of their resonance frequency can thus follow the Doppler shift of the apparent laser frequency in their own reference frame. Consequently, instead of going off resonance, they can be slowed down during their entire propagation through the solenoid, and even come to rest.

frequency: at each point x, the field is calculated so that both Doppler and Zeeman shifts exactly balance each other. This type of apparatus is called a "Zeeman slower".

2-b. Doppler laser cooling of free atoms

α. Doppler laser cooling principle

The "slower" described above concerns the mean velocity of atoms, which can be brought down to zero. However, the root mean square of the atomic velocities remains non-zero, as does the temperature which is characterized not by the mean velocity but by its root mean square. We now describe a method using the velocity dependence of the radiation pressure force, due to the Doppler effect, and which permits reducing the dispersion of the atomic velocities around their mean value, and hence really cooling down the atoms. As this method uses the Doppler effect, it is called "Doppler laser cooling". It was proposed in 1975 for free atoms [19] and for trapped ions [20]. Here, we shall only study the case of free atoms.

The idea is to use two laser waves 1 and 2, having the same angular frequency ω and the same intensity, counterpropagating along the Ox axis, wave 1 toward negative x, and wave 2 toward positive x (Figure 3). Imagine an atom also propagating along the Ox axis with, for example, a positive velocity v. We call ω_0 the angular frequency of the atomic transition $a \leftrightarrow b$ excited by the laser. We assume the lasers are "red-detuned", meaning that:

$$\omega < \omega_0 \tag{16}$$

In the reference frame where the atom is at rest, the apparent frequency of wave i (with $i = 1, 2$) is Doppler shifted, and equal to $\omega - \mathbf{k_i} \cdot \mathbf{v}$. As v is positive, the atom and wave 1 propagate in opposite directions, so that $\mathbf{k_1} \cdot \mathbf{v}$ is negative. The apparent frequency

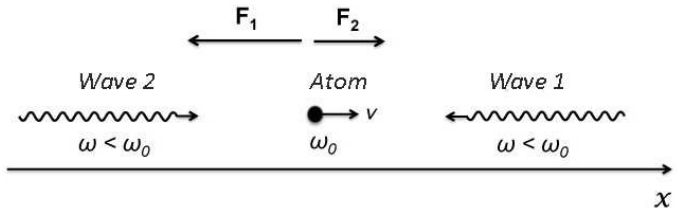


Figure 3: Principle of Doppler laser cooling. An atom moving with velocity v along the Ox axis interacts with two laser waves 1 and 2 having the same angular frequency ω; this frequency is "red-detuned", meaning $\omega < \omega_0$. The two waves have the same intensity but propagate in opposite directions along the Ox axis. The direction of the velocity v is opposite to that of wave 1. In the atom's frame of reference and because of the Doppler effect, the frequency of wave 1 is shifted closer to resonance, whereas the frequency of wave 2 is shifted away from it. Consequently, the modulus of force F_1 exerted by wave 1 increases, whereas that of the force F_2 exerted by wave 2 decreases. The resulting force, zero for $v = 0$, is opposed to v when $v \neq 0$, and proportional to v for small enough values of v. It is therefore equivalent to a friction force.

$\omega - \mathbf{k}\cdot\mathbf{v}$ of wave 1 is therefore increased. The Doppler effect brings the apparent frequency of wave 1 and the atomic resonance frequency closer together; consequently, the modulus of the radiation pressure force F_1 exerted by wave 1 on the atom and directed, like wave 1, towards the negative x, increases with respect to its value for $v = 0$. The conclusions are just the opposite for wave 2, whose frequency is shifted away from resonance by the Doppler effect; the force F_2 it creates along the positive x is weaker than its value for $v = 0$.

The sum of the two forces[5] is zero for $v = 0$ (both forces have the same modulus and opposite directions), but no longer zero when $v \neq 0$. For positive values of v, it has the same direction as F_1 since the modulus of F_1 is larger than that of F_2 (Figure 3); for negative values of v, it is just the opposite since the two waves have switched their roles. The sum $F = F_1 + F_2$ of the two radiation pressure forces exerted by the two waves is thus in the direction opposite to that of the velocity v. For small values of v, it is proportional to v and can be written:

$$F = -\gamma v \tag{17}$$

where γ is a positive friction coefficient.

Under the effect of this friction force, the atomic velocities are constantly reduced. Their dispersion, however, does not tend towards zero for long interaction times, because of the unavoidable fluctuations in the emission and absorption processes. There is actually a competition between the friction effect we just described, which tends to cool down the atoms, and the momentum diffusion that tends to heat them up. We evaluate in the next paragraphs the effect of these two processes to estimate the order of magnitude of the temperatures that can be obtained by Doppler laser cooling.

[5]We shall see below that the effects of interference between the two waves can be neglected.

β. Estimation of the friction coefficient

We will now seek an estimation of the friction coefficient, in order to calculate the evolution of the momentum of the atom, as well as of its energy. In a first step and for the sake of simplicity, we will limit ourselves to a calculation of the average effect of the photon absorption and emission cycles by the atom. This will allow us to determine the evolution of its average momentum $\overline{p}$. Nevertheless, the absorption and emission processes of photons by an atom are actually fluctuating processes, as we will discuss below. Ssuch a calculation is theferore not sufficient to obtain the evolution of the average $\overline{p^2}$ of the square of the momentum, that is of the average kinetic energy, since the average value of a square differs in general from the square of the average value. In a second step, we will use a calculation that takes the fluctuations of the momentum transfers betwee photons and atoms into account.

We set:

$$\delta = \omega - \omega_0 + \omega_R \tag{18}$$

where $\hbar\omega_R$ is the recoil energy defined in (7); δ is the detuning between the laser frequency ω and the atomic frequency $\omega_0 - \omega_R$. We assume from now on that the intensity of the two lasers is weak; the atomic transition is not saturated and consequently the population σ_{bb} of the excited level b remains low. Its variation as a function of the detuning δ follows a Lorentzian curve with a total width at half maximum equal to Γ:

$$\sigma_{bb}(\delta) = \sigma_{bb}(0)\frac{(\Gamma/2)^2}{[(\Gamma/2)^2 + \delta^2]} \tag{19}$$

We shall not need the expression for $\sigma_{bb}(0)$, since it cancels out in the expressions for the friction and diffusion coefficients we shall obtain; it can however be found in Chapter V of reference [21] (optical Bloch equations).

In a perturbative approach to the problem, two types of terms must be considered: the "square" terms, coming from the interaction between the dipole induced by wave i with that same wave i; the "cross" terms coming from the interaction of the dipole induced by wave i with the other wave $j \neq i$. The cross terms correspond to interference between the effects of the two waves. However, as these two waves do not have the same spatial dependence (they propagate in opposite directions), these interference effects vary rapidly in space as $\exp(\pm 2ikx)$. They consequently vanish when the forces are averaged over distances of the order of the laser wavelengths, as we shall assume. It is then possible to consider that the radiation pressure force acting on the atom is simply the sum of the radiation pressure forces exerted by each wave, in the absence of the other.

When the atom is at rest, the two waves have the same frequency in the atom's reference frame, and hence the same detuning δ; remember that δ is supposed to be negative in a laser cooling experiment – see (16). If the atom moves with a velocity $v > 0$, we just saw that the apparent frequency of wave 1 is increased by a quantity kv (where k and v are positive), so that the detuning for the interaction of the atom with wave 1 is equal to:

$$\delta_1 = \delta + kv \tag{20}$$

while it is $\delta_2 = \delta - kv$ for wave 2. For this atom, the population of the excited state is the sum of two contributions: that of wave 1, obtained by replacing δ with $\delta + kv$ in

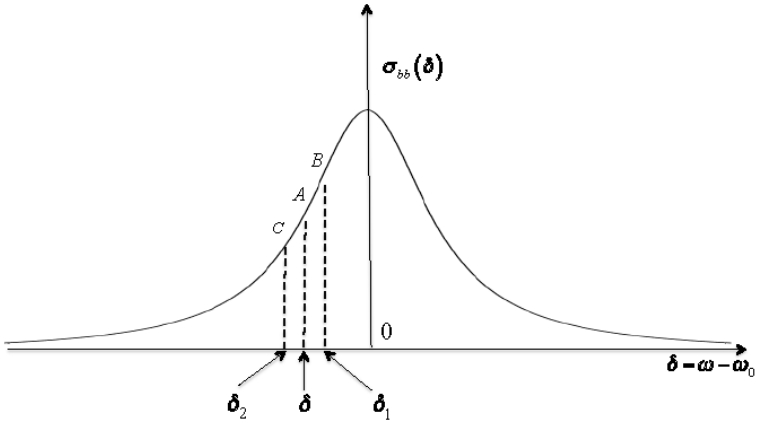


Figure 4: Populations $\sigma_{\rm bb}(\delta_1)$ and $\sigma_{\rm bb}(\delta_2)$ excited in the upper state b by wave 1, with detuning $\delta_1 = \delta + kv$, and wave 2, with detuning $\delta_2 = \delta - kv$. As the detuning δ for $v = 0$ is assumed to be negative, the ordinate $\sigma_{\rm bb}(\delta_1)$ of point B is higher than the ordinate $\sigma_{\rm bb}(\delta_2)$ of point C.

expression (19), and which is proportional to the ordinate of point B whose abscissa is $\delta_1 = \delta + kv$ in Figure 4; that of wave 2, which requires to replace δ with $\delta - kv$ in (19), and which is therefore proportional to the the ordinate of point C having the abscissa $\delta - kv$ in that same figure. Computations similar to those of § 1-d allows computing the total force F acting on the atom, as the sum of the two forces exerted by each wave, added independently of each other:

$$F = -\Gamma\hbar k\sigma_{bb}(\delta + kv) + \Gamma\hbar k\sigma_{bb}(\delta - kv) \tag{21}$$

where $\sigma_{bb}(\delta)$ is the function defined on the right-hand side of (19).

When kv is small compared to the width Γ of the curve in Figure 4, one can expand $\sigma_{bb}(\delta \pm kv)$ to first order in kv and obtain:

$$F = -2kv\ \Gamma\hbar k\ \frac{\mathrm{d}}{\mathrm{d}\delta}\sigma_{bb}(\delta) \tag{22}$$

The last factor on the right-hand side of (22), which is the slope at point A (with abscissa δ) of the curve representing $\sigma_{bb}(\delta)$, can be computed from (19). For the point $\delta = -\Gamma/2$ where the slope is maximum, we find:

$$\frac{\mathrm{d}}{\mathrm{d}\delta}\sigma_{bb}(\delta) = \frac{2\sigma_{bb}(\delta)}{\Gamma} \tag{23}$$

Inserting (23) into (22), we get:

$$F = -\gamma v \tag{24}$$

where the friction coefficient γ is given by:

$$\gamma = 4\hbar k^2\sigma_{bb}(\delta) \tag{25}$$

Using equation (24), we can also compute the damping of the momentum $\overline{p}$ and of its square. As $F = \mathrm{d}\overline{p}/\mathrm{d}t$ and $v = \overline{p}/M$, we have:

$$\frac{\mathrm{d}\overline{p}}{\mathrm{d}t} = -\gamma\frac{\overline{p}}{M} \tag{26}$$

We can also write:

$$\frac{\mathrm{d}}{\mathrm{d}t}\overline{p}^2 = 2\overline{p}\frac{\mathrm{d}\overline{p}}{\mathrm{d}t} = -2\gamma\frac{\overline{p}^2}{M} \tag{27}$$

Remember, however, that the average value of the kinetic energy is proportional to $\overline{p^2}$, the average value of the square of the momentum, and not to the square of its average value $\overline{p}^2$. We should therefore not conclude from relation (27) that the ultimate temperature that can be reached by Doppler laser cooling when $t \to \infty$ is zero. Equation (27) was obtained by considering only the average effect on the atom's velocity of the light beams and of the successive spontaneous emission processes, which introduces a continuous average evolution. But, in reality, the absorption and emission processes fluctuate and yield photons emitted in random directions. Even though their total momentum has a zero average (meaning $\overline{p}$ is not affected), these fluctuations will change $\overline{p^2}$. This effect can be considered as a source of noise (also called momentum diffusion) that increases $\overline{p^2}$, and acts in the direction opposite to the friction. It is therefore the competition between these two opposed mechanisms that leads to an equilibrium state, whose energy $\overline{p^2}/2M$ determines the ultimate temperature that can be obtained.

γ. *Momentum diffusion*

We now study the diffusion of the momentum of one atom, due to the spontaneously emitted photons; the ensemble of particles discussed above then reduces to one single atom ($P = p$). Let us consider a time interval $\mathrm{d}t$ whose value will be specified later. We call $\mathrm{d}N_1$ and $\mathrm{d}N_2$ the photon numbers from waves 1 and 2 that are absorbed during that time interval. As we assume the friction has acted long enough to cancel the average velocity, the detuning δ has become the same for the two beams. We then have, on the average:

$$\mathrm{d}\overline{N_1} = \mathrm{d}\overline{N_2} = \mathrm{d}N \tag{28}$$

Each absorbed photon then yields a spontaneously emitted photon. We use here a simple one-dimension model: each photon is emitted spontaneously in a random way, either in the positive x direction (the atom's recoil is then negative), or in the negative x direction (the recoil is then positive). The variation of the atom's momentum is then $\xi_i\hbar k$, with $\xi_i = -1$ in the first case and $\xi_i = +1$ in the second. There are no correlations in the directions of two consecutive photons. The total momentum $\mathrm{d}P$ gained by the atom during the absorption and re-emission of $\mathrm{d}N_1 + \mathrm{d}N_2 = 2\ \mathrm{d}N$ photons is equal to the sum of the momentum gained by the absorption of photons from beam 1, of the momentum gained by the absorption from beam 2, and finally of the momentum coming from the spontaneous emissions:

$$\mathrm{d}\overline{p} = \hbar k\left[\mathrm{d}N_2 - \mathrm{d}N_1 + \sum_{i=1}^{2\ \mathrm{d}N}\xi_i\right] \tag{29}$$

(i) To begin with, we neglect the fluctuations of the number of photons absorbed from each wave (we shall discuss this point later). The numbers $\mathrm{d}N_1$ and $\mathrm{d}N_2$ are then equal to their average values, and we have:

$$\mathrm{d}\overline{p} = \hbar k \sum_{i=1}^{2\ \mathrm{d}N} \xi_i \tag{30}$$

The summation over i in (29) yields a zero contribution to $\mathrm{d}\overline{p}$, since the sum of the ξ_i is zero on the average, but this is not the case for the contribution to $\mathrm{d}\overline{p^2}$ that we now compute. Taking the square of the summation over i in (30), the cross terms $\overline{\xi_i \xi_j}$ with $j \neq i$ are zero on the average, because there are no correlations between the signs of the ξ_i and the ξ_j. We are left with the square terms $\overline{\xi_i^2}$ which are all equal to 1. We obtain:

$$\mathrm{d}\overline{p^2} = 2\ \mathrm{d}N\ \hbar^2 k^2 \tag{31}$$

which is obviously non-zero.

Let us specify the time interval $\mathrm{d}t$, which should be long enough for the number dN of absorptions and spontaneous emissions to be large during this time, but sufficiently small for the variations of σ_{bb} to stay negligible. We can then write:

$$\mathrm{d}N = \dot{N}\mathrm{d}t = \Gamma\sigma_{bb}(\delta)\mathrm{d}t \tag{32}$$

where we used in the last equality the fact that the average number of photons absorbed per unit time in each wave is equal to $\Gamma\sigma_{bb}(\delta)$, as we have seen above. Inserting this result in (31), we get the increase of $\langle P^2 \rangle$ during the time interval $\mathrm{d}t$:

$$\mathrm{d}\overline{p^2} = 2\hbar^2 k^2 \mathrm{d}N = 2\hbar^2 k^2 \Gamma\sigma_{bb}(\delta)\mathrm{d}t \tag{33}$$

We finally get:

$$\left(\frac{\mathrm{d}\overline{p^2}}{\mathrm{d}t}\right)_{\mathrm{sp}} = 2\hbar^2 k^2 \Gamma\sigma_{bb}(\delta) = D_{\mathrm{sp}} \tag{34}$$

The subscript "sp" of the parenthesis reminds us that this increase of $\langle P^2 \rangle$ per unit time is due to spontaneous emission processes. This expression is often called the "diffusion coefficient" D_{sp}.

(ii) We now study the effect of fluctuations on the numbers $\mathrm{d}N_1$ and $\mathrm{d}N_2$ of absorbed photons in each wave during the time interval $\mathrm{d}t$. If these fluctuations are no longer neglected, we must write:

$$\mathrm{d}N_i = \mathrm{d}\overline{N_i} + \delta n_i = \mathrm{d}N + \delta n_i \tag{35}$$

with $i = 1, 2$. In this equation, δn_i is the fluctuation of the number of absorbed photons in wave i (by definition, the average value of δn_i is zero). The total momentum $\mathrm{d}P$ the atom receives during the absorption of these photons is equal to:

$$\mathrm{d}p = \hbar k(\mathrm{d}N_2 - \mathrm{d}N_1) = \hbar k(\delta n_2 - \delta n_1) \tag{36}$$

Since the average values of δn_1 and δn_2 are zero, the fluctuations in the absorption process do not change the average value of $\overline{p}$, but this is not true for the average value

$\overline{p^2}$. Taking the square of (36) and using the fact that there are no correlations between the fluctuations of δn_1 and δn_2 (and hence the average value of their product is zero), we get:

$$\mathrm{d}\overline{p^2} = \left[\overline{\delta n_1^2} + \overline{\delta n_2^2}\right]\hbar^2 k^2 \tag{37}$$

To compute $\overline{\delta n_i^2}$, we take the square of equation (35). This leads to:

$$\mathrm{d}N_i^2 = \mathrm{d}N^2 + 2\ \mathrm{d}N\delta n_i + \delta n_i^2 \tag{38}$$

We now take the average value of each side of this equation. Using the fact that the average value of δn_i is zero, we get:

$$\overline{\delta n_i^2} = \mathrm{d}\overline{N_i^2} - \mathrm{d}N^2 \tag{39}$$

The quantity $\mathrm{d}\overline{N_i^2} - \mathrm{d}N^2$ is simply the variance of the number of photons absorbed from the wave. In general, for Poisson statistics, we have[6]:

$$\mathrm{d}\overline{N_i^2} - \mathrm{d}\overline{N_i}^2 = \mathrm{d}\overline{N_i} = \mathrm{d}N \tag{40}$$

We deduce that $\overline{\delta n_i^2}$ is simply equal to $\mathrm{d}N$, so that equation (37) is simply written:

$$\mathrm{d}\overline{p^2} = 2\,\mathrm{d}N\ \hbar^2 k^2 \tag{41}$$

We therefore obtain for the increase of $\overline{p^2}$ due to fluctuations in the absorption processes a result identical to (31). The computations leading from (31) to (34) can be repeated and we get the following result for the increase, per unit time, of $\overline{p^2}$ due to the absorption processes:

$$\left(\frac{\mathrm{d}\overline{p^2}}{\mathrm{d}t}\right)_{\mathrm{abs}} = 2\hbar^2 k^2 \Gamma \sigma_{bb}(\delta) = D_{\mathrm{abs}} \tag{42}$$

where the diffusion coefficient D_{abs} due to the absorption processes has the same value as D_{sp} for the spontaneous emission processes:

$$D_{\mathrm{abs}} = D_{\mathrm{sp}} = 2\hbar^2 k^2 \Gamma \sigma_{bb}(\delta) \tag{43}$$

To evaluate the global rate of variation of $\overline{p^2}$, we must add to the rates of variation (34) and (42) the one due to the cooling process $(\mathrm{d}\overline{p^2}/\mathrm{d}t)_{\mathrm{cool}}$. This variation is simply the variation of $\overline{p}^2$ in the absence of fluctuations during the momentum exchanges, so that it can be obtained by assuming that $\overline{p^2}$ and $\overline{p}^2$ are simply equal. Using (27), we can write this variation as:

$$\left(\frac{\mathrm{d}\overline{p^2}}{\mathrm{d}t}\right)_{\mathrm{cool}} = -2\gamma\frac{\overline{p^2}}{M} \tag{44}$$

[6] One could evaluate the effects of the deviations from Poisson statistics, but it will not be done here to keep things simple; this is legitimate for the low laser intensities (unsaturated transition) we assumed here.

Adding (34), (42) and (44), we finally obtain a total rate of change equal to:

$$\left(\frac{\mathrm{d}\overline{p^2}}{\mathrm{d}t}\right)_{\mathrm{tot}} = -2\gamma\frac{\overline{p^2}}{M} + D_{\mathrm{sp}} + D_{\mathrm{abs}} \tag{45}$$

This rate of change goes to zero when :

$$\overline{p^2} = (D_{\mathrm{sp}} + D_{\mathrm{abs}})\frac{M}{2\gamma} \tag{46}$$

Dividing both sides of this equation by $2M$ and using expressions (25) and (43) for the friction and diffusion coefficients, we finally get[7] the expression for the average kinetic energy in the stationary regime:

$$\frac{\overline{p^2}}{2M} = \frac{\hbar\Gamma}{4} \tag{47}$$

This result indicates that the ultimate temperature is directly related to the natural width of the excited level.

δ. Doppler temperature

In a Doppler laser cooling experiment, and for a one-dimensional treatment of the problem, the average kinetic energy of the velocity fluctuations around its average value is related to the Doppler temperature T_D :

$$\frac{\overline{p^2}}{2M} = \frac{k_B T_D}{2} \tag{48}$$

where k_B is the Boltzmann constant. Using (47) and (48), we find that the equilibrium temperature that can be reached by Doppler laser cooling is given by:

$$T_D = \frac{\hbar\Gamma}{2k_B} \tag{49}$$

For sodium atoms, this temperature is equal to 235×10^{-6} K, that is 10^6 times lower than room temperature (of the order of 300 K)!

Our treatment of Doppler laser cooling is based on several approximations, as for example the one-dimensional treatment and the simplified description of spontaneous emission occurring only in two opposite directions. Nevertheless, more precise calculations lead to the conclusion that the average kinetic energy reached in a stationary regime is indeed of the order of $\hbar\Gamma$, within numerical coefficients of a few units.

Finally, equation (47) yields an estimation of the velocity $\bar{v}_D$ of the cold atoms thus obtained:

$$\bar{v}_D \simeq \sqrt{\hbar\Gamma/M} \tag{50}$$

This means that the Doppler shifts $\pm k\bar{v}_D$ of the apparent frequencies of waves 1 and 2 are such that the separations of the dotted lines in Figure 4 is of the order of $\pm k\sqrt{\hbar\Gamma/M}$.

[7] As the diffusion and friction coefficients are all proportional to $\sigma_{bb}(\delta)$, this factor disappears from (46) and is no longer present in (47).

We can compare these Doppler shifts with the width Γ of the curve plotted in Figure 4 and obtain:

$$\frac{k\sqrt{\hbar\Gamma/M}}{\Gamma} \simeq \sqrt{\frac{\hbar k^2/M}{\Gamma}} \simeq \sqrt{\frac{E_{\rm rec}}{\hbar\Gamma}} \tag{51}$$

where $E_{\rm rec}$ is the recoil energy given in (7). The ratio $E_{\rm rec}/\hbar\Gamma$ is in general small for the atomic resonance lines used in laser cooling; it is of the order of 1/100 for sodium, which means that $k\bar{v}_D$ remains small compared to Γ and shows the validity of the limited series expansion used in equation (22).

Other laser cooling methods

Until now we have described laser cooling methods using only the Doppler effect. Other methods have been proposed and demonstrated, such as "Sisyphus cooling" (§ 4 of Complement D_{XX} and [22]), the "subrecoil cooling" and "evaporative cooling" [23]. The reader interested in the two latter methods may read for instance § 13.3 of [24].

2-c. Magneto-optical trap

We wish to introduce an imbalance, depending on the atom's position x, between two laser beams 1 and 2, with the same frequency but propagating in opposite directions along the Ox axis. This requires achieving detunings (between the lasers' frequency and the atomic frequency) that depend on the position x of the atom along the Ox axis; these detunings must be the same for both lasers when $x = 0$, and different when $x \neq 0$. We must then necessarily use an atom with several Zeeman sublevels and different polarizations for the two beams 1 and 2. The principle of the method, suggested for the first time by Jean Dalibard in 1986, is schematized in Figure 5. We assume the atomic transition is between a ground state with a zero angular momentum ($J_g = 0$) and an excited state with an angular momentum equal to 1 ($J_e = 1$). The solid lines in Figure 5 plot the energy of the three Zeeman sublevels e_{+1}, e_0 and e_{-1} of the excited state, and of the sublevel g_0 of ground state, in an inhomegeneous magnetic field B applied along the Ox axis. This field is zero at $x = 0$ and varies linearly with x around $x = 0$; it can be created, for example, by two circular coils having the same axis Ox, placed symmetrically with respect to $x = 0$, and carrying currents of opposite directions. The quantization axis is chosen along Ox and allows defining the magnetic quantum numbers m_e and m_g of the sublevels in the excited and ground states. The two laser waves 1 and 2 propagate in opposite directions and have opposite circular polarizations with respect to the quantization axis [8]. Wave 1, with polarization σ_+, excites the transition $g_0 \leftrightarrow e_{+1}$ whereas wave 2, with polarization σ_-, excites the transition $g_0 \leftrightarrow e_{-1}$. The energy $\hbar\omega_0$ of the zero-field atomic transition is equal to the difference between the energies of states e_0 and g_0 (solid horizontal lines in the figure). The detuning $\hbar\delta = \hbar\omega - \hbar\omega_0$ between the energy $\hbar\omega$ of the laser photons and that of the zero-field atomic transition is shown as the difference in height between the dashed and solid horizontal lines in the figure. We

[8] One generally defines the right-hand and left-hand circular polarization with respect to the propagation direction of the photon. In that case, the polarization of the photons of wave 1 and wave 2 in Figure 5 would be σ_+ (in both cases the electric field of the wave turns around the direction of propagation following the right-hand rule). When studying the selection rules of the various transitions resulting from the conservation of spin angular momentum (see Complements B_{XIX} and C_{XIX}), it is best to define both the quantum numbers and the polarizations of all the beams with respect to the same axis, chosen here to be the quantization axis Ox.

assume here that the detuning is negative (the laser's frequency ω is shifted towards the red of the atomic transition ω_0).

At point $x = 0$, the energies of states e_{+1} and e_{-1} are equal, as are the detunings of waves 1 and 2 which excite the transitions $g_0 \leftrightarrow e_{+1}$ and $g_0 \leftrightarrow e_{-1}$. The radiation pressure forces exerted by the two waves are equal in intensity and opposite in direction, so that the resulting force is zero. This balance does not hold as soon as we move away from $x = 0$. For example, at $x = x_+$, wave 1 which excites transition $g_0 \rightarrow e_{+1}$, is at resonance with this transition and the force it exerts is at its maximum. On the other hand, at that same point, wave 2, which excites transition $g_0 \rightarrow e_{-1}$, is way off-resonance and hence exerts a much weaker force. Consequently, the balance between the two waves is broken in favor of wave 1 and the resulting force exerted by both waves is non-zero and directed towards the right. The conclusions are inverted at point $x = x_-$, where the total force is non-zero and directed towards the left. We obtain a restoring force, proportional to x in the vicinity of $x = 0$, which traps the atoms around $x = 0$. Such a trap is called a "magneto-optical trap" or "MOT".

For the sake of simplicity, we only considered a one-dimensional model, but the extension to three dimensions is possible. Note in particular that the field created by

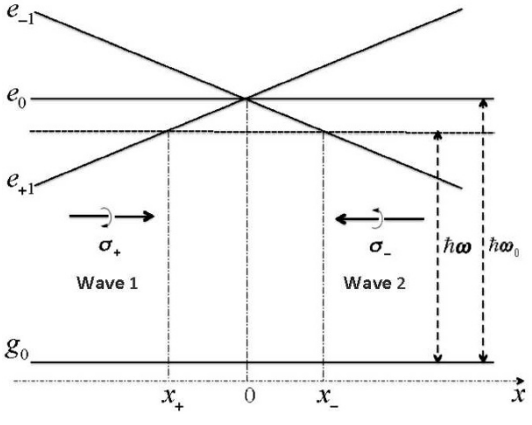


Figure 5: Principle of the magneto-optical trap. The transition used is a $J_g = 0 \rightarrow J_e = 1$ transition, excited by two laser waves 1 and 2 propagating in opposite directions along the Ox axis, with polarizations (σ_+) and (σ_-) with respect to the quantization axis Ox. A magnetic field gradient is applied along the Ox axis, the magnetic field being equal to zero at $x = 0$. Wave 1, which excites the transition $g_0 \leftrightarrow e_{+1}$, is resonant for this transition at point $x = x_+$ and the radiation pressure force it exerts on the atoms is then maximal. At that point, wave 2, which excites the transition $g_0 \leftrightarrow e_-$, is off-resonance for the corresponding atomic frequency, and, consequently, exerts a much weaker force. The two forces are unbalanced, in favor of the wave 1 force. The resulting force is non-zero, directed towards $x > 0$. The conclusions are reversed at point $x = x_-$, where the resulting force is non-zero, directed towards $x < 0$. Finally at $x = 0$ both waves are off-resonance by the same amount, and the resulting force is zero. We obtain a restoring force that traps the atoms around $x = 0$.

two circular coils centered around the Ox axis, placed on each side of point $x = 0$ and carrying currents in opposite directions, is zero at $x = 0$ and exhibits non-zero gradients along the Oy and Oz axis. The detuning towards the red of the laser frequency with respect to the atomic frequency also has the advantage of providing a Doppler laser cooling effect. The magneto-optical trap is nowadays a basic tool of cold atom physics[9].

Other trapping methods

Other methods for trapping atoms with light beams exist, for instance laser dipole trapping methods (Complement D_{XX}, §§ 1, 2 and 3).

3. Blocking recoil through spatial confinement

We now assume the atom or the ion under study is subjected to an external potential that traps it in a region of space. The energy spectrum of the external variables is no longer a continuous spectrum (as would be the case for a free atom), but a spectrum including a discrete part corresponding to the atomic bound states. Furthermore, because of this external potential, the atomic Hamiltonian is no longer translation invariant, and hence the total momentum is no longer a good quantum number. In this section we study how the absorption and emission spectra of the atom are modified by its confinement within the potential, and how the recoil of the atom can be blocked in certain cases.

3-a. Assumptions concerning the external trapping potential

We will assume the external potential acts only on the external variables, and not on the internal variables. This is the case for example for an atomic ion trapped by electric and magnetic fields, which only act on the center of mass via the global ionic charge, but does not act on the internal variables [10]. Figure 6 plots the trapping potentials for an atom or an ion in the internal states a or b. The two potential curves are identical, deduced from one another by a vertical translation of amplitude $\hbar\omega_0$ where ω_0 is the frequency of the internal atomic transition $b \leftrightarrow a$. The spectrum of the vibrational levels, of energies E_v, $E_{v'}$, $E_{v''}$, ... is the same for the two potentials. The atomic states are labeled by two quantum numbers: a quantum number a or b for the internal variables; a vibrational quantum number v, v', v'', ... for the external variables. The atomic transitions between the two internal states a and b now present a structure due to the vibrational motion of the center of mass. The frequency $\omega_{v',v}$ of the transition $b, v' \leftrightarrow a, v$ is given by:

$$\hbar\omega_{v',v} = \hbar\omega_0 + E_{v'} - E_v \tag{52}$$

The quantity $E_{v'} - E_v$ gives the variation of the external energy of the atom during the $b, v' \to a, v$ transition.

[9] See § 14-7 of [24] for a description of the first experimental realizations of such a trap and for a more quantitative study of its performances.

[10] The ion is generally confined in the center of the trap, a region where the electric and magnetic fields are very weak. It is then legitimate to neglect the Stark or Zeeman shifts of the internal states.

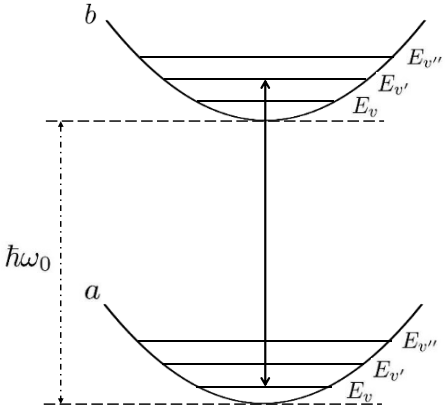


Figure 6: The external potential trapping the ion is the same when the ion is in either of its internal states a or b, separated by an energy difference $\hbar\omega_0$. Consequently, the spectrum of the vibrational levels of the center of mass in the external potential is the same for both internal states.

3-b. Intensities of the vibrational lines

We showed in § C-5 of Chapter XIX that the matrix elements of the interaction Hamiltonian could be factored into three terms pertaining to the three types of variables, the internal and external atomic variables, and the radiation variables – see relations (C-34) and (C-35) of that chapter. The part relative to the external variables is equal to $\langle\psi_{\text{fin}}^{\text{ext}}|\exp(i\,\boldsymbol{k}\cdot\hat{\boldsymbol{R}})|\psi_{\text{in}}^{\text{ext}}\rangle$ for an absorption process, where $|\psi_{\text{fin}}^{\text{ext}}\rangle$ and $|\psi_{\text{in}}^{\text{ext}}\rangle$ are the external final and initial states of the transition, equal here to $|\varphi_{v'}\rangle$ and $|\varphi_v\rangle$. This leads to an intensity of the vibrational line $a, v \leftrightarrow b, v'$ proportional to:

$$I_{v',v} = |\langle\varphi_{v'}|\exp(i\boldsymbol{k}\cdot\hat{\boldsymbol{R}})|\varphi_v\rangle|^2 \tag{53}$$

The $I_{v',v}$ obey the sum rule (obtained by the closure relation on the states $|\varphi_{v'}\rangle$):

$$\sum_{v'} I_{v',v} = \sum_{v'}\langle\varphi_v|\exp(-i\boldsymbol{k}\cdot\hat{\boldsymbol{R}})|\varphi_{v'}\rangle\langle\varphi_{v'}|\exp(i\boldsymbol{k}\cdot\hat{\boldsymbol{R}})|\varphi_v\rangle = 1 \tag{54}$$

It follows that the relative weight of the transition $a, v \to b, v'$ compared to all the transitions starting from a, v is precisely equal to $I_{v',v}$.

Another sum rule

The relative weights $I_{v',v}$ obey another sum rule:

$$\langle E\rangle = \sum_{v'} I_{v',v}(E_{v'} - E_v) = \frac{\hbar^2 k^2}{2M} = E_{\text{rec}} \tag{55}$$

which indicates that the average energy gained by an atom going from a given level a, v to another level b, v' is equal to the recoil energy, whatever the value of v. To prove relation (55), we rewrite the sum over v' in (55) in the form:

$$\sum_{v'} \langle\varphi_v| \left[\exp(-i\boldsymbol{k}\cdot\hat{\boldsymbol{R}}), \hat{H}_{\text{ext}}\right] |\varphi_{v'}\rangle \langle\varphi_{v'}| \exp(i\boldsymbol{k}\cdot\hat{\boldsymbol{R}}) |\varphi_v\rangle \tag{56}$$

where $\hat{H}_{\text{ext}} = \hat{\boldsymbol{P}}^2/2M + V(\hat{\boldsymbol{R}})$ is the external variable Hamiltonian. The only term in H_{ext} that does not commute with $\exp(-i\boldsymbol{k}\cdot\hat{\boldsymbol{R}})$ is the kinetic energy term. We can replace the commutator appearing in (56) by $\left[\exp(-i\boldsymbol{k}\cdot\hat{\boldsymbol{R}}), \hat{\boldsymbol{P}}^2/2M\right]$. We now develop this commutator and use the closure relation on the $|\varphi_{v'}\rangle$ states, as well as relation:

$$\exp(-i\boldsymbol{k}\cdot\hat{\boldsymbol{R}}) \left(\frac{\hat{\boldsymbol{P}}^2}{2M}\right) \exp(+i\boldsymbol{k}\cdot\hat{\boldsymbol{R}}) = \frac{1}{2M}(\hat{\boldsymbol{P}} + \hbar\boldsymbol{k})^2 \tag{57}$$

This yields:

$$\begin{aligned}\langle\varphi_v| \exp(-i\boldsymbol{k}\cdot\hat{\boldsymbol{R}}) \frac{\hat{\boldsymbol{P}}^2}{2M} \exp(+i\boldsymbol{k}\cdot\hat{\boldsymbol{R}}) - \frac{\hat{\boldsymbol{P}}^2}{2M} |\varphi_v\rangle \\ = E_{\text{rec}} + \langle\varphi_v|\hbar\boldsymbol{k}\cdot\hat{\boldsymbol{P}}/M|\varphi_v\rangle = E_{\text{rec}}\end{aligned} \tag{58}$$

since, taking Ehrenfest's theorem into account (Chapter III, § D-1-d-β), the average value of operator $\hat{\boldsymbol{P}}$ (equal to $M\ \mathrm{d}\langle\boldsymbol{R}\rangle/\mathrm{d}t$) in the stationary state $|\varphi_v\rangle$ is zero.

3-c. Effect of the confinement on the absorption and emission spectra

The absorption and emission spectra of an atom are significantly modified by the confinement of its center of mass.

As we saw in § 1, when a free atom has a well defined initial momentum $\boldsymbol{P}_{\text{in}}$, the absorption of a photon with momentum $\hbar\boldsymbol{k}$ places it in a well defined momentum state $\boldsymbol{P}_{\text{fin}} = \boldsymbol{P}_{\text{in}} + \hbar\boldsymbol{k}$. The conservation of the global momentum means that there is only one final state, $|b, \boldsymbol{P}_{\text{fin}}\rangle$, corresponding to an initial state $|a, \boldsymbol{P}_{\text{in}}\rangle$, and hence a single absorption line $a \leftrightarrow b$.

On the other hand, when the global momentum is no longer conserved because of the external potential's trapping of the atom, we get several lines going from an initial given state a, v to several possible final states b, v', whose frequencies $\omega_{v',v}$ are given by (52). One can then ask which of these lines is the strongest.

To answer that question, we go back to expression (53) for the relative weight of the transition $a, v \to b, v'$. In this equation, operator $\exp(i\,\boldsymbol{k}\cdot\hat{\boldsymbol{R}})$ represents a translation operator, in momentum space, of the quantity $\hbar\boldsymbol{k}$. The quantity $I_{v',v}$ is thus proportional to the squared modulus of the scalar product of the vibrational wave function $\varphi_{v'}(\boldsymbol{r})$ and the wave function $\varphi_v(\boldsymbol{r})$ translated in momentum space by the quantity $\hbar\boldsymbol{k}$. Let us assume the atom is trapped in a region of spatial extension Δx, very small compared to the wavelength $\lambda = 2\pi/k$ of the incident photon. The momentum spread Δp of the wave function is then larger than $\hbar k$ since:

$$\Delta x \ll \lambda \tag{59}$$

leads to:

$$\Delta p \simeq \frac{\hbar}{\Delta x} \gg \frac{\hbar}{\lambda} = \frac{\hbar k}{2\pi} \tag{60}$$

This means that, as long as the excited levels' energy is not too large, the translation in momentum space of the wave function $\varphi_v(\boldsymbol{r})$ by a quantity $\hbar\boldsymbol{k}$ much smaller than its width leaves that wave function practically unchanged, and it consequently remains orthogonal to $\varphi_{v'}(\boldsymbol{r})$ for $v' \neq v$. The strongest line in the absorption spectrum is therefore the line with no change in the vibrational quantum number (a line sometimes called the *zero-phonon line*), whose frequency remains unchanged, equal to ω_0. A strong confinement suppresses the Doppler shift and the atom's recoil.

Comment on momentum conservation

One may wonder what happened to the momentum of the absorbed photon in the zero-phonon transition. Remember that we have treated the trapping potential of the atom as an external potential, which breaks the translation invariance of the problem under study: the Hamiltonian no longer commutes with the total momentum, which is no longer conserved. It is thus not surprising that we cannot follow what becomes of the photon momentum. We can, however, describe the trapping potential, not as an externally given potential, but rather coming from the interaction of the atom with another physical object whose dynamics must be taken into account. A quantum treatment of that device and its interaction with the atom permits introducing for the global system, "atom + trapping device", a Hamiltonian that commutes with the total momentum; it is the global system that absorbs the photon momentum. As this momentum is microscopic, whereas the mass of the device is macroscopic, the recoil velocity is so weak that the corresponding frequency shifts are totally undetectable.

3-d. Case of a one-dimensional harmonic potential

We now assume the external trapping potential is harmonic, and we call $\omega_{\rm osc}$ the oscillation frequency of the atoms in this potential. The energies E_v of the vibrational levels are equal to $(v + 1/2)\hbar\omega_{\rm osc}$, where v is an integer, positive or zero. The spatial extension Δx_0 of the ground state wave function $v = 0$ is equal to $\sqrt{\hbar/2M\omega_{\rm osc}}$. To characterize the confinement, we introduce the dimensionless parameter[11]:

$$\eta = k\Delta x_0 = 2\pi\frac{\Delta x_0}{\lambda} \tag{61}$$

If $\eta \ll 1$, the atoms are confined in a region small compared to the radiation wavelength. The square of the parameter η has a simple physical significance since:

$$\eta^2 = \frac{\hbar k^2}{2M\omega_{\rm osc}} = \frac{E_{\rm rec}}{\hbar\omega_{\rm osc}} \tag{62}$$

is the ratio between the recoil energy $E_{\rm rec}$ and $\hbar\omega_{\rm osc}$, which is the energy difference between vibrational levels in the potential well.

It is instructive to compute, as a function of η, the intensities $I_{0,0}$ and $I_{0,1}$ of the vibrational lines $a, 0 \to b, 0$ and $a, 0 \to b, 1$. Assume the photon wave vector $\boldsymbol{k}$ is parallel to the x axis. Exponential $\exp(i\boldsymbol{k}\cdot\hat{\boldsymbol{R}})$ appearing in equation (53) can be replaced by

[11] This parameter is often called the *Lamb Dicke parameter*, after the names of the physicists who first introduced the idea of recoil-free absorption in a trapped system. To get a historical overview of the various studies on the suppression of recoil due to confinement, the interested reader may consult § 6-4-4 of [24] as well as the references cited in that §.

$\exp(ik\hat{X})$. We now use the expression of operator $\hat{X}$ in terms of the annihilation and creation operators of the harmonic oscillator associated with the external potential:

$$\hat{X} = \sqrt{\hbar/2M\omega_{\text{osc}}}\,(\hat{a} + \hat{a}^\dagger) = \Delta x_0\,(\hat{a} + \hat{a}^\dagger) \tag{63}$$

We then get, in the limit $\eta \ll 1$, using (61):

$$\begin{aligned}\exp(ik\hat{X}) &= \exp\left[i\eta(\hat{a} + \hat{a}^\dagger)\right] \\ &= 1 + i\eta(\hat{a} + \hat{a}^\dagger) - \frac{\eta^2}{2}(\hat{a} + \hat{a}^\dagger)^2 + ...\end{aligned} \tag{64}$$

The series expansion (64) used in (53) yields, to order 2 in η:

$$I_{0,0} = 1 - \eta^2 \tag{65a}$$

$$I_{1,0} = \eta^2 \tag{65b}$$

All the other transitions $a, 0 \to b, v$ with $v \geqslant 2$ have relative intensities of a higher order, in η^{2v}. The transition with no change in the vibrational state, and hence with no recoil, is predominant for a strong confinement. The transition $a, 0 \to b, 1$ has a much lower probability, by a factor $E_{\text{rec}}/\hbar\omega_{\text{osc}}$; when it occurs, it increases the atomic energy by a quantity $\hbar\omega_{\text{osc}}$ much larger than E_{rec}. The sum rule (55) shows that, on the average, the energy gained by the atom remains equal to E_{rec}.

3-e. Mössbauer effect

In 1958, Rudolf Mössbauer observed very narrow lines in the resonant absorption spectrum of γ rays by the atomic nuclei in a crystal. Building on the previous work of Lamb [25] on the suppression of the recoil in the resonant absorption of slow neutrons (and not of photons) by the atomic nuclei in a crystalline network, he attributed the narrow spectral structures he observed to a suppression of the recoil. This suppression can occur if, in the crystal phonon spectrum, there are frequencies larger than the recoil frequency $\omega_R = E_{\text{rec}}/\hbar$. The interest of the Mössbauer effect comes from the high value of the frequency of the internal γ transition, which can reach 10^{18} Hz, or even much higher frequencies. If the Doppler width and the recoil shift are suppressed by the confinement, and if the natural width remains of the order of 10^6 to 10^7 Hz (as for an optical transition), the quality factor of the γ transition (ratio between frequency and the spectral width of the resonance) can reach values of the order of 10^{12}. Such a resolution allowed measuring, already in 1960 [26], and for the first time in a laboratory, the gravitational shift predicted by general relativity between the frequencies of an emitter and a receptor, both located in the earth's gravitational field but separated by an altitude of roughly twenty meters.

4. Recoil suppression in certain multi-photon processes

Until now, we only considered one-photon processes. We shall see in Chapter XX that there are multi-photon processes in which the atom goes from an internal state a to another state b by absorbing or emitting several photons. During such processes, the total energy and momentum must of course be conserved[12].

[12] We now consider again free atoms, without an external potential.

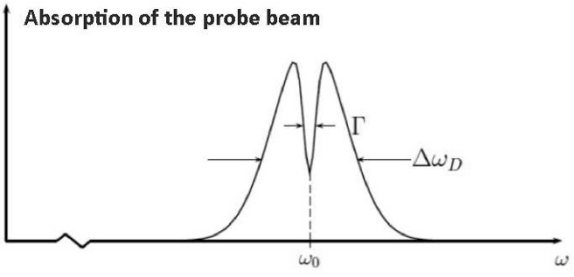


Figure 7: Saturated absorption spectroscopy. This figure plots the absorption profile of the probe beam when scanning the frequency ω of the two laser beams. A narrow hole, of width Γ, appears in the middle of a much larger Doppler profile, of width $\Delta\omega_D$.

Imagine a two-photon process, where the two photons have the same frequency and opposite wave vectors $+\boldsymbol{k}$ and $-\boldsymbol{k}$. The total radiation momentum is zero in this case. If the atom absorbs those two photons, its momentum does not change. Its external energy is not modified, meaning there is neither a Doppler effect, nor a recoil energy. This possibility can be extended to p-photon processes as long as the sum of the wave vectors of the p photons is zero:

$$\boldsymbol{k}_1 + \boldsymbol{k}_2 ... + \boldsymbol{k}_p = \boldsymbol{0} \tag{66}$$

This idea was proposed independently by two groups, in Russia [27] and in France [28]. It led to significant experimental advances in high resolution spectroscopy where the line width is no longer limited by the Doppler width, but rather by the often much smaller natural width. A particularly interesting example is the study of the transition $1s \leftrightarrow 2s$ of the hydrogen atom by Doppler-free two-photon spectroscopy [29]. The upper state $2s$ of this transition is metastable, with a long lifetime (around 120 ms). Consequently, its natural width is very small and the two-photon line very narrow, which allows extremely precise measurements of fundamental constants such as the Rydberg constant. Note must be taken however that the interaction with the laser radiation inducing the two-photon transition leads to shifts of the energy levels[13], proportional to the light intensity; these must be taken precisely into account to determine the non-perturbed frequency of the two-photon transition.

Doppler-free saturated absorption spectroscopy

Nonlinear effects also appear in experimental set-ups where the atom interacts with two counter-propagating light beams, with the same frequency ω, one having a high intensity (pump beam) and the other a weaker one (probe beam). Contrary to the two-photon transitions considered in the previous paragraph, we assume here that the transitions induced by each beam are one-photon transitions between the two atomic internal states a and b; the laser frequency ω is therefore close to the atomic transition ω_0 (and not close to $\omega_0/2$).

We neglect here the recoil energy, in general very small in the optical domain compared with the natural width Γ of the upper state b. However, the Doppler shifts of the absorp-

[13] See Complement B_{XX} , §2-b

tion lines of the various atoms play an important role, as they are different for the pump beam and the probe beam which propagate in opposite directions. The pump beam interacts with an atom of velocity v_{pump} if its apparent frequency $\omega - kv_{\text{pump}}$ for that atom coincides with the atomic frequency ω_0 (within Γ), i.e. if $kv_{\text{pump}} = \omega - \omega_0$ within Γ. In the same way, the probe beam interacts with atoms of velocity v_{probe} if $\omega + kv_{\text{probe}} = \omega_0$ within Γ, i.e. if $kv_{\text{probe}} = -(\omega - \omega_0)$ within Γ. When ω is different from ω_0, we have $v_{\text{pump}} \neq v_{\text{probe}}$: the two beams do not interact with the same atoms in the velocity distribution, so that the absorption of the probe beam is not perturbed by the presence of the pump beam. However, this perturbation becomes important when $\omega = \omega_0$ (within Γ), since the two beams interact with the same sub-set of atoms (those belonging to the same "velocity group" along the beams' axis).

The high intensity pump beam lowers the difference in populations between the a and b levels of the atomic transition, and tends to equalize these populations. The absorption of the probe beam is thus diminished when the two beams interact with the same velocity group, i.e in the vicinity of $\omega = \omega_0$. When scanning the frequency ω of the two laser beams, the absorption of the probe beam varies according to a Doppler profile centered around ω_0, with width $\Delta\omega_D$, in the middle of which (Figure 7) appears a hole with a much smaller width Γ. This method, called *saturated absorption*, allows the determination of the atomic frequency ω_0 with a much better resolution than when using a single beam.

Conclusion

In this complement, we showed how the analysis of the momentum exchanges between atoms and photons allows introducing and interpreting several important physical phenomena. These phenomena include Doppler width, recoil energy, radiation pressure forces, Doppler laser slowing down and cooling of atoms, suppression of the Doppler effect due to confinement or in two-photon transitions, and the Mössbauer effect.

Thanks to these various methods, spectacular improvements in the resolution of spectroscopic measurements have been obtained. This led to high precision measurements and improvements of atomic clocks, which now have a relative stability of the order of 10^{-16}. Placing such a clock in the international spatial station and comparing its frequency with that of a similar clock on the Earth, one hopes to be able to test the value of the gravitational shift predicted by general relativity with a precision better, by a factor close to 100, than all the other existing tests. Another conclusion we can draw is that atom-photon interactions are useful tools for controlling and manipulating atoms.

We shall see in Complement C_{XIX} how the exchanges of angular momentum between atoms and photons allows controlling the angular momentum of the atoms, polarizing them via optical pumping. Such achievements have opened new fields of research, such as atomic interferometry and the study of degenerate quantum gases.

Complement B_{XIX}

Angular momentum of radiation

1 Quantum average value of angular momentum for a spin 1 particle 2044
1-a Wave function, spin operator 2044
1-b Average value of the spin angular momentum 2045
1-c Average value of the orbital angular momentum 2046
2 Angular momentum of free classical radiation as a function of normal variables 2047
2-a Calculation in position space 2047
2-b Reciprocal space 2048
2-c Difference between the angular momenta of massive particles and of radiation 2050
3 Discussion 2050
3-a Spin angular momentum of radiation 2050
3-b Experimental evidence of the radiation spin angular momentum 2051
3-c Orbital angular momentum of radiation 2052

Introduction

Radiation angular momentum plays an important role in many situations, in particular in atomic physics experiments. As will be explained in Complement C_{XIX}, the exchange of angular momentum between atoms and photons is the base of many experimental methods, such as optical pumping, which illustrated for the first time the manipulation of atoms by light.

In Chapter XVIII, a spatial Fourier transform of the classical fields led to the introduction of the field normal variables $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ and $\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$, which are the field components in a basis of transverse plane waves. Upon quantization, these normal variables became the annihilation $\hat{a}_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ and creation $\hat{a}^\dagger_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ operators of a photon in a mode $\boldsymbol{k}, \boldsymbol{\varepsilon}$. Such a plane wave basis is particularly useful for studying the radiation energy and momentum, since the photons of mode $\boldsymbol{k}, \boldsymbol{\varepsilon}$ have a well defined energy $\hbar\omega = \hbar ck$ and momentum $\hbar\boldsymbol{k}$. On the other hand, the expansion of the field angular momentum in terms of the normal variables $\alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ and $\alpha^*_{\boldsymbol{\varepsilon}}(\boldsymbol{k})$ is not as simple, since the photons of mode $\boldsymbol{k}, \boldsymbol{\varepsilon}$ do not have a well determined angular momentum. The aim of this complement is to find another expansion better adapted to the study of the radiation angular momentum, and establish a number of useful results.

In the classical description of radiation, the normal variable $\boldsymbol{\alpha}(\boldsymbol{k}) = \sum_{\boldsymbol{\varepsilon}} \alpha_{\boldsymbol{\varepsilon}}(\boldsymbol{k})\boldsymbol{\varepsilon}$ is a vector function of $\boldsymbol{k}$ presenting a certain analogy with a wave function in the reciprocal space, and which could be seen as the wave function of the radiation (in momentum space). A physical quantity, such as the radiation total energy or the total momentum, does appear as the average value in that wave function of a one-particle operator

representing the energy or the momentum of a photon. We will see other examples of this analogy in this complement. As it is a vector function, this wave function can be regarded as the wave function of a particle of spin 1 whose total angular momentum $\boldsymbol{J}$ would be the sum of the orbital angular momentum $\boldsymbol{L}$ and the spin angular momentum $\boldsymbol{S}$. We will present, in § 1, a quantum mechanical calculation of the average values, in the state of a spin 1 particle described by the vector wave function $\boldsymbol{\Psi}(\boldsymbol{k})$, of the orbital and spin angular momentum of that particle. Returning to classical physics, in § 2 we will establish the expression for the total angular momentum $\boldsymbol{J}_R$ of free radiation; we shall first write it as a function of the fields in reciprocal space, then as a function of the field normal variables $\boldsymbol{\alpha}(\boldsymbol{k})$. The expression thus obtained has the same form as that obtained in § 1, provided we replace the $\boldsymbol{\Psi}(\boldsymbol{k})$ by the $\boldsymbol{\alpha}(\boldsymbol{k})$. This will lead us to the expansion in terms of normal variables of not only the field total angular momentum, but also of its orbital and spin angular momenta. The physical interpretation of these results is discussed in § 3, which highlights, in particular, some important characteristics of these two types of angular momenta.

1. Quantum average value of angular momentum for a spin 1 particle

We first study a spin 1 particle that has a mass, and is therefore not a photon. Our results will be useful as a point of comparison for the next paragraph's computations, where we return to the electromagnetic field.

1-a. Wave function, spin operator

The conclusions of this § 1 will be compared with those of the following § 2 concerning the radiation angular momentum expressed in terms of normal variables. As normal variables characterize the field in reciprocal space, it is important here to describe the state of the spin 1 particle in that same space. The particle state vector can be expanded on a basis $\{|\boldsymbol{k}, \nu\rangle\}$, where $\boldsymbol{k}$ represents the wave vector and ν the spin state:

$$|\boldsymbol{\Psi}\rangle = \sum_{\nu} \int \mathrm{d}^3k \; \psi_\nu(\boldsymbol{k}) \, |\boldsymbol{k}, \nu\rangle \tag{1}$$

with:

$$\psi_\nu(\boldsymbol{k}) = \langle \boldsymbol{k}, \nu | \boldsymbol{\Psi} \rangle \tag{2}$$

In general, we choose for the states $|\nu\rangle$ the eigenstates $|+1\rangle, |0\rangle, |-1\rangle$ of the S_z spin component. Here, we shall choose another basis:

$$\begin{aligned} |x\rangle &= (1/\sqrt{2})\Big[|-1\rangle - |+1\rangle\Big] \\ |y\rangle &= (i/\sqrt{2})\Big[|-1\rangle + |+1\rangle\Big] \\ |z\rangle &= |0\rangle \end{aligned} \tag{3}$$

The action of the $\boldsymbol{S}$ components S_x, S_y, S_z on these basis vectors can easily be computed. We must use $S_x = (S_+ + S_-)/2$ and $S_y = -i\,(S_+ - S_-)/2$, as well as the action of

S_+, S_-, S_z on the states $|+1\rangle, |0\rangle, |-1\rangle$ (Chapter VI, relations (C-50)). As an example, we obtain:

$$S_x|x\rangle = \frac{1}{2}(S_+ + S_-)\frac{1}{\sqrt{2}}\Big[|-1\rangle - |+1\rangle\Big] = \frac{\hbar}{2\sqrt{2}}\Big[\sqrt{2}|0\rangle - \sqrt{2}|0\rangle\Big] = 0$$
$$S_x|y\rangle = \frac{i}{2}(S_+ + S_-)\frac{1}{\sqrt{2}}\Big[|-1\rangle + |+1\rangle\Big] = \frac{i\hbar}{2\sqrt{2}}\Big[\sqrt{2}\,|0\rangle + \sqrt{2}\,|0\Big] = i\hbar|z\rangle$$
$$S_x|z\rangle = \frac{1}{2}(S_+ + S_-)|0\rangle = \frac{\hbar}{2}\Big[\sqrt{2}\,|+1\rangle + \sqrt{2}\,|-1\rangle\Big] = -i\hbar|y\rangle \tag{4}$$

These equations, and those similar for the action of S_y and S_z, can be written in a more compact way as:

$$S_a|b\rangle = i\hbar \sum_c \epsilon_{abc}\,|c\rangle \tag{5}$$

where a, b, c are the indices x, y, z and ϵ_{abc} is the three-dimensional completely antisymmetric tensor[1]. Equation (5) also leads to:

$$\langle c|S_a|b\rangle = i\hbar\,\epsilon_{abc} \tag{6}$$

1-b. Average value of the spin angular momentum

Taking (1) into account, the average value of S_a is written:

$$\langle\boldsymbol{\Psi}|S_a|\boldsymbol{\Psi}\rangle = \int \mathrm{d}^3k \int \mathrm{d}^3k' \sum_{b,c} \psi_c^*(\boldsymbol{k}')\langle\boldsymbol{k}', c|S_a|\boldsymbol{k}, b\rangle\psi_b(\boldsymbol{k}) \tag{7}$$

As S_a does not act on the orbital degrees of freedom, described by the variable $\boldsymbol{k}$, we get, taking (6) into account :

$$\langle\boldsymbol{k}', c|S_a|\boldsymbol{k}, b\rangle = i\hbar\,\epsilon_{abc}\,\delta(\boldsymbol{k} - \boldsymbol{k}') \tag{8}$$

Inserting this result in (7) then yields:

$$\langle\boldsymbol{\Psi}|S_a|\boldsymbol{\Psi}\rangle = i\hbar\int \mathrm{d}^3k \sum_{b,c} \epsilon_{abc}\,\psi_c^*(\boldsymbol{k})\psi_b(\boldsymbol{k}) = -i\hbar\int \mathrm{d}^3k \sum_{b,c} \epsilon_{abc}\,\psi_b^*(\boldsymbol{k})\psi_c(\boldsymbol{k}) \tag{9}$$

(using the fact that ϵ_{abc} is antisymmetric). As the a component of the cross product of two vectors $\boldsymbol{V}$ and $\boldsymbol{W}$ is written:

$$(\boldsymbol{V}\times\boldsymbol{W})_a = \sum_{b,c} \epsilon_{abc}\,V_b\,W_c \tag{10}$$

we get:

$$\langle\boldsymbol{\Psi}|S_a|\boldsymbol{\Psi}\rangle = -i\hbar\int \mathrm{d}^3k\,(\boldsymbol{\Psi}^*\times\boldsymbol{\Psi})_a \tag{11}$$

[1] By definition $\epsilon_{abc} = +1$ if abc are xyz or can be deduced from xyz by an even permutation; $\epsilon_{abc} = -1$ if abc is deduced from xyz by an odd permutation; finally $\epsilon_{abc} = 0$ if two of the three indices abc (or all three) are equal.

1-c. Average value of the orbital angular momentum

The orbital angular momentum is written:

$$\boldsymbol{L} = \boldsymbol{R} \times \boldsymbol{P} \tag{12}$$

Its L_a component acts in position space as:

$$L_a = (\boldsymbol{R} \times \boldsymbol{P})_a = \sum_{b,c} \epsilon_{abc}\, R_b\, P_c \Rightarrow -i\hbar \sum_{b,c} \epsilon_{abc} r_b\, \partial_c \qquad \text{with} \qquad \partial_c = \partial/\partial r_c \tag{13}$$

Going to the reciprocal space amounts to performing a spatial Fourier transform. The operators which in $\boldsymbol{r}$ space correspond to a multiplication by r_b and a derivation with respect to r_c, respectively become, in $\boldsymbol{k}$ space, the operators derivation with respect to k_b and multiplication by k_c (multiplied by appropriate factors):

$$r_b \underset{\text{FT}}{\leftrightarrow} i\partial/\partial k_b = i\tilde{\partial}_b \qquad\qquad \partial/\partial r_c = \partial_c \underset{\text{FT}}{\leftrightarrow} ik_c \tag{14}$$

with the notation:

$$\tilde{\partial}_b \equiv \partial/\partial k_b \tag{15}$$

In reciprocal space, the action of the L_a component of the orbital angular momentum is therefore:

$$\begin{aligned} L_a &\Rightarrow -i\hbar \sum_{b,c} \epsilon_{abc}\, (i\tilde{\partial}_b)\, (ik_c) = i\hbar \sum_{b,c} \epsilon_{abc}\, k_c\, \tilde{\partial}_b \\ &= -i\hbar(\boldsymbol{k} \times \boldsymbol{\nabla}_{\boldsymbol{k}})_a \end{aligned} \tag{16}$$

In the last equality of the first line, we have moved $\tilde{\partial}_b$ to the right of k_c, which is allowed since the presence of ϵ_{abc} means that all the terms with equal indices b and c must be zero. To obtain the equality in the second line of (16), we use the anstisymetry of ϵ_{abc} under the exchange of b and c, and relation (10) of the vector product; $\boldsymbol{\nabla}_{\boldsymbol{k}}$ is the gradient with respect of $\boldsymbol{k}$.

We finally compute the average value of L_a in the state $\boldsymbol{\Psi}$. Taking (1) into account, we get:

$$\langle \boldsymbol{\Psi} | L_a | \boldsymbol{\Psi} \rangle = \int \mathrm{d}^3k \int \mathrm{d}^3k' \sum_{d,e} \psi_d^*(\boldsymbol{k}') \langle \boldsymbol{k}', d | L_a | \boldsymbol{k}, e \rangle \psi_e(\boldsymbol{k}) \tag{17}$$

As L_a does not act on the spin degrees of freedom, we must have $d = e$ in the matrix element on the right-hand side, which yields, using the differential form (16) for L_a:

$$\begin{aligned} \langle \boldsymbol{\Psi} | L_a | \boldsymbol{\Psi} \rangle &= \int \mathrm{d}^3k \int \mathrm{d}^3k' \sum_{d} \psi_d^*(\boldsymbol{k}') \langle \boldsymbol{k}', d | L_a | \boldsymbol{k}, d \rangle \psi_d(\boldsymbol{k}) \\ &= -i\hbar \int \mathrm{d}^3k \sum_{b,c,d} \psi_d^*(\boldsymbol{k})\; \epsilon_{abc}\; k_b\; \tilde{\partial}_c\; \psi_d(\boldsymbol{k}) \\ &= -i\hbar \int \mathrm{d}^3k \sum_{d} \psi_d^*(\boldsymbol{k})\, (\boldsymbol{k} \times \boldsymbol{\nabla}_{\boldsymbol{k}})_a\, \psi_d(\boldsymbol{k}) \end{aligned} \tag{18}$$

Finally, adding (11) and (18), we obtain the following expression for the average value of the particle total angular momentum in the state $\boldsymbol{\Psi}$:

$$\langle \boldsymbol{\Psi} | \boldsymbol{J} | \boldsymbol{\Psi} \rangle = -i\hbar \int \mathrm{d}^3 k \left\{ \underbrace{\boldsymbol{\Psi}^*(\boldsymbol{k}) \times \boldsymbol{\Psi}(\boldsymbol{k})}_{\text{spin}} + \sum_d \underbrace{\psi_d^*(\boldsymbol{k})\, (\boldsymbol{k} \times \boldsymbol{\nabla}_{\boldsymbol{k}})\, \psi_d(\boldsymbol{k})}_{\text{orbital}} \right\} \tag{19}$$

2. Angular momentum of free classical radiation as a function of normal variables

We now show that the classical calculation of the radiation angular momentum presents a certain analogy with the results of § 1.

2-a. Calculation in position space

Relation (A-53) of Chapter XVIII yields for the total angular momentum of free radiation (in the absence of particles):

$$\boldsymbol{J} = \varepsilon_0 \int \mathrm{d}^3 r \Big[\boldsymbol{r} \times [\boldsymbol{E}_\perp(\boldsymbol{r}) \times \boldsymbol{B}(\boldsymbol{r})] \Big] \tag{20}$$

Let us replace $\boldsymbol{B}$ by $\boldsymbol{\nabla} \times \boldsymbol{A}$ and use the triple product expansion $\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \cdot \boldsymbol{c})\, \boldsymbol{b} - (\boldsymbol{a} \cdot \boldsymbol{b})\, \boldsymbol{c}$. We obtain, keeping the right order between $\boldsymbol{\nabla}$ and $\boldsymbol{A}$:

$$\boldsymbol{E} \times [\boldsymbol{\nabla} \times \boldsymbol{A}] = \sum_d E_d \boldsymbol{\nabla} A_d - (\boldsymbol{E} \cdot \boldsymbol{\nabla}) \boldsymbol{A} \tag{21}$$

where E_d is the component of $\boldsymbol{E}_\perp(\boldsymbol{r})$ on the Ox, Oy or Oz axis, labeled by the index d. Inserting (21) in (20) leads to:

$$\boldsymbol{J} = \epsilon_0 \int \mathrm{d}^3 r \left\{ \sum_d E_d (\boldsymbol{r} \times \boldsymbol{\nabla}) A_d - \boldsymbol{r} \times (\boldsymbol{E} \cdot \boldsymbol{\nabla}) \boldsymbol{A} \right\} \tag{22}$$

Consider first the contribution $\boldsymbol{J}^{(1)}$ of the second term in (22). Its $J_a^{(1)}$ component is written:

$$\begin{aligned} J_a^{(1)} &= -\epsilon_0 \int \mathrm{d}^3 r \left\{ \boldsymbol{r} \times (\boldsymbol{E} \cdot \boldsymbol{\nabla}) \boldsymbol{A} \right\}_a \\ &= -\epsilon_0 \int \mathrm{d}^3 r \sum_{b,c,d} \epsilon_{abc}\, r_b\, E_d\, \partial_d\, A_c \end{aligned} \tag{23}$$

We now move r_b to the right of ∂_d, using:

$$r_b\, \partial_d = \partial_d\, r_b - \delta_{bd} \tag{24}$$

The contribution of the term $-\delta_{bd}$ to (23) yields:

$$\epsilon_0 \int \mathrm{d}^3 r \sum_{b,c} \epsilon_{abc}\, E_b\, A_c = \epsilon_0 \int \mathrm{d}^3 r\, (\boldsymbol{E} \times \boldsymbol{A})_a \tag{25}$$

As for the contribution of the term $\partial_d\, r_b$ to (23), we perform an integration by parts. The contribution of the integrated term yields a zero surface integral if the fields decrease fast enough at infinity. We obtain for the contribution of the term $\partial_d\, r_b$ to (23) :

$$-\epsilon_0 \int \mathrm{d}^3 r \sum_{b,c,d} \epsilon_{abc}\, E_d\, \partial_d\, (r_b\, A_c) = 0 + \epsilon_0 \int \mathrm{d}^3 r \sum_{b,c,d} \epsilon_{abc}\, r_b\, A_c\, \partial_d\, E_d \tag{26}$$

In the last term of (26), we note the quantity

$$\sum_d \partial_d\, E_d = \boldsymbol{\nabla} \cdot \boldsymbol{E} \tag{27}$$

which is equal to zero since the electric field, in the absence of sources, is purely transverse (and hence of zero divergence). This term therefore disappears.

Finally, the average value of $\boldsymbol{J}$ is the sum of (25) and of the first term of (22):

$$\boldsymbol{J} = \epsilon_0 \int \mathrm{d}^3 r \left\{ \underbrace{\boldsymbol{E}(\boldsymbol{r}) \times \boldsymbol{A}(\boldsymbol{r})}_{\text{spin}} + \sum_d \underbrace{E_d(\boldsymbol{r})(\boldsymbol{r} \times \boldsymbol{\nabla}) A_d(\boldsymbol{r})}_{\text{orbital}} \right\} \tag{28}$$

2-b. Reciprocal space

Expression (28) for $\boldsymbol{J}$ can be rewritten as a function of the Fourier transforms of the fields $\boldsymbol{E}(\boldsymbol{r})$ and $\boldsymbol{A}(\boldsymbol{r})$. Using the Parseval-Plancherel equality (Appendix I, § 2-c) and relations (14), we get:

$$\boldsymbol{J} = \epsilon_0 \int \mathrm{d}^3 k \left\{ \underbrace{\tilde{\boldsymbol{E}}^*(\boldsymbol{k}) \times \tilde{\boldsymbol{A}}(\boldsymbol{k})}_{\text{spin}} + \sum_d \underbrace{\tilde{E}_d^*(\boldsymbol{k})(\boldsymbol{k} \times \boldsymbol{\nabla}_{\boldsymbol{k}}) \tilde{A}_d(\boldsymbol{k})}_{\text{orbital}} \right\} \tag{29}$$

We now use expressions (B-22a) and (B-22b) of Chapter XVIII to obtain the fields $\tilde{\boldsymbol{E}}(\boldsymbol{k})$ and $\tilde{\boldsymbol{A}}(\boldsymbol{k})$ as a function of the normal variables:

$$\tilde{\boldsymbol{E}}(\boldsymbol{k}) = \frac{i\omega}{2\mathcal{N}(k)} \left[\boldsymbol{\alpha}(\boldsymbol{k}) - \boldsymbol{\alpha}^*(-\boldsymbol{k})\right] \tag{30a}$$

$$\tilde{\boldsymbol{A}}(\boldsymbol{k}) = \frac{1}{2\mathcal{N}(k)} \left[\boldsymbol{\alpha}(\boldsymbol{k}) + \boldsymbol{\alpha}^*(-\boldsymbol{k})\right] \tag{30b}$$

where $\mathcal{N}(k)$ was determined by relation (A-3) of Chapter XIX:

$$\mathcal{N}(k) = \sqrt{\frac{\epsilon_0 \omega}{2\hbar}} \tag{31}$$

Inserting (30) into (29), we get:

$$\begin{aligned} J_a = -\, i\frac{\hbar}{2} \int \mathrm{d}^3 k \sum_{b,c} \Big\{ & \left[\alpha_b^*(\boldsymbol{k}) - \alpha_b(-\boldsymbol{k})\right] \epsilon_{abc} \left[\alpha_c(\boldsymbol{k}) + \alpha_c^*(-\boldsymbol{k})\right] \\ & + \sum_{b,c,d} \left[\alpha_d^*(\boldsymbol{k}) - \alpha_d(-\boldsymbol{k})\right] \epsilon_{abc}\, k_b\, \tilde{\partial}_c \left[\alpha_d(\boldsymbol{k}) + \alpha_d^*(-\boldsymbol{k})\right] \Big\} \end{aligned} \tag{32}$$

Each line in (32) contains four terms: two of these terms include either $\boldsymbol{\alpha}^*$ twice for the first one, or $\boldsymbol{\alpha}$ twice for the second – both will be shown below to be equal to zero; the other two terms contain either $\boldsymbol{\alpha}^*$ once or $\boldsymbol{\alpha}$ once – we show below that they are equal. We finally obtain:

$$\boldsymbol{J} = -i\hbar \int \mathrm{d}^3k \left\{ \underbrace{\boldsymbol{\alpha}^*(\boldsymbol{k}) \times \boldsymbol{\alpha}(\boldsymbol{k})}_{\text{spin}} + \sum_d \underbrace{\alpha_d^*(\boldsymbol{k})\, (\boldsymbol{k} \times \boldsymbol{\nabla}_{\boldsymbol{k}})\, \alpha_d(\boldsymbol{k})}_{\text{orbital}} \right\} \tag{33}$$

This expression has the same form as (19): the angular momentum is the sum of a spin term and an orbital term involving spatial derivatives. This result confirms that the normal variable $\boldsymbol{\alpha}(\boldsymbol{k})$ can be regarded as the wave function in reciprocal space of the photon field, and that the photon is indeed a spin 1 particle. This result also gives the explicit expressions of the radiation spin angular momentum (first term in the bracket of (33)) and orbital angular momentum (second term).

For a massive particle, we know (Chapitre VI, § D-1-a) that in spherical (or cylindrical) coordinates, the action of the Oz angular momentum component corresponds to a derivation with respect to the azimuthal angle φ:

$$L_z \Rightarrow -i\hbar(k_x\, \tilde{\partial}_y - k_y\, \tilde{\partial}_x) = \frac{\hbar}{i} \frac{\partial}{\partial \varphi} \tag{34}$$

This result simply comes from a calculation of partial derivatives; it is thus also valid for a field.

Computation of the various terms appearing in equation (32)

Consider, in the first line of (32), the terms involving a product of two $\boldsymbol{\alpha}^*$ or two $\boldsymbol{\alpha}$, for example:

$$-i(\hbar/2) \int \mathrm{d}^3k \sum_{b,c} [\alpha_b^*(\boldsymbol{k})\, \epsilon_{abc}\, \alpha_c^*(-\boldsymbol{k})] \tag{35}$$

Changing $\boldsymbol{k}$ into $-\boldsymbol{k}$, inverting the indices b and c, and using $\epsilon_{acb} = -\epsilon_{abc}$, we can show that (35) is equal to its opposite, and hence must be zero. The same approach, followed for the term:

$$+i(\hbar/2) \int \mathrm{d}^3k \sum_{b,c} [\alpha_b(-\boldsymbol{k})\, \epsilon_{abc}\, \alpha_c^*(-\boldsymbol{k})] \tag{36}$$

shows that this term is equal to:

$$-i(\hbar/2) \int \mathrm{d}^3k \sum_{b,c} [\alpha_c(\boldsymbol{k})\, \epsilon_{abc}\, \alpha_b^*(\boldsymbol{k})] \tag{37}$$

which is identical to the other term on the first line of (32) containing one $\boldsymbol{\alpha}^*$ and one $\boldsymbol{\alpha}$, provided we change the relative order in which the term $\alpha_b^*(\boldsymbol{k})$ and $\alpha_c(\boldsymbol{k})$ appear. In classical theory, these quantities are numbers, and hence commute: their order does not matter. It is however useful to keep track of that order in order to obtain an expression still valid when, upon quantization, the $\boldsymbol{\alpha}^*$ and $\boldsymbol{\alpha}$ will be replaced by the non-commuting creation and annihilation operators.

This computation can be extended to the terms on the second line of (32). In addition to changing $\boldsymbol{k}$ into $-\boldsymbol{k}$, we must also perform an integration by parts. The integrated term,

which yields a surface integral, is zero if the fields tend to zero fast enough at infinity. Added to it is a contribution that shows that the terms containing two $\boldsymbol{\alpha}^*$ or two $\boldsymbol{\alpha}$ are equal to their opposite, and hence equal to zero. On the other hand, the integration by parts shows that the two terms containing one $\boldsymbol{\alpha}^*$ and one $\boldsymbol{\alpha}$ are equal if the order of the $\boldsymbol{\alpha}^*$ and $\boldsymbol{\alpha}$ can be switched. In the case where the order of the $\boldsymbol{\alpha}^*$ and $\boldsymbol{\alpha}$ is not taken into account, we obtain expression (33).

2-c. Difference between the angular momenta of massive particles and of radiation

In spite of the strong analogy between equations (19) and (33), we should not forget an important difference between the two angular momenta, arising from the fact that the normal variables $\boldsymbol{\alpha}(\boldsymbol{k})$ are transverse. Maxwell's equation $\boldsymbol{\nabla}\cdot\boldsymbol{E}=0$ for the free field does require $\boldsymbol{\alpha}(\boldsymbol{k})$ to always be perpendicular to the wave vector $\boldsymbol{k}$:

$$\boldsymbol{k}\cdot\boldsymbol{\alpha}(\boldsymbol{k})=0 \qquad\qquad \forall \quad \boldsymbol{k} \tag{38}$$

while the wave function $\boldsymbol{\Psi}(\boldsymbol{k})$ of a massive particle in the reciprocal space is not necessarily perpendicular to $\boldsymbol{k}$. Another difference, of course, is that the norm of this wave function does not have any particular physical meaning (it can arbitrarily be put equal to unity), while changing the norm of the normal variables of the field changes its amplitude.

3. Discussion

3-a. Spin angular momentum of radiation

The spin angular momentum, first term on the right-hand side of (33), can be written as:

$$(S_R)_a=-i\hbar\int \mathrm{d}^3k\ \sum_{b,c}\alpha_b^*(\boldsymbol{k})\,\varepsilon_{b,c}\,\alpha_c(\boldsymbol{k})=-i\hbar\int \mathrm{d}^3k\ \left[\boldsymbol{\alpha}^*(\boldsymbol{k})\times\boldsymbol{\alpha}(\boldsymbol{k})\right]_a \tag{39}$$

Instead of using the components $\boldsymbol{\alpha}^*(\boldsymbol{k})$ and $\boldsymbol{\alpha}(\boldsymbol{k})$ on a basis of three vectors $\boldsymbol{e}_x$, $\boldsymbol{e}_y$, $\boldsymbol{e}_z$ independent of the wave vector $\boldsymbol{k}$ direction, we can choose a basis of three vectors $\boldsymbol{\varepsilon}_1(\boldsymbol{k})$, $\boldsymbol{\varepsilon}_2(\boldsymbol{k})$, $\boldsymbol{\varepsilon}_3(\boldsymbol{k})=\boldsymbol{\kappa}=\boldsymbol{k}/k$, including the unit vector $\boldsymbol{\kappa}$ along $\boldsymbol{k}$ and two other vectors $\boldsymbol{\varepsilon}_1(\boldsymbol{k})$ and $\boldsymbol{\varepsilon}_2(\boldsymbol{k})$, orthogonal to each other and to $\boldsymbol{\kappa}$, and forming a right-handed reference frame. As the normal variables $\boldsymbol{\alpha}^*(\boldsymbol{k})$ and $\boldsymbol{\alpha}(\boldsymbol{k})$ are transverse, their components on $\boldsymbol{\kappa}$ are zero. In addition, we introduce the two complex linear combinations of $\boldsymbol{e}_1(\boldsymbol{k})$ and $\boldsymbol{e}_2(\boldsymbol{k})$:

$$\begin{aligned}\boldsymbol{\varepsilon}_+(\boldsymbol{k})&=-\left[\boldsymbol{\varepsilon}_1(\boldsymbol{k})+i\boldsymbol{\varepsilon}_2(\boldsymbol{k})\right]/\sqrt{2}\\ \boldsymbol{\varepsilon}_-(\boldsymbol{k})&=+\left[\boldsymbol{\varepsilon}_1(\boldsymbol{k})-i\boldsymbol{\varepsilon}_2(\boldsymbol{k})\right]/\sqrt{2}\end{aligned} \tag{40}$$

corresponding to right and left circular polarizations with respect to the $\boldsymbol{k}$ direction. The transverse normal variables $\boldsymbol{\alpha}^*(\boldsymbol{k})$ and $\boldsymbol{\alpha}(\boldsymbol{k})$ can be expanded on these two vectors:

$$\begin{aligned}\boldsymbol{\alpha}(\boldsymbol{k})&=\alpha_+(\boldsymbol{k})\boldsymbol{e}_+(\boldsymbol{k})+\alpha_-(\boldsymbol{k})\boldsymbol{e}_-(\boldsymbol{k})\\ \boldsymbol{\alpha}^*(\boldsymbol{k})&=\alpha_+^*(\boldsymbol{k})\boldsymbol{e}_+^*(\boldsymbol{k})+\alpha_-^*(\boldsymbol{k})\boldsymbol{e}_-^*(\boldsymbol{k})\end{aligned} \tag{41}$$

Using these two expansions, we compute the cross product $\boldsymbol{\alpha}^*(\boldsymbol{k}) \times \boldsymbol{\alpha}(\boldsymbol{k})$. Since:

$$\begin{aligned}
\boldsymbol{\varepsilon}_+^* \times \boldsymbol{\varepsilon}_+ &= \frac{1}{2}(\boldsymbol{\varepsilon}_1 - i\boldsymbol{\varepsilon}_2) \times (\boldsymbol{\varepsilon}_1 + i\boldsymbol{\varepsilon}_2) = \frac{i}{2}(\boldsymbol{\varepsilon}_1 \times \boldsymbol{\varepsilon}_2 - \boldsymbol{\varepsilon}_2 \times \boldsymbol{\varepsilon}_1) = i\boldsymbol{\kappa} \\
\boldsymbol{\varepsilon}_-^* \times \boldsymbol{\varepsilon}_- &= \frac{1}{2}(\boldsymbol{\varepsilon}_1 + i\boldsymbol{\varepsilon}_2) \times (\boldsymbol{\varepsilon}_1 - i\boldsymbol{\varepsilon}_2) = \frac{-i}{2}(\boldsymbol{\varepsilon}_1 \times \boldsymbol{\varepsilon}_2 - \boldsymbol{\varepsilon}_2 \times \boldsymbol{\varepsilon}_1) = -i\boldsymbol{\kappa} \\
\boldsymbol{\varepsilon}_+^* \times \boldsymbol{\varepsilon}_- &= \frac{-1}{2}(\boldsymbol{\varepsilon}_1 - i\boldsymbol{\varepsilon}_2) \times (\boldsymbol{\varepsilon}_1 - i\boldsymbol{\varepsilon}_2) = 0 = \boldsymbol{\varepsilon}_-^* \times \boldsymbol{\varepsilon}_+
\end{aligned} \tag{42}$$

we get:

$$\boldsymbol{S}_R = \int \mathrm{d}^3k \left\{ \left[\alpha_+^*(\boldsymbol{k})\alpha_+(\boldsymbol{k})\right] \hbar\boldsymbol{\kappa} - \left[\alpha_-^*(\boldsymbol{k})\alpha_-(\boldsymbol{k})\right] \hbar\boldsymbol{\kappa} \right\} \tag{43}$$

The form of this expression, "diagonal with respect to the spin variables", has a clear physical significance: to each plane wave with wave vector $\boldsymbol{k}$ and a right (left) polarization with respect to $\boldsymbol{k}$, correspond photons of momentum $\hbar\boldsymbol{k}$ and spin angular momentum $+\hbar$ $(-\hbar)$ along the direction $\boldsymbol{\kappa}$ of $\boldsymbol{k}$. Upon quantization, when the normal variables $\boldsymbol{\alpha}^*(\boldsymbol{k})$ and $\boldsymbol{\alpha}(\boldsymbol{k})$ are replaced by creation and annihilation operators, expression (43) becomes:

$$\hat{\boldsymbol{S}}_R = \int \mathrm{d}^3k \left\{ \left[\hat{a}_+^\dagger(\boldsymbol{k})\hat{a}_+(\boldsymbol{k})\right] \hbar\boldsymbol{\kappa} - \left[\hat{a}_-^\dagger(\boldsymbol{k})\hat{a}_-(\boldsymbol{k})\right] \hbar\boldsymbol{\kappa} \right\} \tag{44}$$

Operator $\hat{a}_+^\dagger(\boldsymbol{k})$ creates a photon with momentum $\hbar\boldsymbol{k}$ and spin angular momentum $+\hbar$ along the direction $\boldsymbol{\kappa}$ of $\boldsymbol{k}$; operator $\hat{a}_+(\boldsymbol{k})$ annihilates that photon, and operator $\hat{a}_+^\dagger(\boldsymbol{k})\hat{a}_+(\boldsymbol{k})$ corresponds to the number of photons in that mode. An analogous definition applies to the second term of (44) with a change of sign for the angular momentum.

Helicity

These results, which arise from the transverse character of free radiation, lead us to introduce the so-called "*helicity*". It is the projection of the photon spin angular momentum onto the direction $\boldsymbol{\kappa}$ of the wave vector $\boldsymbol{k}$, equal to $+1$ for photons with a right circular polarization with respect to $\boldsymbol{\kappa}$, and -1 for photons with a left circular polarization. The transverse character of the free radiation field forbids the photons to have zero helicity. Note also that helicity is a *pseudoscalar*: upon reflection in space, the polar vector $\boldsymbol{\kappa}$ changes sign whereas the spin vector $\boldsymbol{S}$, an axial vector, remains unchanged. Consequently, the scalar product of $\boldsymbol{\kappa}$ and $\boldsymbol{S}$ changes sign (as opposed to a scalar).

3-b. Experimental evidence of the radiation spin angular momentum

Consider a plane wave of wave vector $\boldsymbol{k}$ and polarization $\boldsymbol{\varepsilon}$. Using the expressions for the electric field $\boldsymbol{E}(\boldsymbol{r})$ and the vector potential $\boldsymbol{A}(\boldsymbol{r})$ given in Chapter XVIII, one can easily compute the two terms of equation (28) yielding, in position space, the radiation spin angular momentum and orbital angular momentum[2]. The result is that the radiation orbital angular momentum – second term of (28) – is always zero, whatever the

[2]These are the two terms that, transformed into the reciprocal space, and after the introduction of the normal variables, yield the two terms on the right-hand side of equation (33); by comparison with the expression of the angular momentum of a spin 1 particle, these terms have been interpreted as the two components of the radiation angular momentum.

polarization $\boldsymbol{\varepsilon}$ is. As for the spin angular momentum – first term of (28) – it is zero for a linear polarization, but different from zero for a circular polarization, with opposite signs for the right and left circular polarizations. This validates, in the simple plane wave case, the general conclusions of the previous paragraph.

Such a result suggests sending a linearly polarized light beam through a quarter-wave plate. Assuming the plate transforms the incident linear polarization into a right (left) circular polarization, the incident photons have a zero spin angular momentum before they go through the plate, and equal to $+\hbar$ $(-\hbar)$ as they come out of the plate. The radiation spin angular momentum thus changes as beam goes through the plate, and this must be accompanied by a change, in the opposite direction, of the plate's angular momentum. Suspending the plate by a thin torsion fiber, one should observe a rotation of the plate induced by the incident radiation, in a direction opposite to that of the circular polarization of the outgoing beam. This experiment, suggested by A. Kastler [30] was performed by R. Beth [31], confirming the existence of angular momentum transfer.

Comment

A paradox arises when computing, again for a plane wave, the angular momentum of the radiation, given by equation (20). In a plane wave, the Poynting vector $\boldsymbol{\Pi}(\boldsymbol{r}) = \boldsymbol{E}(\boldsymbol{r}) \times \boldsymbol{B}(\boldsymbol{r})$ is always parallel to the wave vector $\boldsymbol{k}$ at any point $\boldsymbol{r}$, and for any polarization. The integral over the entire space of $\boldsymbol{r} \times \boldsymbol{\Pi}(\boldsymbol{r})$ must then be zero. As the orbital angular momentum is also zero, it seems that the spin angular momentum should also be zero, whatever the polarization is. This paradox arises because infinite plane waves do not exist in the physical world: any real light beam has a finite spatial extension. The authors of [32] (see also [33]) show that the circular polarization at the center of the beam changes when the field amplitude changes around the edge of the beam. Taking this effect into account quantitatively confirms the result obtained above, namely that the beam spin angular momentum is equal to the sum of the angular momenta $\pm\hbar$ of that beam's photons.

3-c. Orbital angular momentum of radiation

An important difference between the radiation orbital and spin angular momenta is clearly seen in expression (28): the definition of orbital angular momentum involves a reference point O, since the vector $\boldsymbol{r}$, defined with respect to that point, explicitly appears in the expression for the orbital angular momentum. This is not the case for the spin angular momentum which, for this reason, is sometimes called "intrinsic" angular momentum. There are actually at least two cases where the choice of the point O is obvious, cases that we now analyze.

α. *Multipolar waves*

When studying the radiation emitted or absorbed by an atom or a nucleus between two discrete states, a natural choice for analyzing the exchanges of angular momentum between the system internal variables and the photons is the center of mass of that atom or that nucleus. In the next complement C_{XIX}, we study for example the exchange of angular momentum between photons and the internal variables of an atom in a particular case: an electric dipole transition, in the long wavelength approximation. Consequently, the expressions describing the photon absorption only involve the radiation polarization variables, and hence only the photon spin angular momentum; the radiation orbital

angular momentum does not actually play any role[3].

There are other transitions, especially for atomic nuclei, where the variation of the internal angular momentum between the two transition states is larger than or equal to 2; the photon spin angular momentum, equal to 1, can then no longer ensure the conservation of the angular momentum. Radiation states having a total angular momentum J larger than 1 must come into play, which implies a contribution from the radiation orbital angular momentum. Waves corresponding to such states are called "multipolar waves".

The simplest way to build multipolar waves having a total angular momentum characterized by the quantum numbers J and M is to associate a spherical harmonic $Y_\ell^{m_\ell}(\boldsymbol{\kappa})$ with a spin $S = 1$; the spherical harmonic is an eigenfunction of $\boldsymbol{L}^2$ and L_z with eigenvalues $\ell(\ell+1)\hbar^2$ and $m_\ell\hbar$; the spin $S = 1$ has three eigenstates $|S, m_S\rangle$, with $m_S = +1, 0, -1$, isomorphic to the three polarization states $\boldsymbol{e}_+ = -(\boldsymbol{e}_x + i\boldsymbol{e}_y)/\sqrt{2}$, $\boldsymbol{e}_z$, $\boldsymbol{e}_- = (\boldsymbol{e}_x - i\boldsymbol{e}_y)/\sqrt{2}$. We therefore obtain a *vector spherical harmonic*:

$$\boldsymbol{Y}_{J,\ell,1}^M(\boldsymbol{\kappa}) = \sum_{m_\ell}\sum_{m_S} \langle JM|\ell 1 m_\ell m_S\rangle \; Y_\ell^{m_\ell}(\boldsymbol{\kappa})\boldsymbol{e}_{m_S} \tag{45}$$

which is an eigenfunction of $\boldsymbol{J}^2$, $\boldsymbol{L}^2$, J_z, with eigenvalues $J(J+1)\hbar^2$, $\ell(\ell+1)\hbar^2$, $M\hbar$. In this equation, the first term on the right-hand side is a Clebsch-Gordan coefficient (Chapter X, § C-4-c), J can take one of the three values $J = \ell - 1, J = \ell, J = \ell + 1$ and $M = m_\ell + m_S$.

A difficulty is that the vector spherical harmonics are not all transverse functions and hence cannot be used as a basis of normal transverse functions for expanding the radiation field. For a given value of J, one can nevertheless build linear superpositions of vector spherical harmonics $\boldsymbol{Y}_{J,\ell,1}^M(\boldsymbol{\kappa})$ with $\ell = J, \ell = J \pm 1$ that are transverse and have, in addition, a well defined parity $\pi = \pm 1$. Each vector spherical harmonic, which only depends on the direction $\boldsymbol{\kappa}$ of $\boldsymbol{k}$, can also be multiplied by $\delta(k - \omega_0/c)$, hence yielding a function that is also an eigenfunction of the energy, with eigenvalue $\hbar\omega_0$. These functions form a possible basis of normal transverse variables for expanding the field; they are characterized by four quantum numbers: the energy $\hbar c k_0$, the total angular momentum J, M, and the parity π. They are called electric (for $\pi = -1$) or magnetic (for $\pi = +1$) multipolar waves. As this book is limited to the study of electric or magnetic dipole transitions, we do not give here the general expressions for multipolar waves. More details can be found in complement B_I of [16] and in [34].

β. Beams with cylindrical symmetry around one axis

It often happens that the beams under study have a cylindrical symmetry. This is the case, for example, for Gaussian beams propagating along a z axis, and whose transverse sections are circular. If the reference point O is taken on the axis, the beam symmetry causes the orbital angular momentum to be necessarily along the z axis and to have the same value regardless of the position of O along this axis. If the reference point O is taken outside this axis, the orbital angular momentum will change, but not its component L_z, which exhibits an intrinsic character.

[3] This orbital angular momentum may, however, come into play during the angular momentum exchanges with the atom's external variables for certain types of light beams, the Laguerre-Gaussian light beams (§ 3-b of complement C_{XIX}).

A particularly interesting case concerns the Laguerre-Gaussian beam (LG) whose field has an $\exp(im\varphi)$ dependence with respect to the azimuth angle φ that defines the direction of a point in the plane perpendicular to the beam axis. The cylindrical symmetry is preserved since a rotation of the beam of an angle φ_0 around the z axis yields the same field to within a global phase factor $\exp(-im\varphi_0)$. Relation (34) then shows that the Oz component of the orbital angular momentum of each photon of the LG beam is $m\hbar$.

Consider an LG beam propagating along the z axis, with a k wave vector along that axis. The phase at a point with cylindrical coordinates z, ρ, φ is that of $\exp i(kz + m\varphi)$. For an ordinary Gaussian beam (for which $m = 0$) the surfaces of constant phase are, in the vicinity of the focal point, planes perpendicular to the z axis. When $m \neq 0$, z and φ must both vary for the phase to remain constant, following the relation $k\mathrm{d}z + m\mathrm{d}\varphi = 0$. The surfaces of constant phase are therefore helicoidal surfaces spanned by a half-line perpendicular to the z axis, starting from this axis, and which rotates of an angle $-2\pi/m$ when z increases by λ. It is not surprising that under such conditions the field has a non-zero orbital angular momentum. Note also that the field must be zero on the z axis (otherwise its phase would vary discontinuously upon crossing that axis).

In conclusion, we showed in this complement that there are two types of radiation angular momenta, the spin angular momentum and the orbital angular momentum, and we studied their properties. The photon can be viewed as a spin 1 particle, except for the fact that it only has two (instead of three) internal states, with respective heliticity $+1$ and -1. We shall see in the next complement how the interactions between radiation and atoms permit transferring angular momentum from the first to the others.

Complement C_{XIX}

Angular momentum exchange between atoms and photons

1 Transferring spin angular momentum to internal atomic variables 2056

1-a Electric dipole transitions 2056

1-b Polarization selection rules 2056

1-c Conservation of total angular momentum 2058

2 Optical methods 2058

2-a Double resonance method 2059

2-b Optical pumping 2062

2-c Original features of these methods 2064

3 Transferring orbital angular momentum to external atomic variables 2065

3-a Laguerre-Gaussian beams 2065

3-b Field expansion on Laguerre-Gaussian modes 2065

Introduction

Angular momentum exchanges between atoms and photons are at the base of several experimental methods that played an important role in atomic physics and laser spectroscopy. The aim of this complement is to analyze the selection rules appearing in the photon absorption and emission processes by an atom; they express the conservation of the total angular momentum of the system atom + radiation during these processes.

We shall mainly focus on the transfer of angular momentum to the atomic internal (electronic) variables (§ 1). The polarization of the field, characterizing the spin angular momentum of that field (Complement B_{XIX}), then plays an essential role. For the absorption process, we will establish the selection rules relating the field polarization to the variation of the "magnetic quantum number" m characterizing the projection of the total internal angular momentum of the atom on a given axis. Two important applications of these selection rules will be described in § 2: the double resonance method, and optical pumping. We shall see that the proper choice of the polarization of the exciting light beam, and of both the direction and polarization of the detected light emitted by the excited atoms, allows controlling the atomic Zeeman sublevels that can be populated and detected by light. We shall emphasize in § 2 the importance of this selectivity. The transfer of the radiation orbital angular momentum to the atomic external variables will be briefly addressed in § 3.

1. Transferring spin angular momentum to internal atomic variables

1-a. Electric dipole transitions

We shall limit ourselves to the case where the transitions between the atomic internal states are electric dipole transitions. As seen in Chapter XIX (§ C-4), the interaction Hamiltonian between atom and radiation can then be written in the form:

$$\hat{H}_I = -\hat{\boldsymbol{D}} \cdot \hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}}) \tag{1}$$

where $\hat{\boldsymbol{D}}$ is the operator associated with the atomic electric dipole moment and $\hat{\boldsymbol{E}}_\perp(\hat{\boldsymbol{R}})$ is the radiation transverse electric field operator at point $\hat{\boldsymbol{R}}$, the position of the atom's center of mass. Note that all the results established in this complement are still valid for magnetic dipole transitions; one must simply replace $\hat{\boldsymbol{D}}$ by the atomic magnetic dipole moment operator $\hat{\boldsymbol{M}}$ and $\hat{\boldsymbol{E}}_\perp$ by the radiation magnetic field $\hat{\boldsymbol{B}}$ operator. For the transitions where one photon is absorbed, the expressions (B-3) and (B-4) of Chapter XIX of the fields $\hat{\boldsymbol{E}}_\perp$ and $\hat{\boldsymbol{B}}$ can be replaced by the part containing only destruction operators, called "positive frequency component" (*cf.* § A-3 of Chapiter XX) and denoted $\hat{\boldsymbol{E}}_\perp^{(+)}$ and $\hat{\boldsymbol{B}}^{(+)}$. For the transitions where one photon is emitted, these fields can be replaced their components $\hat{\boldsymbol{E}}_\perp^{(-)}$ and $\hat{\boldsymbol{B}}^{(-)}$ containing only creation operators, called "negative frequency components".

Replacing $\hat{\boldsymbol{E}}_\perp^{(+)}$ and $\hat{\boldsymbol{B}}^{(+)}$ by their plane wave expansions, the internal variables only appear in the scalar products of $\hat{\boldsymbol{D}}$, or $\hat{\boldsymbol{M}}$, with the polarization vector $\boldsymbol{\varepsilon}$ of the plane wave $\boldsymbol{k}$. We shall assume in § 1 and § 2 that all the incident radiation states are plane waves, or linear superpositions of plane waves with wave vectors having directions very close to an optical axis (paraxial approximation). These waves are supposed to have the same polarization $\boldsymbol{\varepsilon}$, which means that the beam angular aperture must be sufficiently small. The transitions between the internal atomic states[1] g and e we shall consider are thus entirely characterized by the matrix elements $\langle e|\boldsymbol{\varepsilon} \cdot \hat{\boldsymbol{D}}|g\rangle$.

1-b. Polarization selection rules

We start with the simple case of an atom with a single electron, and a transition between a ground state g with an orbital angular momentum $\ell = 0$ and an excited state e with an angular momentum $\ell = 1$. We assume the radiation has a right circular polarization, noted σ_+, with respect to an axis noted z: this means that the radiation electric field rotates around that axis *following the right-hand rule*, at the angular frequency ω. We have:

$$\boldsymbol{\varepsilon} = -\frac{1}{\sqrt{2}}(\boldsymbol{e}_x + i\boldsymbol{e}_y) \tag{2}$$

Since $\hat{\boldsymbol{D}} = q\hat{\boldsymbol{r}}$, where q is the electron charge and $\hat{\boldsymbol{r}}$ its position with respect to the nucleus, we have:

$$\boldsymbol{\varepsilon} \cdot \hat{\boldsymbol{D}} = -\frac{q}{\sqrt{2}}(x + iy) = -\frac{q}{\sqrt{2}} r \sin\theta \exp(i\varphi) \tag{3}$$

[1] In this complement and as is usually done in the literature, we shall note g the ground state and e the excited state (instead of using our previous notation of a and b for the two atomic levels).

where r, θ, φ are the spherical coordinates of the electron with respect to the nucleus. Starting from a ground state with a magnetic quantum number $m_g = 0$ (defined with respect to the z axis) and with no dependence on φ, the excitation with σ_+ polarized light creates an excited state wave function that now varies with φ as $\exp(i\varphi)$. This wave function is an eigenfunction of the component $\hat{L}_z = (\hbar/i)\partial/\partial\varphi$ of the orbital angular momentum, with eigenvalue $+1$ (it corresponds to the state $m_e = +1$). In a similar way, an excitation with a σ_- polarization, for which $\boldsymbol{\varepsilon} = \frac{1}{\sqrt{2}}(\boldsymbol{e}_x - i\boldsymbol{e}_y)$, transfers the atom into the state $m_e = -1$. Consider finally an excitation with a π polarized light, for which $\boldsymbol{\varepsilon} = \boldsymbol{e}_z$: the electric field has a linear polarization parallel[2] to the Oz axis. We then have $\boldsymbol{\varepsilon} \cdot \hat{\boldsymbol{D}} = qz = qr\cos\theta$, with no φ dependence; the excitation brings the atom to the state $m_e = 0$. To sum up, the polarization selection rules for a transition $\ell_g = 0 \longrightarrow \ell_e = 1$ are given by:

$$\sigma_+ \longrightarrow m_e = +1, \qquad \pi \longrightarrow m_e = 0 \qquad \sigma_- \longrightarrow m_e = -1. \tag{4}$$

The previous results can easily be generalized to any transition going from a ground state g with angular momentum J_g to an excited state e with angular momentum J_e, for an atom with any number of electrons. One simply has to use the Wigner-Eckart theorem (Complement $\mathrm{D_X}$ and exercise 8 in Complement $\mathrm{G_X}$). The dipole operator $\hat{\boldsymbol{D}}$ (sum of the dipole operators for each individual electron) is a vector operator whose three spherical components $\hat{\boldsymbol{D}}_s$ with $s = +1, 0, -1$ are equal to :

$$\hat{D}_{+1} = \frac{-1}{\sqrt{2}}(\hat{D}_x + i\hat{D}_y) \tag{5a}$$

$$\hat{D}_0 = \hat{D}_z \tag{5b}$$

$$\hat{D}_{-1} = \frac{1}{\sqrt{2}}(\hat{D}_x - i\hat{D}_y) \tag{5c}$$

The Wigner-Eckart theorem (Complement $\mathrm{D_X}$) states that

$$\langle J_e, m_e|\hat{\boldsymbol{D}}_s|J_g, m_g\rangle = \langle e||\hat{D}||g\rangle \ \langle J_e, m_e|J_g, 1; m_g, s\rangle \tag{6}$$

where the last term on the right-hand side is a Clebsch-Gordan coefficient (Chapter X, § C-4-c) and the first term is a "reduced matrix element" independent of m_g, s and m_e. The Clebsch-Gordan coefficient is different from zero only if:

- a triangle can be formed with J_e, 1 and J_g, which means that J_e is equal either to J_g, or to $J_g \pm 1$, the transition $J_e = 0 \leftrightarrow J_g = 0$ being forbidden;
- $m_e = m_g + s$.

As the three values of s ($s = +1, 0, -1$) correspond to the three polarizations σ_+, π, σ_- respectively, the selection rules (4) for any given transition $J_g \leftrightarrow J_e$ can be generalized to (see Figure 1):

$$\sigma_+ \longrightarrow m_e = m_g + 1, \qquad \pi \longrightarrow m_e = m_g \qquad \sigma_- \longrightarrow m_e = m_g - 1 \tag{7}$$

[2]Because of the transversality of the field, the light beam must then propagate in a direction perpendicular to the Oz axis.

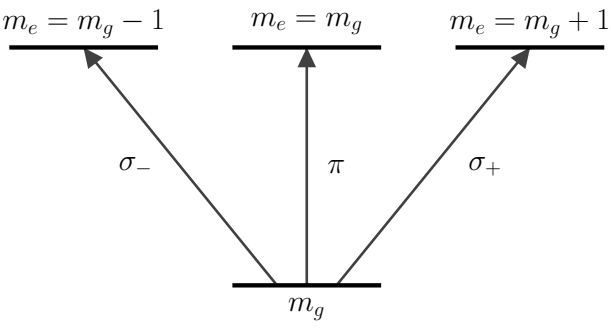


Figure 1: Selection rules for an electric or magnetic dipole transition. The magnetic quantum number m increases by one unit for an excitation with a σ_+ polarization, remains unchanged for a π polarization and decreases by one unit for a σ_- polarization.

1-c. Conservation of total angular momentum

We saw that, when an atom absorbs a photon having a polarization σ_+ with respect to a z axis, the component of the atom's angular momentum along that axis increases by one (in $\hbar$ units). The conservation of the total angular momentum means that the absorbed σ_+ photon must have an angular momentum $+\hbar$ along the z axis.

This result can also be obtained from the study of the radiation angular momentum presented in Complement B_{XIX}, as we now show. Although the interaction Hamiltonian pertaining to the atomic internal variables does not depend on the photon wave vector[3] $\boldsymbol{k}$, one can always choose a $\boldsymbol{k}$ wave vector parallel to the atomic quantization axis (since these two directions must be perpendicular to the polarization vector of the circular wave). Now we saw in Complement B_{XIX} that in a plane wave (or in a beam of plane waves with a small angular aperture) a photon with polarization σ_+ with respect to the wave vector has a total angular momentum (actually reduced to its spin angular momentum) equal to $+\hbar$ and parallel to its wave vector. The total angular momentum is indeed conserved.

2. Optical methods

The polarization selection rules show that it is possible to selectively excite a Zeeman sublevel of an atomic excited state. In a similar way, we shall see below that the observation of the emitted light in a given direction with a specific polarization allows determining from which excited sublevel the light was emitted. Such a selectivity in excitation and detection is the base of the optical methods for Hertzian spectroscopy, as will be illustrated below with two examples.

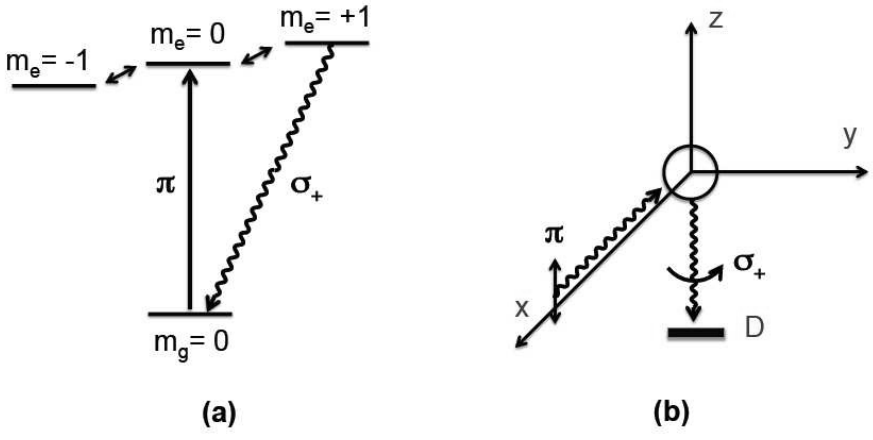


Figure 2: Energy diagram and Zeeman structure of the $6^1S_0 \longrightarrow 6^3P_1$ transition of the even isotopes of mercury at 253.7 nm (Fig.a). A static magnetic field, applied along the z axis, lifts the degeneracy of the excited state, which is split into three equidistant Zeeman sublevels: $m_e = -1, 0, +1$. A resonant light beam, propagating along the x axis, with a linear polarization π parallel to the z axis (Fig.b), selectively excites the atoms into the $m_e = 0$ sublevel (upward arrow on Fig.a). When they are in the excited state, the atoms are transferred from the $m_e = 0$ state to both states $m_e = \pm 1$ by a resonant radiofrequency field (oblique small double arrows on Fig.a). A detector D placed along the z axis, just behind an analyser that only transmits light with σ_+ polarization along the z axis, detects the light emitted by the atoms (Fig.b). Consequently, the detector is only sensitive to the light emitted from the $m_e = +1$ sublevel (wiggly arrow on Fig.a); it yields a signal proportional to the population of that sublevel.

2-a. Double resonance method

We now explain the principle of the method taking as an example the even isotopes of mercury, for which the first theoretical predictions were made by Brossel and Kastler [35]; the first experimental evidences were obtained by Brossel and Bitter [36]. Since for these isotopes the nuclear spin is zero, and as in a mercury atom all the electron shells are filled in the ground state, the energy diagram is particularly simple. The ground state g has a zero angular momentum ($J_g = 0$), and the first accessible excited states, an angular momentum $J_e = 1$. In the presence of a static magnetic field B applied along the z axis, the three Zeeman sublevels $m_e = -1, 0, +1$ of the excited state undergo Zeeman shifts proportional to m_e and to the applied field, so that the energy of the m_e state is written:

$$E_{m_e} = E_0 + g_L m_e \mu_B B \tag{8}$$

[3] In the long wavelength approximation, the radiation wave vector $\boldsymbol{k}$ no longer appears in the interaction Hamiltonian for the atomic internal variables. This wave vector only appears in the part of the interaction Hamiltonian, $\exp(i\boldsymbol{k} \cdot \hat{\boldsymbol{R}})$, pertaining to the external variables.

where E_0 is the excited state energy in the absence of the field, g_L is the Landé g-factor of that state, and μ_B the Bohr magneton. The ground state $J_g = 0, m_g = 0$, whose energy is chosen to be zero, is not affected by the B field (Figure 2a).

In the double resonance method, the atoms are selectively excited by a resonant light beam with polarization π into the sublevel $m_e = 0$. For example, the exciting beam propagates along the x axis, with a polarization $\boldsymbol{\varepsilon}_z$ perpendicular to the x axis, and parallel to the z axis, hence with a π polarization (Figure 2b).

If the atoms were left alone, with no perturbation while in the excited state during its radiative lifetime $\tau_R = 1/\Gamma$ (of the order of 1.5 10^{-7} sec), they would remain in that sublevel for the entire time they stay in the excited state. On the other hand, if they are subjected to a resonant radiofrequency field[4] that is strong enough to make them go from $m_e = 0$ to $m_e = \pm 1$ during the excited state lifetime τ_R, the two states $m_e = \pm 1$ will be equally populated.

Is it possible, observing the light emitted by the atoms as they go back to the ground state by spontaneous emission of a photon, to determine the sublevel m_e from which the light was emitted, and hence obtain a signal proportional to this sublevel population? This problem must be analyzed with more care than the absorption process for the following reason. Once the atom has reached the excited sublevel m_e, it can emit in any direction with all possible polarizations, which are not necessarily the three basic polarizations σ_+, π or σ_-. We are going to show that one can place the detector in a specific direction, far from the atom, and put in front of it a polarization analyzer suitably chosen so as to be able to determine from which sublevel the detected photon was emitted.

To demonstrate this result, it is important to first study the oscillations of the atomic dipole associated with a transition $m_g \longleftrightarrow m_e$. Let us assume the detector is placed on the z axis where the atom is located.

(i) For the transition $m_e = 0 \longleftrightarrow m_g = 0$, as the dipole oscillates along the z axis, it does not emit along that axis. The detector does not receive any fraction of the light emitted by an atom in the $m_e = 0$ state.

(ii) On the other hand, the dipole associated with the transition $m_g = 0 \longleftrightarrow m_e = +1$, which rotates around the z axis following the right-hand rule, in a plane perpendicular to that axis, emits along that axis light with a σ_+ polarization; this light yields a signal on the detector equipped with an analyzer selecting right circular polarization. This analyzer will, however, block the light emitted by an atom in the $m_e = -1$ sublevel, which has a left circular polarization.

To sum up, placing a detector in a well chosen direction, preceded by an analyzer selecting a suitable polarization, one can block out all light except the one coming from a specific unique m_e sublevel, and get a signal proportional to that sublevel population.

The principle of this double resonance experiment is to selectively excite the atoms in the $m_e = 0$ sublevel, and to detect, by observing the σ_+ light emitted along the z axis, the variations of the number of atoms transferred into the $m_e = +1$ sublevel by a radiofrequency field with angular frequency ω_{RF} close to the Zeeman frequency $\omega_Z = g_L \mu_B B/\hbar$. Observing the variation of the emitted light as one scans ω_{RF} for a fixed value of the magnetic field, or as one scans the magnetic field for a fixed ω_{RF}, one

[4] The atoms are thus submitted to two resonant excitations: an optical resonant excitation that brings them from $m_g = 0$ to $m_e = 0$; a radiofrequency resonant excitation that brings them from $m_e = 0$ to $m_e = \pm 1$. This is why this method is called "double resonance method".

can optically detect the magnetic resonance in the excited state.

Comment

What happens if the light emitted by the atom is not observed along the z axis, with a right circular analyzer, but with a detector placed along the y axis, behind an analyzer selecting a linear polarization parallel to the x axis, and hence perpendicular to the z axis, a polarization called "sigma"[5]? The light emitted along the y axis by the atom in the $m_e = 0$ state has a linear polarization parallel to the oscillating dipole, and hence parallel to the z axis. It is blocked by the analyzer that only lets through the orthogonal polarization. On the other hand, whether the atom is in the $m_e = +1$ or $m_e = -1$ sublevel, the rotating dipole in the xOy plane (following the right-hand or left-hand rule) emits in that plane light with a linear polarization perpendicular to the z axis; that light can be detected by the sigma polarization detector and it yields a signal proportional to the sum of the populations of both sublevels $m_e = +1$ and $m_e = -1$. Actually, the resonant radiofrequency field excites a linear superposition of these two Zeeman sublevels. It was later discovered that the waves emitted from these two $m_e = +1$ and $m_e = -1$ states gave rise to interference, hence modulating at the frequency $2\omega_{RF}$ the detected sigma light intensity[6]. The detector signal therefore contains a component modulated at the frequency $2\omega_{RF}$, on top of a continuous component (in steady state), proportional to the sum of the two populations of the Zeeman sublevels. This continuous component was the signal used in the first double resonance experiment.

The shape of the magnetic resonance line can be exactly computed; it leads to analytical expressions in excellent agreement with experimental observations. The center of the resonance line yields the Landé g-factor of the excited state, i.e. the magnetic moment of that state. The resonance width, extrapolated to zero radiofrequency intensity to eliminate radiative broadening, yields the natural Γ width of the excited state.

Calculation of the line shape

Broadband excitation with a π polarization prepares, in a quasi-instantaneous way, the atom in the $m_e = 0$ excited state. In the steady state, n_0 atoms per unit time are excited into that state. Each atom then evolves, because of its interaction with the radiofrequency field $\boldsymbol{B}_1$, and its state becomes a linear superposition of the three sublevels $m_e = -1, 0, +1$. Let us assume the radiofrequency field is a rotating field, which allows introducing the associated rotating reference frame (Complement $\mathrm{F_{IV}}$) where the atom's evolution actually becomes a simple rotation around $\boldsymbol{B}_1$. Using the rotation matrix for a spin 1, one can find the expression for $P(m_e = 0 \longrightarrow m_e = +1, t)$, the probability for the atom, initially in the $m_e = 0$ state, to be found after a time t in the $m_e = +1$ state [38]. Because of the radiative lifetime $\tau_R = 1/\Gamma$ of the excited sublevels, this probability is reduced by a factor $e^{-\Gamma t}$. Consequently, in steady state, the number of atoms transferred to the $m_e = +1$ state is equal to[7]:

$$N_{+1} = n_0 \int_0^\infty P(m_e = 0 \longrightarrow m_e = +1, t) e^{-\Gamma t} \mathrm{d}t \tag{9}$$

Using the expression for $P(m_e = 0 \longrightarrow m_e = +1, t)$, one finally obtains:

$$N_{+1} = \frac{n_0 \Omega^2}{2\Gamma} \frac{4\delta^2 + \Gamma^2 + \Omega^2}{(\delta^2 + \Gamma^2 + \Omega^2)(4\delta^2 + \Gamma^2 + 4\Omega^2)} \tag{10}$$

[5]This set-up was the one used in the first double resonance experiment [36].
[6]Such modulations, called "light beats", were observed in 1959 [37].
[7]A similar computation can be carried out for N_{-1}.

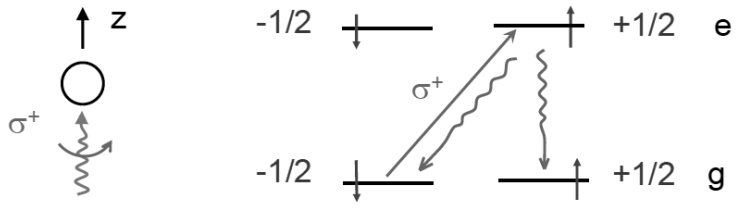


Figure 3: Principle of optical pumping for a $J_g = 1/2 \rightarrow J_e = 1/2$ transition. The resonant absorption of a photon with σ^+ polarization selectively excites the $m_g = -1/2 \longrightarrow m_e = 1/2$ transition. Once it has reached its excited state, the atom falls back, through spontaneous emission, into the $m_g = \pm 1/2$ states. If it falls in the $m_g = +1/2$ state, it can no longer absorb an incident photon, and remains in that state (since there is no transition σ^+ originating from the $m_g = +1/2$ state where the atoms accumulate). In addition, any change in the population difference between the $m_g = \pm 1/2$ sublevels can be detected by a change in the incident beam absorption, since this absorption is only possible starting from the $m_g = -1/2$ sublevel.

In this expression, Ω is the Rabi frequency associated with the radiofrequency field $\boldsymbol{B}_1$, proportional to that field's amplitude; $\delta = \omega_{RF} - \omega_Z$ is the difference between the frequency ω_{RF} of the RF field and the Zeeman frequency ω_Z associated with the gap between the Zeeman sublevels, which is proportional to the static field B_0. The resonance is plotted by scanning either the frequency of the RF field, or the static field, which amounts to scanning ω_Z.

2-b. Optical pumping

Optical pumping, proposed by A. Kastler [39], extends to atomic ground states the essential ingredients of the double resonance method. It also opens the possibility of achieving, in a steady state, large population differences between Zeeman sublevels in the ground state.

We shall explain this method's principle in the simple case where the ground state g (as well as the excited state e) has only two Zeeman sublevels $m_g = \pm 1/2$ (or $m_e = \pm 1/2$ for the excited state). The ideas introduced for that example remain valid for more complex transitions where the e and g states have a higher degeneracy.

The principle of optical pumping is illustrated in Figure 3. Atoms are excited by a resonant beam with σ^+ polarization, propagating along the z axis (left-hand side of the figure). A static magnetic field is also applied along that same axis. The absorption of a resonant photon with σ^+ polarization is selective, meaning it can only excite the $m_g = -1/2 \longrightarrow m_e = 1/2$ transition, during which the magnetic quantum number m increases by one unit. Once in the excited state, and after an average time $\tau_R = 1/\Gamma$, the atom falls back, through spontaneous emission, to the $m_g = \pm 1/2$ states. The probabilities of the various possible transitions are proportional to the square of the Clebsch-Gordan coefficients. If the atom falls into the $m_g = -1/2$ state, it can reabsorb a σ_+ photon. After a certain number of cycles, it will eventually fall in the $m_g = +1/2$ state, where it can no longer absorb an incident photon. It will remain in that state (since no σ^+ transition originates from the $m_g = +1/2$ state) and the atoms will accumulate in

that sublevel. The cycles of absorption of a σ_+ photon starting from the $m_g = -1/2$ state followed by a spontaneous emission a photon transferring the atom in the $m_g = +1/2$ level can be considered as an "optical pump" that empties the $m_g = -1/2$ sublevel to fill the $m_g = +1/2$ sublevel, hence the name "optical pumping".

Balance of angular momentum exchanges

During an optical pumping cycle, the atom gains an angular momentum $+\hbar$ since it goes from the state $m_g = -1/2$ to the state $m_g = +1/2$. The radiation looses one unit $\hbar$ of angular momentum as a σ_+ photon is absorbed. Since the total angular momentum of the system atom + radiation is conserved, the angular momentum of the radiation emitted by the atom going by spontaneous emission from the state $m_e = +1/2$ to the state $m_g = +1/2$ must be zero.

In order to directly demonstrate that the radiation does not lose, on average, any angular momentum during this transition, one must take into account that, as it goes from the $m_e = +1/2$ state to the $m_g = +1/2$ state, the atom actually emits a spherical wave in all directions, with all possible transverse polarizations; it is therefore not correct to say that the atom emits a single photon with a π polarization, which would imply the atom only emits photons with a wave vector perpendicular to the z axis. The correct way is to calculate the total (orbital and spin) angular momentum of the spherical wave emitted by the atom going from $m_e = +1/2$ to $m_g = +1/2$. We give a brief outline of the computation[8] as this involves expanding the field no longer onto plane waves but onto multipolar waves (Complement B_{XIX}, § 3-c-α). Those waves are field modes characterized by four quantum numbers: wave number k (or energy), parity equal to $+1$ or -1, total angular momentum J (which must be an integer), and finally the M_J component of the angular momentum J on the quantization axis z (which varies by unit steps between $-J$ and $+J$). For an electric dipole transition such as the one studied here, the parity equals -1 and the total angular momentum J equals 1. It can be shown that M_J is equal to 0 when the atom goes from m_e to $m_g = m_e$, and is equal to ± 1 when the atom goes from m_e to $m_g = m_e \mp 1$.

The correct language to describe the spontaneous emission into the ground state is as follows: for a $J_e = 1/2 \longleftrightarrow J_g = 1/2$ transition, an atom in the $m_e = +1/2$ state can fall back either into the state $m_g = -1/2$ by spontaneous emission of an electric dipole photon $M_J = +1$, or into the state $m_g = +1/2$ by spontaneous emission of an electric dipole photon $M_J = +0$. This correlation between m_g and M_J is due to conservation of total angular momentum, arising from the rotational invariance of the interaction Hamiltonian. The final state of the system after the spontaneous emission is thus an entangled state, a superposition of the $|m_g = -1/2\rangle \otimes |M_J = +1\rangle$ state and the $|m_g = +1/2\rangle \otimes |M_J = 0\rangle$ state, with coefficients weighted by the Clebsch-Gordan coefficients of the two atomic transitions, coming from the dipole matrix elements. This computation also yields the speed with which the two ground state sublevels repopulate, the branching ratio being given by the square of the Clebsch-Gordan coefficients.

Light plays a double role in these experiments. As we just saw, it polarizes the atoms by accumulating them in a ground state sublevel; it also permits the optical detection of the atoms' polarization. As the atoms can only absorb the incident σ_+ light if they are in the $m_g = -1/2$ sublevel, the absorption of that light yields a signal proportional to the population of that sublevel. Any change in the population differences

[8] The interested reader will find more details on multipolar waves properties in Complement B_I of [16] and in [34].

between the $m_g = -1/2$ and $m_g = +1/2$ sublevels, induced by a resonant radiofrequency field or by collisions, can therefore be detected by a change in the absorbed light.

2-c. Original features of these methods

We now review some of the original features of these optical methods, to understand their prominent role in the development of atomic physics.

- At the time it was suggested, the double resonance was among the first methods to extend the magnetic resonance techniques to atomic excited states. These techniques had been developed for ground states or metastables states with very long lifetimes, using essentially atomic or molecular beams: Stern-Gerlach type experimental set-ups were used to select atoms in given internal states; the flipping of the spins by a RF field would change the trajectories of the atoms or molecules in a detectable way, hence allowing, in most cases, the monitoring of the magnetic resonance (see for example reference [40]). These techniques could not be extended to the excited states because of their very short lifetimes.

- An interesting feature of these optical methods is their selectivity, both for the excitation and the detection. This selectivity comes from the light polarization, and not from the light frequency. The width of the spectral sources used at the time, and the Doppler width of the spectral lines of the atoms contained in a glass cell were considerably larger than the frequency differences between optical lines going from the ground state Zeeman sublevels to the excited states Zeeman sublevels; it was thus out of the question to try to excite or detect a single Zeeman component of the optical line.

- The measurements of the Zeeman or hyperfine structures in the excited states by the double resonance method are high resolution measurements. The structures under study are not determined by the measurement of the difference between two optical line frequencies, but by a direct measurement of the structure. In the radiofrequency or microwave domain, the Doppler width is negligible and the measurement resolution is only limited by the natural width.

- The optical methods are highly sensitive methods. A radiofrequency transition between two sublevels of the excited or ground state is not detected by the loss or gain in energy of the radiofrequency field, but via an absorbed or reemitted optical photon, which has a much higher energy than a RF photon, and whose polarization depends on which sublevel the atom is in. It is therefore possible to detect magnetic resonances in a very dilute medium, such as a vapor.

- Very high polarization ratios in the ground state, up to 90%, can be achieved by optical pumping. Such ratios are considerably larger than those expected at thermodynamic equilibrium: because of the very weak Zeeman shifts between the ground state sublevels, and the high temperature at which the experiments are conducted, the Boltzmann factors $\exp(-\hbar\omega_Z/k_BT)$ are all very close to 1. Note in addition that optical pumping may easily result, by a suitable choice of the polarization, in a larger population for the ground state Zeeman sublevel having the higher energy. This is one of the first examples of a method for achieving a population inversion, an essential condition for obtaining a maser or a laser effect.

More details on the optical methods and their applications can be found in the well documented work [24] and the references suggested therein.

3. Transferring orbital angular momentum to external atomic variables

3-a. Laguerre-Gaussian beams

In optics or atomic physics experiments, one often uses Gaussian beams, linear superpositions of plane waves with wave vectors nearly parallel to an optical axis (paraxial approximation). If all the plane waves forming the Gaussian beam have the same polarization $\boldsymbol{\varepsilon}$, the field phase in planes perpendicular to the beam axis does not depend on the azimuth angle φ that determines the direction around the beam axis. New types of Gaussian beams have recently been realized[9], called *Laguerre-Gaussian* (LG) beams, for which the field has an azimuthal dependence in $\exp im\varphi$, where $m = \pm 1, \pm 2, ...$, in planes perpendicular to the beam axis. We already mentioned the existence of such beams in § 3-c-β of Complement $\mathrm{B_{XIX}}$. We now show how the absorption of photons from such beams can transfer to the atomic center of mass a non-zero orbital angular momentum with respect to the beam axis.

3-b. Field expansion on Laguerre-Gaussian modes

The LG modes form a possible basis for expanding any field. They are characterized by three quantum numbers: the wave number k, the number of nodes p in the radial direction, and the integer number m characterizing the phase dependence on the azimuth angle φ. We assume the polarization $\boldsymbol{\varepsilon}$ to be uniform in the beam. We now place an atom in that beam, and use the LG modes basis $\boldsymbol{\varepsilon} F^m_{k,p}(\boldsymbol{r})$. Instead of being written $\langle \psi^{\mathrm{ext}}_{\mathrm{fin}} | \exp(i\,\boldsymbol{k}_i \cdot \hat{\boldsymbol{R}})\ |\psi^{\mathrm{ext}}_{\mathrm{in}}\rangle$, the matrix element pertaining to the external variables of the interaction Hamiltonian in this basis is now written :

$$\left\langle \psi^{\mathrm{ext}}_{\mathrm{fin}} \right|\, F^m_{k,p}(\hat{\boldsymbol{R}})\ \left|\psi^{\mathrm{ext}}_{\mathrm{in}}\right\rangle \tag{11}$$

This relation shows that the initial wave function $\psi^{\mathrm{ext}}_{\mathrm{in}}(\boldsymbol{R})$ of the atomic center of mass is now multiplied by the function $F^m_{k,p}(\boldsymbol{R})$ characterizing the mode. The phase factor $\exp(im\varphi)$ is, in a manner of speaking, "imprinted" on the initial wave function. Equation (11) means that the transition amplitude, concerning the external variables and induced by the interaction Hamiltonian, is equal to the scalar product of $F^m_{k,p}(\boldsymbol{R})\psi^{\mathrm{ext}}_{\mathrm{in}}(\boldsymbol{R})$ and the final center of mass wave function $\psi^{\mathrm{ext}}_{\mathrm{fin}}(\boldsymbol{R})$.

Imagine that the initial external state of the atom has a zero angular momentum with respect to the z axis, i.e. that $\psi^{\mathrm{ext}}_{\mathrm{in}}(\boldsymbol{R})$ does not depend on the angle φ. The absorption of a photon from such an LG beam, with quantum numbers k, p, m, gives to the product $F^m_{k,p}(\boldsymbol{R})\psi^{\mathrm{ext}}_{\mathrm{in}}(\boldsymbol{R})$ a φ dependence given by $\exp(im\varphi)$. This means that in its final state, the center of mass must have an orbital angular momentum $m\hbar$ with respect to the z axis, since $L_z = (\hbar/i)\partial/\partial\varphi$. The LG beam has transferred to the atom's center of mass an orbital angular momentum $m\hbar$. It is important to note that the transfer's efficiency, described by the matrix element (11), will only be significant if the

[9] The main method used to achieve such beams is to numerically design and fabricate holograms, then use them to diffract a Gaussian beam. For a review of the properties and applications of these new types of beams, see reference [41].

spatial extent of the initial and final wave functions are well adapted to the geometrical characteristics of the LG beam. In the vicinity of the focal point, the width of the beam is of the order of its "waist" w_0, i.e. of the order of a few microns, much larger than the atomic wave packets, of the order of nanometers at normal temperatures ($T \simeq 300K$). This explains why the orbital angular momentum transfer to the atoms' centers of mass became operational only when atoms could be cooled down to very low temperatures, in the microkelvin, or even nanokelvin range. The matter waves thus obtained, in Bose-Einstein condensates for example, can have spatial extensions of the order of a few microns. The transfer of orbital angular momentum by "phase imprint" can then be used to generate quantum vortices (see Complement D_{XV}, § 3-b-β) in matter waves, where the atoms rotate in phase around an axis. This method was actually used to create such vortices in a Bose-Einstein condensate of trapped atoms [42].

Chapter XX

Absorption, emission and scattering of photons by atoms

A A basic tool: the evolution operator 2068
A-1 General properties 2069
A-2 Interaction picture 2070
A-3 Positive and negative frequency components of the field . . . 2072
B Photon absorption between two discrete atomic levels . . . 2073
B-1 Monochromatic radiation 2073
B-2 Non-monochromatic radiation 2075
C Stimulated and spontaneous emissions 2080
C-1 Emission rate 2080
C-2 Stimulated emission 2081
C-3 Spontaneous emission 2081
C-4 Einstein coefficients and Planck's law 2083
D Role of correlation functions in one-photon processes . . . 2084
D-1 Absorption process 2084
D-2 Emission process 2085
E Photon scattering by an atom 2085
E-1 Elastic scattering 2086
E-2 Resonant scattering 2089
E-3 Inelastic scattering - Raman scattering 2091

Introduction

In this chapter, we will use the results established in the previous chapter to study some elementary processes concerning the absorption or emission of photons by atoms. Knowing the Hamiltonians describing the atomic energy levels and the radiation, as well

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

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

as their interactions, we can now focus on solving Schrödinger's equation that governs the evolution of the system atom plus field. Our objective is to compute the probability amplitude for that system to go from a given initial state $|\varphi_i\rangle$ at time t_i to a certain final state $|\varphi_f\rangle$ at a later time t_f.

In quantum mechanics, the evolution of the system's state vector between the instants t_i and t_f is controlled by the evolution operator $U(t_f, t_i)$, which is the basic tool for computing the amplitudes of the various processes studied in this chapter. This is why we start in § A by reviewing a number of equations satisfied by $U(t_f, t_i)$, which will be useful for the forthcoming computations.

The first processes we shall study in § B concern photon absorption or emission by an atom undergoing a transition between two discrete states. We shall first consider monochromatic incident radiation, and then a broadband excitation. We will then show in § C that two types of emission can occur during these interaction processes: stimulated emission, which is also predicted in a semiclassical treatment, and spontaneous emission, which requires a quantum treatment of the radiation. We will make a connection with the method used by Einstein to reestablish Planck's law (giving the spectral distribution of the black body radiation), and deduce the absorption and emission coefficients. The role of correlation functions (pertaining to the atomic dipole and to the incident field) in the computation of transition probabilities is discussed in § D.

An important example that involves not one, but two photons, is the scattering of a photon by an atom: during that process, an incident photon is absorbed and a new one is created either by spontaneous or induced emission. This process is studied in § E. When the frequency of the incident photon is close to the atomic transition frequency, the scattering is said to be "quasi-resonant". Its description requires a non-perturbative treatment that will be developed, based on the results of §§ 4 and 5 of Complement $\mathrm{D_{XIII}}$. In this entire chapter, we shall only consider cases where the atomic levels are discrete; a case where those levels include a continuum will be treated in Complement $\mathrm{B_{XX}}$.

> ***Notation:*** In Chapter XVIII, it was important to distinguish between the classical quantities and the corresponding quantum operators, so that the latter were denoted with "hats". In the present chapter, this distinction becomes less important, and we will come back to a more standard and simpler notation, without the hats; for instance, the annihilation and creation operators will be denoted a_i and $a_i^\dagger$, instead of $\hat{a}_i$ and $\hat{a}_i^\dagger$.

A. A basic tool: the evolution operator

The unitary evolution operator $U(t, t_0)$ has been defined in Complement $\mathrm{F_{III}}$; it yields the state of a quantum system at instant t knowing the state of that system at a previous time t_0:

$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle \tag{A-1}$$

It is a unitary operator:

$$U^\dagger(t, t_0)U(t, t_0) = \mathbb{1} \tag{A-2}$$

If $H(t)$ is the system Hamiltonian, $U(t,t_0)$ obeys the differential equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{dt}}U(t,t_0) = H(t)U(t,t_0) \tag{A-3}$$

with the initial condition $U(t_0,t_0) = \mathbb{1}$. The integral equation:

$$U(t,t_0) = \mathbb{1} - \frac{i}{\hbar}\int_{t_0}^{t} dt' H(t')U(t',t_0) \tag{A-4}$$

is equivalent to the differential equation together with its initial condition. If the Hamiltonian is time-independent, the evolution operator is simply:

$$U(t,t_0) = e^{-iH(t-t_0)/\hbar} \tag{A-5}$$

A-1. General properties

In this entire chapter, we use the evolution operator to express the probability amplitude $A_{fi}(t_f,t_i)$ for the system, starting from the initial state $|\varphi_i\rangle$ at instant t_i, to be found in the state $|\varphi_f\rangle$ at time t_f:

$$A_{fi}(t_f,t_i) = \langle\varphi_f|U(t_f,t_i)|\varphi_i\rangle \tag{A-6}$$

Consider the total Hamiltonian H:

$$H = H_A + H_R + H_I = H_0 + H_I \tag{A-7}$$

where $H_0 = H_A + H_R$ is the non-perturbed Hamiltonian (sum of the isolated atom Hamiltonian H_A and the free radiation Hamiltonian H_R) and H_I is the interaction Hamiltonian (between the atom and the field). The evolution operators U_0 and U associated respectively with H_0 and H read:

$$U_0(t,t_0) = e^{-iH_0(t-t_0)/\hbar} \tag{A-8a}$$

$$U(t,t_0) = e^{-iH(t-t_0)/\hbar} \tag{A-8b}$$

These operators are related through the integral relation:

$$U(t,t_0) = U_0(t,t_0) + \frac{1}{i\hbar}\int_{t_0}^{t} \mathrm{d}t'\; U_0(t,t')\, H_I\, U(t',t_0) \tag{A-9}$$

To demonstrate this relation, we take the derivative of each side (taking into account the derivative with respect to the integral's upper bound, which appears in addition to that of the function to be integrated):

$$i\hbar\frac{d}{dt}U(t,t_0) = H_0\; U_0(t,t_0) + U_0(t,t)H_I U(t,t_0) + \frac{1}{i\hbar}\int_{t_0}^{t} \mathrm{d}t'\; H_0\; U_0(t,t')\, H_I\, U(t',t_0) \tag{A-10}$$

that is, taking into account the relation $U_0(t,t) = \mathbb{1}$ and relation (A-9):

$$\begin{aligned} i\hbar\frac{d}{dt}U(t,t_0) &= H_I U(t,t_0) + H_0\left[U_0(t,t_0) + \frac{1}{i\hbar}\int_{t_0}^{t} \mathrm{d}t'\; U_0(t,t')\, H_I\, U(t',t_0)\right] \\ &= [H_I + H_0]\, U(t,t_0) = HU(t,t_0) \end{aligned} \tag{A-11}$$

The operator defined in (A-9) therefore obeys relation (A-3) with the total Hamiltonian H; since, in addition, $U(t_0, t_0) = \mathbb{1}$, this operator is an evolution operator.

Inserting expression (A-9) for $U(t, t_0)$ into the integral on the right-hand side of that same expression (A-9), we obtain an expression of that operator containing a double integral in time. Reiterating this process several times, we get the series expansion of the evolution operator in powers of H_I:

$$U(t,t_0) = U_0(t,t_0) + \frac{1}{i\hbar}\int_{t_0}^{t} \mathrm{d}t'\; U_0(t,t')\, H_I\, U_0(t',t_0) + \\ + \left(\frac{1}{i\hbar}\right)^2 \int_{t_0}^{t} \mathrm{d}t' \int_{t_0}^{t'} \mathrm{d}t''\; U_0(t,t')\, H_I\, U_0(t',t'') H_I\, U_0(t'',t_0) + ... \qquad \text{(A-12)}$$

The n-order term of this series is a succession of $n+1$ non-perturbed evolutions, each described by U_0, separated by n interactions H_I.

The evolution operator also obeys another integral equation, symmetric to (A-9), which will be useful for what follows:

$$U(t,t_0) = U_0(t,t_0) + \frac{1}{i\hbar}\int_{t_0}^{t} \mathrm{d}t'\; U(t,t')\, H_I\, U_0(t',t_0) \qquad \text{(A-13)}$$

Its demonstration is similar to that of (A-9).

Finally, we can also insert expression (A-13) for U in the integral of (A-9), replacing t by t' and t' by t''. We then obtain:

$$U(t,t_0) = U_0(t,t_0) + \frac{1}{i\hbar}\int_{t_0}^{t} \mathrm{d}t'\; U_0(t,t')\, H_I\, U_0(t',t_0) + \\ + \left(\frac{1}{i\hbar}\right)^2 \int_{t_0}^{t} \mathrm{d}t' \int_{t_0}^{t'} \mathrm{d}t''\; U_0(t,t')\, H_I\, U(t',t'') H_I\, U_0(t'',t_0) \qquad \text{(A-14)}$$

Contrary to (A-12), the right-hand side of (A-14) only has three terms, and not an infinity. It is, however, the perturbed evolution operator U, and not U_0, that appears in the last term on the right-hand side of (A-14), in between the two interaction Hamiltonians H_I. This form of the evolution operator U will be used in § E-2.

A-2. Interaction picture

For the following computations, it will often be useful to write Schrödinger's equation in the interaction picture. Let $|\psi(t)\rangle$ be the Schrödinger state vector. Setting:

$$|\bar{\psi}(t)\rangle = U_0^{\dagger}(t,t_0)|\psi(t)\rangle = \exp\left[i\,(t-t_0)\,H_0/\hbar\right]|\psi(t)\rangle \qquad \text{(A-15)}$$

the new state vector $|\bar{\psi}(t)\rangle$ obeys the time evolution equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\bar{\psi}(t)\rangle = \bar{H}_I(t)|\bar{\psi}(t)\rangle \qquad \text{(A-16)}$$

where $\bar{H}_I(t)$ is defined as:

$$\bar{H}_I(t) = U_0^{\dagger}(t,t_0) H_I U_0(t,t_0) \qquad \text{(A-17)}$$

The evolution operator $\bar{U}(t,t_0)$ in the interaction representation yields the evolution of $|\bar{\psi}(t)\rangle$:

$$|\bar{\psi}(t)\rangle = \bar{U}(t,t_0)|\bar{\psi}(t_0)\rangle \tag{A-18}$$

From equations (A-1) and (A-15), we can deduce the relation between evolution operators in the two points of view:

$$\bar{U}(t,t_0) = U_0^\dagger(t,t_0)U(t,t_0) = \exp\left[i\,(t-t_0)\,H_0/\hbar\right]U(t,t_0) \tag{A-19}$$

In addition, insertion of (A-18) into (A-16) shows that:

$$i\hbar\frac{d}{dt}\bar{U}(t,t_0) = \bar{H}_I(t)\bar{U}(t,t_0) \tag{A-20}$$

which leads to the following series expansion for $\bar{U}(t,t_0)$ (perturbation expansion):

$$\bar{U}(t,t_0) = \mathbb{1} + \frac{1}{i\hbar}\int_{t_0}^{t} \mathrm{d}t'\ \bar{H}_I(t') + \left(\frac{1}{i\hbar}\right)^2 \int_{t_0}^{t}\mathrm{d}t' \int_{t_0}^{t'} \mathrm{d}t''\ \bar{H}_I(t')\bar{H}_I(t'') + ... \tag{A-21}$$

The great advantage of the interaction picture is that the state vector only evolves under the effect of the interaction – since $\bar{H}_I(t)$ is the only operator appearing on the right-hand side of (A-16). We shall see that this point of view also allows expressing the transition probabilities in terms of time correlation functions of dipole and field operators, i.e. as average values of products of physical quantities taken at two different instants, and evolving freely (under the effect of only H_0). Finally, when trying to calculate the transition probability between two eigenstates $|\varphi_i\rangle$ and $|\varphi_f\rangle$ of H_0, with respective energies E_i and E_f, it is often convenient to use the interaction picture since, according to (A-19), the transition amplitude is of the form:

$$\langle\varphi_f|\,U(t,t_0)\,|\varphi_i\rangle = \langle\varphi_f|\exp\left[-i\,(t-t_0)\,H_0/\hbar\right]\bar{U}(t,t_0)\,|\varphi_i\rangle \tag{A-22}$$

$$= e^{-iE_f(t-t_0)/\hbar}\,\langle\varphi_f|\,\bar{U}(t,t_0)\,|\varphi_i\rangle \tag{A-23}$$

As the phase factor $e^{-iE_f(t-t_0)/\hbar}$ disappears from the probability (modulus squared of the amplitude), it can be ignored; this allows simply replacing the evolution operator U by $\bar{U}$ and directly using the more compact expansion (A-21).

In this entire chapter and its complements, we describe the interaction between atom and field by the electric dipole Hamiltonian $H_I = -\boldsymbol{D}\cdot\boldsymbol{E}_\perp(\boldsymbol{R})$ introduced in § C-4 of Chapter XIX. For the sake of simplicity, we assume that $t_0 = 0$. In the interaction picture[1], this operator becomes:

$$\bar{H}_I(t) = -\bar{\boldsymbol{D}}(t)\cdot\bar{\boldsymbol{E}}_\perp(\boldsymbol{R},t) \tag{A-24}$$

where:

$$\bar{\boldsymbol{D}}(t) = \exp\left(iH_0t/\hbar\right)\boldsymbol{D}\exp\left(-iH_0t/\hbar\right) \tag{A-25a}$$

$$\bar{\boldsymbol{E}}_\perp(t) = \exp\left(iH_0t/\hbar\right)\boldsymbol{E}_\perp\exp\left(-iH_0t/\hbar\right) \tag{A-25b}$$

[1]Here, the atom's external degrees of freedom are treated classically. The atom is supposed to be at rest at point $\boldsymbol{R}$, meaning $\boldsymbol{R}$ is not modified when going to the interaction picture.

A-3. Positive and negative frequency components of the field

In the present chapter, we assume the system is contained in a cubic box of volume L^3. The transverse electric field is given by relation (B-3) of Chapter XIX. It is a linear combination of annihilation a_i and creation $a_i^\dagger$ operators, and can be expressed as the sum of two terms:

$$\boldsymbol{E}_\perp(\boldsymbol{R}) = \boldsymbol{E}_\perp^{(+)}(\boldsymbol{R}) + \boldsymbol{E}_\perp^{(-)}(\boldsymbol{R}) \tag{A-26}$$

where:

$$\begin{aligned}
\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R}) &= i\sum_i \left(\frac{\hbar\omega_i}{2\epsilon_0 L^3}\right)^{1/2} a_i \boldsymbol{\varepsilon}_i e^{i\boldsymbol{k}_i\cdot\boldsymbol{R}} \\
\boldsymbol{E}_\perp^{(-)}(\boldsymbol{R}) &= -i\sum_i \left(\frac{\hbar\omega_i}{2\epsilon_0 L^3}\right)^{1/2} a_i^\dagger \boldsymbol{\varepsilon}_i^* e^{-i\boldsymbol{k}_i\cdot\boldsymbol{R}} = \left[\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R})\right]^\dagger
\end{aligned} \tag{A-27}$$

Operator $\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R})$, obtained by keeping only the annihilation operators a_i in the expansion of $\bar{\boldsymbol{E}}_\perp(\boldsymbol{R})$, is called [2] the electric field "*positive frequency component*". As for the operator $\boldsymbol{E}_\perp^{(-)}(\boldsymbol{R})$, it is the "*negative frequency component*". These two operators are not Hermitian, and do not commute. In a product of field operators, the order is said to be *normal* if the creation operators are to the left of the annihilation operators, as in $\boldsymbol{E}_\perp^{(-)}\boldsymbol{E}_\perp^{(+)}$; the order is said to be *antinormal* for a product in the inverse order, as in $\boldsymbol{E}_\perp^{(+)}\boldsymbol{E}_\perp^{(-)}$.

In the interaction picture, the annihilation and creation operators become:

$$\begin{aligned}
\bar{a}_i(t) &= \exp\left(iH_0 t/\hbar\right) a_i \exp\left(-iH_0 t/\hbar\right) = a_i e^{-i\omega_i t} \\
\bar{a}_i^\dagger(t) &= \exp\left(iH_0 t/\hbar\right) a_i^\dagger \exp\left(-iH_0 t/\hbar\right) = a_i^\dagger e^{+i\omega_i t}
\end{aligned} \tag{A-28}$$

(these equalities can be verified by computing the matrix elements in the Fock state basis, eigenvectors of H_0, and using the fact that the only action of operator a_i is to annihilate a photon in the mode i). The positive and negative frequency components of the field are thus:

$$\begin{aligned}
\bar{\boldsymbol{E}}_\perp^{(+)}(\boldsymbol{R},t) &= i\sum_i \left(\frac{\hbar\omega_i}{2\epsilon_0 L^3}\right)^{1/2} a_i \boldsymbol{\varepsilon}_i e^{i(\boldsymbol{k}_i\cdot\boldsymbol{R}-\omega_i t)} \\
\bar{\boldsymbol{E}}_\perp^{(-)}(\boldsymbol{R},t) &= -i\sum_i \left(\frac{\hbar\omega_i}{2\epsilon_0 L^3}\right)^{1/2} a_i^\dagger \boldsymbol{\varepsilon}_i^* e^{-i(\boldsymbol{k}_i\cdot\boldsymbol{R}-\omega_i t)} = \left[\bar{\boldsymbol{E}}_\perp^{(+)}(\boldsymbol{R},t)\right]^\dagger
\end{aligned} \tag{A-29}$$

Suppose now that we wish to study the lowest order process of photon absorption by atoms. To compute the action of $\bar{H}_I(t)$ on the system's initial state, we can keep

[2] The positive frequency component annihilates a photon, the negative frequency component creates one. Furthermore, in the Heisenberg picture, we shall see that the free evolution of the positive component goes as $e^{-i\omega t}$, and that of the negative one, as $e^{+i\omega t}$. This labeling as "positive frequency" may seem somewhat counter intuitive, but is widely accepted.

only the terms in $\bar{H}_I(t)$ that annihilate one photon, i.e. the Hamiltonian terms containing $\bar{a}_i(t)$. This amounts to keeping only the positive frequency component of the field operator, and using the simplified interaction Hamiltonian:

$$\bar{H}_I(t) \simeq -\bar{\boldsymbol{D}}(t) \cdot \boldsymbol{E}_\perp^{(+)}(\boldsymbol{R}, t) \qquad \text{(absorption)} \tag{A-30}$$

In a similar way, to study the lowest order photon emission process, we can use the expression:

$$\bar{H}_I(t) \simeq -\bar{\boldsymbol{D}}(t) \cdot \boldsymbol{E}_\perp^{(-)}(\boldsymbol{R}, t) \qquad \text{(emission)} \tag{A-31}$$

B. Photon absorption between two discrete atomic levels

We start with monochromatic radiation (§ B-1), and will study later the broadband radiation case (§ B-2). The base of the computation is the study of the transition rate between stationary states of the non-perturbed Hamiltonian H_0; Complement $\mathrm{D_{XX}}$ will present a more detailed study in terms of wave packets propagating in free space, built from coherent superpositions of stationary states.

B-1. Monochromatic radiation

B-1-a. Probability amplitude (absorption)

We call a and b two discrete levels with respective energies E_a and E_b, and set:

$$E_b - E_a = \hbar\omega_0 \tag{B-1}$$

where $\omega_0/2\pi$ is the atomic eigenfrequency, assumed to be positive (the b level has an energy higher than the a level). For the sake of simplicity, we shall ignore the external variables, which amounts to considering the atom as infinitely heavy and at rest[3].

We assume the radiation is at the initial time $t_i = 0$ in a state $|\,\phi_{\text{in}}^R\rangle = |n_i\rangle$ containing n_i photons i with wave vector $\boldsymbol{k}_i$, polarization $\boldsymbol{\varepsilon}_i$ and frequency $\omega_i/2\pi$; it is a monochromatic radiation. The initial state of the system atom + radiation is written:

$$|\psi_{\text{in}}\rangle = |a; n_i\rangle \qquad \text{with energy } E_{\text{in}} = E_a + n_i\hbar\omega_i \tag{B-2}$$

We are trying to compute the probability amplitude for the atom to absorb a photon and be in the excited state b at instant $t_f = \Delta t$. The final state of the system must then be:

$$|\psi_{\text{fin}}\rangle = |b; n_i - 1\rangle \qquad \text{with energy } E_{\text{fin}} = E_b + (n_i - 1)\hbar\omega_i. \tag{B-3}$$

As mentioned above, when $|\psi_{\text{in}}\rangle$ and $|\psi_{\text{fin}}\rangle$ are eigenstates of H_0, it is easier to carry out the calculation in the interaction representation. In the expansion (A-21) of $\bar{U}$, the lowest order term that can link those two states is the first order term in H_I. Calling

[3] Complement $\mathrm{A_{XIX}}$ shows how taking into account the external variables and momentum conservation allows introducing the Doppler effect and the recoil effect in the absorption and emission of a photon. It also shows how the confinement of the atom in a region of space by a trapping potential allows controlling those effects.

Δt the duration of the interaction, the probability amplitude for the system, initially in the state $|\psi_{\rm in}\rangle$ at time $t_0 = 0$, to be at time $t = \Delta t$ in the final state $|\psi_{\rm fin}\rangle$ is:

$$\begin{aligned}\langle\psi_{\rm fin}|\bar{U}(t_f,t_i)|\psi_{\rm in}\rangle &= \frac{1}{i\hbar}\int_0^{\Delta t} {\rm d}t'\ \langle\psi_{\rm fin}|e^{iH_0t'/\hbar}H_I e^{-iH_0t'/\hbar}|\psi_{\rm in}\rangle \\ &= \frac{1}{i\hbar}\langle\psi_{\rm fin}|H_I|\psi_{\rm in}\rangle\int_0^{\Delta t}{\rm d}t'\ e^{i(E_{\rm fin}-E_{\rm in})t'/\hbar}\end{aligned} \tag{B-4}$$

or else, taking into account the relation $E_{\rm fin} - E_{\rm in} = \hbar(\omega_0 - \omega_i)$:

$$\begin{aligned}\langle\psi_{\rm fin}|\bar{U}(\Delta t,0)|\psi_{\rm in}\rangle &= \frac{1}{i\hbar}\langle\psi_{\rm fin}|H_I|\psi_{\rm in}\rangle\int_0^{\Delta t}{\rm d}t'\ e^{i(\omega_0-\omega_i)t'} \\ &= -\langle\psi_{\rm fin}|H_I|\psi_{\rm in}\rangle\frac{1}{\hbar\,(\omega_0-\omega_i)}\left[e^{i(\omega_0-\omega_i)\Delta t}-1\right]\end{aligned} \tag{B-5}$$

which leads to:

$$\langle\psi_{\rm fin}|\bar{U}(\Delta t,0)|\psi_{\rm in}\rangle = -\frac{i}{\hbar}\langle\psi_{\rm fin}|H_I|\psi_{\rm in}\rangle\ e^{i(\omega_0-\omega_i)\Delta t/2}\left[\frac{2\sin\left(\omega_0-\omega_i\right)\Delta t/2}{\left(\omega_0-\omega_i\right)}\right] \tag{B-6}$$

The absorption probability is the squared modulus of that expression:

$$\left|\langle\psi_{\rm fin}|\bar{U}(\Delta t,0)|\psi_{\rm in}\rangle\right|^2 = \frac{1}{\hbar^2}\left|\langle\psi_{\rm fin}|H_I|\psi_{\rm in}\rangle\right|^2\frac{4\sin^2\left(\omega_0-\omega_i\right)\Delta t/2}{\left(\omega_0-\omega_i\right)^2} \tag{B-7}$$

Taking relations (A-24) and (A-27) into account, the matrix element of H_I, appearing in the amplitude (B-5), can be written:

$$\langle\psi_{\rm fin}|H_I|\psi_{\rm in}\rangle = -i\sqrt{\frac{\hbar\omega_i}{2\epsilon_0 L^3}}\sqrt{n_i}\ \langle b|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|a\rangle e^{i\boldsymbol{k}_i\cdot\boldsymbol{R}} \tag{B-8}$$

so that the probability becomes:

$$\left|\langle\psi_{\rm fin}|\bar{U}(\Delta t,0)|\psi_{\rm in}\rangle\right|^2 = \frac{\omega_i}{2\hbar\epsilon_0 L^3}n_i\left|\langle b|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|a\rangle\right|^2\frac{4\sin^2\left(\omega_0-\omega_i\right)\Delta t/2}{\left(\omega_0-\omega_i\right)^2} \tag{B-9}$$

It is proportional to the number of incident photons n_i, i.e. to the incident intensity in the state $|\psi_{\rm in}\rangle$, as well as to the squared modulus of the atomic dipole matrix element between the states a and b (since $\boldsymbol{D}$ is an odd operator, the absorption of the photon can only occur between two states of different parity). This probability is an oscillating function of the time Δt.

B-1-b. Energy conservation

The presence in the denominator of (B-9) of the factor $(\omega_i - \omega_0)^2$ means that, the closer ω_i gets to ω_0, the larger the absorption probability will be. A photon absorption is said to be resonant when the absorbed photon energy is exactly equal to the energy the atom must gain to go from a to b (energy conservation). The width of the resonance described by (B-9) is of the order of $\Delta\omega_i = 1/\Delta t$ or, in energy, of the order of $\Delta E = \hbar/\Delta t$. This is consistent with the time-energy uncertainty principle: in a process extending over a length time Δt, the energy is only conserved to within $\hbar/\Delta t$.

Actually, when the variable is the angular frequency ω_i, one may consider the function within brackets in (B-6) to be an approximate delta function $\delta^{(\Delta t)}(\omega_0 - \omega_i)$, having a non-zero width of the order of $1/\Delta t$. As relation (10) of Appendix II in Volume II shows that the integral over ω_i of $\sin[(\omega_0 - \omega_i)\Delta t/2]/(\omega_0 - \omega_i)$ is equal to π, and since $(\omega_0 - \omega_i)$ is proportional to the difference in total energy between the final and initial states, equation (B-6) can be written (ignoring the phase factor):

$$\langle \psi_{\text{fin}} | \bar{U}(\Delta t, 0) | \psi_{\text{fin}} \rangle \simeq -2i\pi \ \langle \psi_{\text{fin}} | H_I | \psi_{\text{in}} \rangle \ \delta^{(\Delta t)}(E_{\text{fin}} - E_{\text{in}}) \tag{B-10}$$

B-1-c. Limits of the perturbative treatment

The lowest order perturbative treatment in H_I that we have used cannot remain valid for arbitrarily long times. To understand why, imagine for example that the resonance condition $\omega_i = \omega_0$ is satisfied. Since $\sin(x\Delta t)/x$ tends toward Δt when x goes to 0, the absorption probability predicted by (B-9) becomes proportional to Δt^2, which gets very large as the time interval Δt increases. Now a probability can never be larger than one; it is therefore obvious that expression (B-9) is no longer valid for long times. The same is true if ω_i gets close to ω_0 without being strictly equal to it: it is then quite possible for the oscillation amplitude of the right-hand side of (B-9) to be larger than one. The previous results can thus only be used for short enough times, ensuring the validity of the perturbative treatment.

We shall use in Complement C_{XX} the "dressed atom method" to get a more precise treatment of the coupling effects between a two-level atom and a single mode of the field. The coupling intensity will be characterized by a constant Ω_1 called the "Rabi frequency". At resonance, one shows that the probability amplitude presents a "Rabi oscillation"[4] in $\sin \Omega_1 \Delta t$. The quadratic behavior in Δt^2 found above for the absorption probability at resonance is simply the first term of the series expansion of $\sin^2(\Omega_1 \Delta t)$ in powers of $\Omega_1 \Delta t$. We shall also discuss the extent to which relation (B-9) can be used far from resonance.

B-2. Non-monochromatic radiation

We now study the absorption and emission processes when the radiation is no longer monochromatic. The transitions are still between two discrete atomic levels a and b; the case of a continuum of atomic levels is discussed in Complement B_{XX}.

B-2-a. Absorption of a broadband radiation

We now assume the initial state of the system atom + radiation to be of the form:

$$|\psi_{\text{in}}\rangle = |a; \phi^R_{\text{in}}\rangle \tag{B-11}$$

The atom is still in the internal state of lower energy a, but the radiation is now in a state $|\phi^R_{\text{in}}\rangle$ where photons occupy several modes with different frequencies. The radiation frequency distribution is characterized by a certain spectral band $\Delta\omega$. We shall see below (very last part of § B-2-a-γ) the condition $\Delta\omega$ must satisfy for the results obtained in

[4] We shall also show how this Rabi oscillation is modified when taking into account the width of the excited level due to spontaneous emission.

this section to be valid. The calculations will be carried out in the same way as in § B-1, while focusing more on the time correlation properties of the incident field.

We introduce the notation:

$$\langle b\,|\boldsymbol{D}|a\rangle = \boldsymbol{\mathcal{D}} \qquad \text{with:} \qquad \boldsymbol{\mathcal{D}} = \mathcal{D}\mathbf{e}_d \tag{B-12}$$

where $\mathbf{e}_d$ is the unit vector[5] (a priori complex) parallel to $\boldsymbol{\mathcal{D}}$, and $\mathcal{D}$ the (real) modulus of that vector. In the interaction picture (with respect to the free atom Hamiltonian), this equality becomes:

$$\langle b\,|\bar{\boldsymbol{D}}(t)|a\rangle = \langle b\,|e^{iH_0t}\boldsymbol{D}e^{-iH_0t}|a\rangle = \boldsymbol{\mathcal{D}}e^{i\omega_0 t} \tag{B-13}$$

α. Transition probability

Let us insert expression (A-24) for $\bar{H}_I(t)$ into (A-21). The first order term in $\bar{H}_I(t)$ yields the probability amplitude for the system, starting at $t_i = 0$ from the state $|\,a;\phi^R_{\text{in}}\rangle$, to be found at time $t_f = \Delta t$ in the state $|b;\phi^R_{\text{fin}}\rangle$. Taking (B-13) into account, we obtain:

$$\begin{aligned}\langle\psi_{\text{fin}}|\bar{U}(t_f,t_i)|\psi_{\text{in}}\rangle &= -\frac{1}{i\hbar}\int_0^{\Delta t} \mathrm{d}t'\ \langle b|\ \bar{\boldsymbol{D}}(t')\,|a\rangle\cdot\left\langle\phi^R_{\text{fin}}\right|\bar{\boldsymbol{E}}_\perp(\boldsymbol{R},t')\left|\phi^R_{\text{in}}\right\rangle \\ &= -\frac{\mathcal{D}}{i\hbar}\int_0^{\Delta t} \mathrm{d}t'\ e^{i\omega_0 t'}\left\langle\phi^R_{\text{fin}}\right| E^{(+)}_{\mathbf{d}}(\boldsymbol{R},t')\left|\phi^R_{\text{in}}\right\rangle\end{aligned} \tag{B-14}$$

where we have included only the positive frequency component of the field, which is the only one involved since $\left|\phi^R_{\text{fin}}\right\rangle$ contains less photons than $\left|\phi^R_{\text{in}}\right\rangle$ (we are dealing with an atomic absorption process); in this equality and the following, we shall use the convenient notation:

$$E^{(+)}_{\mathbf{d}}(\boldsymbol{R},t) = \mathbf{e}_d\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t) \quad ; \quad E^{(-)}_{\mathbf{d}^*}(\boldsymbol{R},t) = \mathbf{e}^*_d\cdot\bar{\boldsymbol{E}}^{(-)}_\perp(\boldsymbol{R},t) \tag{B-15}$$

where $\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t)$ and $\bar{\boldsymbol{E}}^{(-)}_\perp(\boldsymbol{R},t)$ are the field operators in the interaction picture defined in (A-29); the atom is supposed to be fixed at point $\boldsymbol{R}$. The corresponding transition probability $P^{\text{abs}}_{a\to b}(\Delta t)$ is obtained by squaring the modulus of amplitude (B-14), and then summing over all possible radiation final states. Replacing in (B-14) the integral variable t' by t'', we obtain:

$$P^{\text{abs}}_{a\to b}(\Delta t) = \frac{\mathcal{D}^2}{\hbar^2}\int_0^{\Delta t}\mathrm{d}t'\int_0^{\Delta t}\mathrm{d}t''\ e^{i\omega_0(t''-t')}G^N_R(t',t'') \tag{B-16}$$

where $G^N_R(t',t'')$ is defined as:

$$\begin{aligned}G^N_R(t',t'') &= \sum_{\left|\phi^R_{\text{fin}}\right\rangle}\langle\phi^R_{\text{in}}|E^{(-)}_{\mathbf{d}^*}(\boldsymbol{R},t')\left|\phi^R_{\text{fin}}\right\rangle\left\langle\phi^R_{\text{fin}}\right| E^{(+)}_{\mathbf{d}}(\boldsymbol{R},t'')|\phi^R_{\text{in}}\rangle \\ &= \langle\phi^R_{\text{in}}|E^{(-)}_{\mathbf{d}^*}(\boldsymbol{R},t')E^{(+)}_{\mathbf{d}}(\boldsymbol{R},t'')|\phi^R_{\text{in}}\rangle\end{aligned} \tag{B-17}$$

[5]The matrix elements $\mathcal{D}_{x,y,z}$ of the three components of $\boldsymbol{\mathcal{D}}$ are three a priori complex numbers. We set $\mathcal{D} = \sqrt{|\mathcal{D}_x|^2+|\mathcal{D}_y|^2+|\mathcal{D}_z|^2}$ and introduce the vector $\mathbf{e}_d = \boldsymbol{\mathcal{D}}/\mathcal{D}$. It is a unit vector since $\mathbf{e}^*_d\cdot\mathbf{e}_d = 1$; we have $\mathcal{D} = \boldsymbol{e}^*_d\cdot\boldsymbol{\mathcal{D}} = \langle b|\boldsymbol{e}^*_d\cdot\boldsymbol{D}|a\rangle$.

This function characterizes the role of the incident beam in the absorption process under study. It is the average value in the initial state of the product of two field operators, arranged in normal order (§ A-3), and taken at two different times.

Changing the integral variable t'' for $\tau = t'' - t'$ (the jacobian of this change of variables is equal to unity), equation (B-16) becomes:

$$P_{a\to b}^{\text{abs}}(\Delta t) = \frac{\mathcal{D}^2}{\hbar^2} \int_0^{\Delta t} \mathrm{d}t' \int_{-t'}^{\Delta t - t'} \mathrm{d}\tau \; e^{i\omega_0 \tau} G_R^N(t', t'+\tau) \tag{B-18}$$

The transition probability is therefore proportional to the time integral of the Fourier transform[6] with respect to τ of the function $G_R^N(t', t'+\tau)$, limited to the time interval $[-t', \Delta t - t']$.

β. Excitation spectrum

If the radiation initial state $\left|\phi_{\text{in}}^R\right\rangle$ is an eigenstate of the free radiation, the function $G_R^N(t', t'')$ depends only on the difference $t' - t''$. For example, if $\left|\phi_{\text{in}}^R\right\rangle$ is a Fock state[7]:

$$\left|\phi_{\text{in}}^R\right\rangle = |n_1, ..., n_i, ...\rangle \tag{B-19}$$

where the populated modes have the polarization $\boldsymbol{\varepsilon}_i$, inserting expansions (A-29) for the field operators leads to:

$$\begin{aligned} G_R^N(t'-t'') &= \sum_i \frac{\hbar\omega_i}{2\varepsilon_0 L^3} \underbrace{\left\langle \phi_{\text{in}}^R \left| a_i^\dagger a_i \right| \phi_{\text{in}}^R \right\rangle}_{=\langle n_i \rangle} \left| \mathbf{e}_d \cdot \boldsymbol{\varepsilon}_i \right|^2 \; e^{i\omega_i (t'-t'')} \\ &= \int \mathrm{d}\omega \; I_R^N(\omega) e^{i\omega(t'-t'')} \end{aligned} \tag{B-20}$$

where:

$$I_R^N(\omega) = \sum_i \frac{\langle n_i \rangle \hbar\omega_i}{2\varepsilon_0 L^3} \; \left| \mathbf{e}_d \cdot \boldsymbol{\varepsilon}_i \right|^2 \delta(\omega_i - \omega) \tag{B-21}$$

If the radiation is initially in the Fock state (B-19), the value of $\langle n_i \rangle$ is simply n_i, the number of photons in that state. On the other hand, if the radiation is in a statistical mixture of such states (see note 7), $\langle n_i \rangle$ represents the corresponding statistical average.

Expression (B-20) thus appears as the value at $t' - t''$ of the Fourier transform of a function $I_R^N(\omega)$ of ω, which depends on the initial photon populations $\langle n_i \rangle$. As $\langle n_i \rangle \hbar\omega_i$ is the average energy of mode i with frequency $\omega_i/2\pi$, the function $I_R^N(\omega)$ actually gives the variation of the energy density of the radiation as a function of frequency; this function is also referred to as the spectral distribution of the incident radiation (excitation spectrum). This distribution can have, a priori, any shape, but the energy density often presents a single peak of width $\Delta\omega$, with no other particular structure. Its Fourier transform $G_R^N(t' - t'')$ is then a function of width $\simeq 1/\Delta\omega$; as a result, when $|t' - t''|$

[6] The components of the field $E_{\mathbf{d}}^{(+)}$ vary in $\mathrm{e}^{-i\omega t}$. This is why the Fourier component of G_R^N at $-\omega_0$ appears in (B-18).

[7] Instead of a Fock state as the initial state, one could choose a statistical mixture of such states with arbitrary weights; this would not significantly change the following calculations and conclusions.

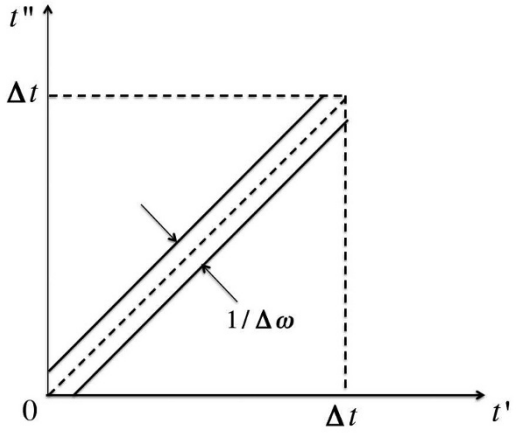


Figure 1: The function to be integrated in expression (B-16) yielding the absorption probability is significant only in a band along the first bisector, of width $1/\Delta\omega$, where $\Delta\omega$ is the width of the incident radiation spectrum. If $\Delta t \gg 1/\Delta\omega$, the area of the domain actually contributing to that integral increases linearly with the diagonal of the square, and hence with Δt (and not with its square).

becomes large compared to $1/\Delta\omega$, the correlation function $G_R^N(t'-t'')$ goes to zero and can be neglected. We shall see that in such a case the probability becomes proportional to Δt, which naturally leads to introducing a probability per unit time. Our reasoning will be similar to that of Complement E_{XIII} for a classical perturbation, but here the radiation is treated quantum mechanically.

γ. Probability per unit time

The integration domain of the double integral in (B-16) is plotted in Figure 1. As $G_R^N(t'-t'')$ goes to zero as soon as $|t'-t''|$ is large compared to $1/\Delta\omega$, the portion of that domain where the function to be integrated is not negligible is a band of width $1/\Delta\omega$, along the first bisector; the width of this band is very small compared to the domain extension if $\Delta t \gg 1/\Delta\omega$. To make use of that property, we again replace the integral variable t'' by $\tau = t'' - t'$ (the associated Jacobian for this change of variables in the double integral is equal to 1):

$$\int_0^{\Delta t} \mathrm{d}t' \int_0^{\Delta t} \mathrm{d}t'' = \int_0^{\Delta t} \mathrm{d}t' \int_{-t'}^{\Delta t - t'} \mathrm{d}\tau \tag{B-22}$$

In the second integral, the values of τ that actually yield a non negligible contribution are of the order of the correlation time $1/\Delta\omega$ of $G_R^N(t'-t'')$. If we assume $\Delta t \gg 1/\Delta\omega$, we can replace the limits of that second integral by $-\infty$ and $+\infty$. Inserting then (B-20) into (B-18) yields the integral:

$$\int_0^{\Delta t} \mathrm{d}t' \int_{-\infty}^{+\infty} \mathrm{d}\tau \int_{-\infty}^{+\infty} \mathrm{d}\omega \; e^{i(\omega_0-\omega)\tau} I_R^N(\omega) = 2\pi \, \Delta t \, I_R^N(\omega_0) \tag{B-23}$$

since the summation over $\mathrm{d}\tau$ yields the function $2\pi\delta(\omega_0-\omega)$, which allows integration over $\mathrm{d}\omega$; we get a function independent of t' whose integration is proportional to Δt.

We finally obtain:

$$P_{a\to b}^{\mathrm{abs}}(\Delta t)=\frac{2\pi}{\hbar^2}\ \Delta t\ \mathcal{D}^2\ I_R^N(\omega_0)=\frac{2\pi}{\hbar^2}\ \Delta t\ \mathcal{D}^2\sum_i\frac{\langle n_i\rangle\hbar\omega_i}{2\varepsilon_0 L^3}\ \left|\mathbf{e}_d\cdot\boldsymbol{\varepsilon}_i\right|^2\delta(\omega_i-\omega_0) \tag{B-24}$$

This means that $P_{a\to b}^{\mathrm{abs}}(\Delta t)$ increases **linearly** with Δt, which leads us to define an absorption probability per unit time (absorption rate):

$$W_{a\to b}^{\mathrm{abs}}=\frac{P_{a\to b}^{\mathrm{abs}}(\Delta t)}{\Delta t}=2\pi\frac{\mathcal{D}^2}{\hbar^2}I_R^N(\omega_0) \tag{B-25}$$

(remember that, for monochromatic radiation, we found an absorption probability increasing not linearly but as the square of Δt). This absorption rate is proportional to the radiation energy density at the atomic transition frequency ω_0. Formulas (B-21) and (B-25) give the dependence of the absorption rate on the various parameters of the incident radiation (populations of the modes, polarization $\boldsymbol{\varepsilon}_i$).

Our calculation is perturbative since it was carried out to the lowest order in H_I. It is therefore only valid for times Δt such that $P_{a\to b}^{\mathrm{abs}}(\Delta t)=W_{a\to b}^{\mathrm{abs}}\Delta t\ll 1$, i.e. such that $\Delta t\ll 1/W_{a\to b}^{\mathrm{abs}}$. In addition, we saw above that the linear variation of $P_{a\to b}^{\mathrm{abs}}(\Delta t)$ with Δt is obtained only if $\Delta t\gg 1/\Delta\omega$. These two inequalities are compatible if $\Delta\omega\gg W_{a\to b}^{\mathrm{abs}}$. The approximation to lowest order is thus valid only if the broad band radiation has a spectral width large compared to the absorption rate.

B-2-b. A specific case: isotropic radiation

The previous calculations can be pushed a step further when the radiation is isotropic, meaning when $\langle n_i\rangle$, the average number of photons in mode i, depends only on ω_i, and neither on the direction of the wave vector $\boldsymbol{k}_i$ nor on the polarization $\boldsymbol{\varepsilon}_i$. The results we shall obtain for this specific case will be useful for later computations (§ C-4) on the spontaneous emission rate, as well as for the isotropic radiation at thermal equilibrium.

Consider the limit of (B-25) when the volume L^3 containing the system goes to infinity. The summation over the index i can be replaced by an integral:

$$\sum_i\Rightarrow\frac{L^3}{(2\pi)^3}\int k^2\ \mathrm{d}k\ \mathrm{d}\Omega\sum_{\boldsymbol{\varepsilon}\perp\boldsymbol{k}} \tag{B-26}$$

where, for isotropic radiation, a sum is taken over two linear polarizations perpendicular to à $\boldsymbol{k}$. We assume that the vector $\mathbf{e}_d$ is real as well, and choose an Oz axis that is parallel to it[8]. We first calculate, for a given direction of $\boldsymbol{k}$, the sum of the quantities ε_{iz}^2 for the two polarizations. We take two polarization vectors, $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$, both perpendicular to $\boldsymbol{k}$, and perpendicular to each other. Imagine the first one is in the plane containing $\boldsymbol{k}$ and the z axis, so that $\varepsilon_{1z}=\sin\theta$, where θ is the angle between $\boldsymbol{k}$ and the z axis; the

[8] If $\mathbf{e}_d$ is complex, it is easy to see that the contributions of its real and imaginary part simply add in (B-24). The sum of these two contributions then gives (B-27) again.

second is necessarily perpendicular to that plane, which means $\varepsilon_{2z} = 0$. We then have $\sum_{\boldsymbol{\varepsilon}\perp\boldsymbol{k}} \varepsilon_{iz}^2 = \sin^2\theta$. The integral over all the directions of $\boldsymbol{k}$ is then trivial:

$$\int \mathrm{d}\Omega\ \sin^2\theta = 2\pi \int_0^\pi \mathrm{d}\theta\ \sin^3\theta = \frac{8\pi}{3} \tag{B-27}$$

We are left with the integral of the modulus k of $\boldsymbol{k}$. Using (B-25), (B-26), (B-27) and changing the variable k into the variable $\omega = ck$, we finally obtain:

$$W_{a\to b}^{\mathrm{abs}} = \frac{\mathcal{D}^2\omega_0^3}{3\varepsilon_0\hbar\pi c^3}\langle n(\omega_0)\rangle \tag{B-28}$$

C. Stimulated and spontaneous emissions

C-1. Emission rate

We now assume that the atom is initially in the upper state b, and that the radiation is in the initial Fock state (B-19); we study the emission processes where the atom falls back into the state a while emitting a photon. We still assume that the spectrum of the incident radiation is broad, so that its correlation function tends to zero in a time that is much shorter than Δt. The computations are then similar to the one we just did for the absorption process, but with a certain number of changes. First of all, we must now use expression (A-31) for the interaction Hamiltonian, the one that contains the negative frequency component of the field operator (with solely creation operators). Secondly, concerning the electric dipole operator, we must only keep the term connecting b to a, which amounts to replacing (B-13) by the complex conjugate expression:

$$\langle a\,|\tilde{\boldsymbol{D}}(t)|b\rangle = \boldsymbol{\mathcal{D}}^* e^{-i\omega_0 t} = \mathcal{D}\ \mathbf{e}_d^*\ e^{-i\omega_0 t} \tag{C-1}$$

The correlation function (B-17) of the field in normal order is now replaced by the correlation function in antinormal order $G_R^A(t', t'')$:

$$G_R^A(t', t'') = \left\langle \phi_{\mathrm{in}}^R \right|\ E_{\mathbf{d}}^{(+)}(\boldsymbol{R}, t')E_{\mathbf{d}^*}^{(-)}(\boldsymbol{R}, t'')\ \left|\phi_{\mathrm{in}}^R\right\rangle \tag{C-2}$$

For the radiation state (B-19), this function is given by:

$$\begin{aligned} G_R^A(t'-t'') &= \sum_i \frac{\hbar\omega_i}{2\varepsilon_0 L^3}\langle n_i\,|a_i a_i^\dagger|\,n_i\rangle\ |\mathbf{e}_d\cdot\boldsymbol{\varepsilon}_i|^2\ e^{-i\omega_i(t'-t'')} \\ &= \sum_i \frac{(\langle n_i\rangle+1)\hbar\omega_i}{2\varepsilon_0 L^3}\ |\mathbf{e}_d\cdot\boldsymbol{\varepsilon}_i|^2\ e^{-i\omega_i(t'-t'')} \end{aligned} \tag{C-3}$$

In the present case, it is now $a_i a_i^\dagger$ that is present in (C-3), and not $a_i^\dagger a_i$ as in (B-20). Since a_i and $a_i^\dagger$ do not commute, this leads to an important difference compared with expression (B-20) for $G_R^N(t'-t'')$: the number of photons $\langle n_i\rangle$ initially populating the mode i is replaced by $\langle n_i\rangle + 1$. It is the quantum character of the field (non-commuting operators) that is responsible for these essential differences between the absorption and emission processes.

The emission rate is obtained by computations similar to those that led to the absorption rate (B-25) (with a change of sign for ω_0 and for the ω_i). We obtain:

$$W_{b\to a}^{\mathrm{em}} = \frac{P_{b\to a}^{\mathrm{em}}(\Delta t)}{\Delta t} = 2\pi \frac{\mathcal{D}^2}{\hbar^2} \sum_i \frac{[\langle n_i\rangle + 1]\,\hbar\omega_i}{2\varepsilon_0 L^3} \left|\mathbf{e}_d \cdot \boldsymbol{\varepsilon}_i\right|^2 \delta(\omega_i - \omega_0) \tag{C-4}$$

This result differs from the absorption rate only by the replacement of $\langle n_i\rangle$ by $\langle n_i\rangle + 1$. This formula gives the general expression of the emission rate as a function of the population $\langle n_i\rangle$ of the modes and of their polarization $\boldsymbol{\varepsilon}_i$. We shall now discuss its two components, one proportional to $\langle n_i\rangle$, and one that does not depend on $\langle n_i\rangle$.

C-2. Stimulated emission

We first consider the terms in (C-4) that contain $\langle n_i\rangle$, i.e. the contribution of the modes containing initially at least one photon. These terms correspond to an emission induced by the incident radiation; their rate is proportional to the incident light intensity. For this reason it is called *stimulated* emission. Its rate is the same as the absorption rate, since the terms depending on $\langle n_i\rangle$ are identical in (B-24) and (C-4):

$$W_{b\to a}^{\mathrm{stim.em}} = W_{a\to b}^{\mathrm{abs}} \tag{C-5}$$

In particular, for isotropic radiation, we get a result identical to (B-28):

$$W_{b\to a}^{\mathrm{stim.em}} = \frac{\mathcal{D}^2\omega_0^3}{3\varepsilon_0\hbar\pi c^3}\langle n(\omega_0)\rangle \tag{C-6}$$

A photon resulting from stimulated emission is created in the same mode i as the photons inducing that emission: the number n_i of photons in that mode goes from n_i to $n_i + 1$. The added photon has the same energy, same direction and same polarization as the initial n_i photons. If the incident radiation is coherent, one can show that radiation emitted by stimulated emission has the same phase as the incident one. This results in a constructive interference effect (in the direction of the incident radiation) between the radiation emitted by the induced dipole and the incident radiation, hence leading to an amplification effect. For this phenomenon to occur, the atomic populations must be inverted, meaning that the probability of occupation of the upper level b must be larger than that of the lower level a. However, if this is not the case, the interference is destructive, which explains the attenuation of an incident beam by the absorption process. The coherent amplification by stimulated emission of an incident beam propagating through atoms with an inverted population plays an essential role in laser systems. The word laser is an acronym of "Light Amplification by Stimulated Emission of Radiation".

C-3. Spontaneous emission

If all the $\langle n_i\rangle$ are equal to zero, the radiation is initially in the vacuum state; the absorption rate (B-25) is then equal to zero. On the other hand, the emission rate (C-4) is not, because of the term 1 in $\langle n_i\,|a_i a_i^\dagger|\,n_i\rangle = \langle n_i\rangle + 1$. It follows that an atom, initially in the upper state b and placed in a vacuum of photons, has a non-zero probability per unit time of emitting a photon and falling back into the lower state a. This is called the *spontaneous emission* process.

The term 1 appearing in (C-4) is the same for all the modes, as opposed to the term $\langle n_i \rangle$ that only exists for the modes already containing photons. As this term 1 does not favor any particular direction or polarization, the situation is similar to that where the radiation is isotropic; the calculations leading to (B-28) remain valid provided we replace $\langle n(\omega_0) \rangle$ by 1. This leads to the rate of spontaneous emission:

$$W_{b\to a}^{\text{spont.em.}} = \frac{\mathcal{D}^2 \omega_0^3}{3\varepsilon_0 \hbar \pi c^3} \tag{C-7}$$

This result could also be obtained by Fermi's golden rule (see Chapter XIII, § C-3-b) in the following way. The system atom + radiation starts from the initial state $|b;0\rangle$ (atom in state b, photon vacuum). This discrete state is coupled to an infinity of final states $|a;1_i\rangle$, atom in state a with one photon in mode i. The golden rule enables one to compute the probability per unit time of the transition from the discrete initial state towards the continuum of final states, which is simply the spontaneous emission rate:

$$W_{b\to a}^{\text{em.spont.}} = \frac{2\pi}{\hbar} \sum_i |\langle a;1_i|H_I|b;0\rangle|^2 \, \delta(\hbar\omega_i - \hbar\omega_0) \tag{C-8}$$

Using $\langle b;0|H_I|a;1_i\rangle = i\mathcal{D}(\mathbf{e}_d \cdot \boldsymbol{\varepsilon}_i)(\hbar\omega_i/2\varepsilon_0 L^3)^{1/2}$, for a real polarization, as well as equations (B-26) and (B-27), we get the same result as (C-7).

This equation shows that the spontaneous emission rate increases as the cube of the atomic transition frequency; this explains why spontaneous emission is negligible in the radiofrequency domain, becomes important in the optical domain, and even more so in the ultraviolet or X ray domain. This ω_0^3 factor has two origins: on one hand, the square of the $\omega^{1/2}$ factor appearing in the electric field expression, on the other, the ω^2 factor present in the density of final states.

The spontaneous emission rate $W_{b\to a}^{\text{spont.em.}}$ is also called the "*natural width*" of the excited state b, and noted Γ. The inverse of Γ is the "*radiative lifetime*" of the excited state, which is the average time necessary for the atom to undergo radiative decay:

$$\Gamma = W_{b\to a}^{\text{spont.em.}} = \frac{1}{\tau_R} \tag{C-9}$$

It is instructive to compare the rate $W_{b\to a}^{\text{spont.em.}}$ to ω_0. It follows from (C-7) that:

$$\frac{\Gamma}{\omega_0} = \frac{\mathcal{D}^2 \omega_0^2}{3\varepsilon_0 \hbar \pi c^3} \tag{C-10}$$

The dipole $\mathcal{D}$ is of the order of qa_0 where q is the electron charge and a_0 the Bohr radius. This yields the quantity $q^2/(3\pi\varepsilon_0\hbar c)$ which, within a factor 4/3, is the fine-structure constant $\alpha \simeq 1/137$ multiplied by $a_0^2\omega_0^2/c^2$. Since $a_0\omega_0$ is of the order of the electron velocity in the first Bohr orbit, which is α times smaller than the speed of light c, $a_0^2\omega_0^2/c^2$ is of the order of α^2, so we finally have:

$$\frac{\Gamma}{\omega_0} \simeq \alpha^3 \tag{C-11}$$

The natural width of the excited level is therefore much smaller than the atomic transition frequency: the atomic dipole may oscillate a great number of times before these oscillations are damped. Typically, in the optical domain, Γ is of the order of 10^7 to 10^9 s^{-1} whereas $\omega_0/2\pi$ is much larger, of the order of 10^{14} to 10^{15} s^{-1}.

C-4. Einstein coefficients and Planck's law

Let us now assume the radiation is at thermal equilibrium at temperature T (black body radiation). In this case, it is current practice to use another notation for the various absorption and spontaneous or stimulated emission rates: the Einstein A and B coefficients from his 1917 article. In that article, he introduces for the first time the concept of stimulated emission:

$$W_{a\to b}^{\text{abs}} = B_{a\to b} \qquad W_{b\to a}^{\text{stim.em.}} = B_{b\to a} \qquad W_{b\to a}^{\text{spont.em.}} = A_{b\to a} \tag{C-12}$$

One can then write the change per unit time of the populations σ_{bb} and σ_{aa} of the b and a levels due to the various absorption and emission processes:

$$\begin{aligned} \dot{\sigma}_{bb} &= B_{a\to b}\sigma_{aa} - B_{b\to a}\sigma_{bb} - A_{b\to a}\sigma_{bb} \\ \dot{\sigma}_{aa} &= -B_{a\to b}\sigma_{aa} + B_{b\to a}\sigma_{bb} + A_{b\to a}\sigma_{bb} \end{aligned} \tag{C-13}$$

As an example, on the right-hand side of the first line of (C-13), the first term describes how level b fills up when a absorbs a photon, the second term how it is emptied by stimulated emission towards a, the third one how it is emptied by spontaneous emission. Similar explanations can be given for the second line.

In a steady state, there is a balance between the various processes, and we have:

$$\dot{\sigma}_{aa} = \dot{\sigma}_{bb} = 0 \tag{C-14}$$

We then get from the first equation:

$$\sigma_{bb} = \frac{B_{a\to b}}{B_{b\to a} + A_{b\to a}}\sigma_{aa} \tag{C-15}$$

Now, according to the Boltzmann distribution law, σ_{bb}/σ_{aa} must be equal to $e^{-(\hbar\omega_0/k_BT)}$. Relations (B-28), (C-6) and (C-7) then show that, at equilibrium, the populations σ_{bb} and σ_{aa} obey the relation:

$$\frac{\sigma_{bb}}{\sigma_{aa}} = e^{-(\hbar\omega_0/k_BT)} = \frac{B_{a\to b}}{B_{b\to a} + A_{b\to a}} = \frac{\langle n(\omega_0)\rangle}{\langle n(\omega_0)\rangle + 1} \tag{C-16}$$

which means that:

$$\langle n(\omega_0)\rangle = \frac{e^{-(\hbar\omega_0/k_BT)}}{1 - e^{-(\hbar\omega_0/k_BT)}} = \frac{1}{e^{(\hbar\omega_0/k_BT)} - 1} \tag{C-17}$$

Multiplying the average energy per mode $\hbar\omega_0\langle n(\omega_0)\rangle$ by the mode density $8\pi\nu_0^2/c^3$ in the vicinity of ω_0, yields Planck's law for the energy density per unit volume of the black body radiation, as a function of the frequency $\nu_0 = \omega_0/2\pi$:

$$u(\nu_0) = \frac{8\pi h\nu_0^3}{c^3}\frac{1}{e^{(h\nu_0/k_BT)} - 1} \tag{C-18}$$

In other words, when an ensemble of two-level atoms reaches Maxwell-Boltzmann equilibrium through absorption and spontaneous or stimulated emission of radiation, that radiation must necessarily obey Planck's law. This is the essence of the argument used by Einstein to establish this law[9].

[9]Einstein could not reason in 1917 in terms of the quantum theory of radiation, which was not available. The heuristic introduction of the A and B coefficients illustrates his remarkable intuition.

D. Role of correlation functions in one-photon processes

The probabilities associated with one-photon processes can be expressed in terms of atom and field correlation functions.

D-1. Absorption process

Let us write again the probability amplitude for the system, starting at $t_i = 0$ from the state $|\, a; \phi^R_{\rm in}\rangle$, to end up at time $t_f = \Delta t$ in the state $|b; \phi^R_{\rm fin}\rangle$ after undergoing a one-photon transition – see (B-14):

$$\langle \psi_{\rm fin}|\bar{U}(t_f,t_i)|\psi_{\rm in}\rangle = -\frac{1}{i\hbar}\int_0^{\Delta t} {\rm d}t'\, \langle b|\bar{\boldsymbol{D}}(t')|a\rangle \cdot \langle \phi^R_{\rm fin}|\bar{\boldsymbol{E}}_\perp(\boldsymbol{R},t')|\phi^R_{\rm in}\rangle \tag{D-1}$$

When dealing with an absorption of radiation by the atom, the number of photons in the final state $\left|\phi^R_{\rm fin}\right\rangle$ is lower than in the initial state $\left|\phi^R_{\rm in}\right\rangle$; only the positive frequency component of the field $\bar{\boldsymbol{E}}^{(+)}_\perp$, which can destroy photons, yields a contribution to the matrix element $\left\langle \phi^R_{\rm fin}\right| \bar{\boldsymbol{E}}_\perp(\boldsymbol{R},t') \left|\phi^R_{\rm in}\right\rangle$. Let us start with a radiation initial state containing photons with a single given polarization $\boldsymbol{\varepsilon}_{\rm in}$; only modes with this particular polarization are involved in the matrix element. If the polarization is not linear (circular for instance), $\boldsymbol{\varepsilon}_{\rm in}$ is complex, and we must replace it by its complex conjugate $\boldsymbol{\varepsilon}^*_{\rm in}$ is all negative frequency components of the field. Moreover, since $\boldsymbol{\varepsilon}^*_{\rm in} \cdot \boldsymbol{\varepsilon}_{\rm in} = 1$, the $\boldsymbol{\varepsilon}_{\rm in}$ polarization components are obtained by a scalar product of $\boldsymbol{\varepsilon}^*_{\rm in}$ by the field. We then have:

$$\langle \psi_{\rm fin}|\bar{U}(t_f,t_i)|\psi_{\rm in}\rangle = -\frac{1}{i\hbar}\int_0^{\Delta t} {\rm d}t'\, \langle b|\, \boldsymbol{\varepsilon}_{\rm in} \cdot \bar{\boldsymbol{D}}(t')|a\rangle \langle \phi^R_{\rm fin}|\boldsymbol{\varepsilon}^*_{\rm in} \cdot \bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')|\phi^R_{\rm in}\rangle \tag{D-2}$$

The probability $P^{\rm abs}_{a\to b}(\Delta t)$ of the absorption process is then:

$$\begin{aligned} P^{\rm abs}_{a\to b}(\Delta t) = \frac{1}{\hbar^2}\int_0^{\Delta t} {\rm d}t' \langle a|\, \boldsymbol{\varepsilon}^*_{\rm in} \cdot \bar{\boldsymbol{D}}(t')|b\rangle \langle \phi^R_{\rm in}|\boldsymbol{\varepsilon}_{\rm in} \cdot \bar{\boldsymbol{E}}^{(-)}_\perp(\boldsymbol{R},t')|\phi^R_{\rm fin}\rangle \\ \times \int_0^{\Delta t} {\rm d}t'' \langle b|\, \boldsymbol{\varepsilon}_{\rm in} \cdot \bar{\boldsymbol{D}}(t'')|a\rangle \langle \phi^R_{\rm fin}|\boldsymbol{\varepsilon}^*_{\rm in} \cdot \bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t'')|\phi^R_{\rm in}\rangle \end{aligned} \tag{D-3}$$

To obtain the probability of finding the atom in any final state b other than a, whatever final state $|\phi^R_{\rm fin}\rangle$ the radiation is in, we must sum this result over all possible $|b\rangle$ and $|\phi^R_{\rm fin}\rangle$ states. This yields two closure relations, one in the atom state space[10], the other in the radiation state space. This leads to the following result:

$$\sum_b P^{\rm abs}_{a\to b}(\Delta t) = \frac{1}{\hbar^2}\int_0^{\Delta t} {\rm d}t' \int_0^{\Delta t} {\rm d}t''\; G_{\rm a}(t',t'')\; G^N_R(t',t'') \tag{D-4}$$

with the definitions:

$$G_{\rm a}(t',t'') = \langle a|\left[\boldsymbol{\varepsilon}^*_{\rm in} \cdot \bar{\boldsymbol{D}}(t')\right]\left[\boldsymbol{\varepsilon}_{\rm in} \cdot \bar{\boldsymbol{D}}(t'')\right]|a\rangle \tag{D-5}$$

[10] In the summation over b, the initial atomic state a can be included since the matrix element of the atomic dipole in that state is zero (because of a parity argument).

and:

$$G_R^N(t',t'') = \left\langle \phi_{\rm in}^R \right| \left[\boldsymbol{\varepsilon}_{\rm in} \cdot \bar{\boldsymbol{E}}_\perp^{(-)}(\boldsymbol{R},t') \right] \left[\boldsymbol{\varepsilon}_{\rm in}^* \cdot \bar{\boldsymbol{E}}_\perp^{(+)}(\boldsymbol{R},t'') \right] \left| \phi_{\rm in}^R \right\rangle \tag{D-6}$$

The two functions we just defined correspond, respectively, to the correlation function of the atomic dipole and to that of the electric field expressed in normal order.

In the more general case where the field initial state includes several polarizations, (D-1) must now include the matrix elements:

$$\sum_{j=x,y,z} \langle b| \ \bar{D}_j(t') \left|a\right\rangle \left\langle \phi_{\rm fin}^R \right| \bar{E}_j(\boldsymbol{R},t') \left| \phi_{\rm in}^R \right\rangle \tag{D-7}$$

with:

$$\bar{D}_j(t) = \mathbf{e}_j \cdot \bar{\boldsymbol{D}}(t) \quad \text{and} \quad \bar{E}_j(\boldsymbol{R},t) = \mathbf{e}_j \cdot \bar{\boldsymbol{E}}_\perp(\boldsymbol{R},t) \tag{D-8}$$

where $\mathbf{e}_j$ is the unit vector of each of the three $j = x, y, z$ axes. Probability (D-4) then becomes:

$$\sum_b P_{a\to b}^{\rm abs}(\Delta t) = \sum_{j,j'=x,y,z} \frac{1}{\hbar^2} \int_0^{\Delta t} {\rm d}t' \int_0^{\Delta t} {\rm d}t'' \ G_{\rm a}^{jj'}(t',t'') \ G_R^{jj'}(t',t'') \tag{D-9}$$

with the following definitions of the 9 dipole correlation functions, and the other 9 field correlation functions (the vectors $\mathbf{e}_j$ are real) :

$$\begin{aligned} G_{\rm a}^{jj'}(t',t'') &= \langle a| \left[\mathbf{e}_j \cdot \bar{\boldsymbol{D}}(t') \right] \left[\mathbf{e}_{j'} \cdot \bar{\boldsymbol{D}}(t'') \right] |a\rangle \\ G_R^{jj'}(t',t'') &= \left\langle \phi_{\rm in}^R \right| \left[\mathbf{e}_j \cdot \bar{\boldsymbol{E}}_\perp^{(-)}(\boldsymbol{R},t') \right] \left[\mathbf{e}_{j'} \cdot \bar{\boldsymbol{E}}_\perp^{(+)}(\boldsymbol{R},t'') \right] \left| \phi_{\rm in}^R \right\rangle \end{aligned} \tag{D-10}$$

We thus get a correlation tensor, which slightly complicates the equations, but does not change the essence of the results.

D-2. Emission process

For the photon emission processes, spontaneous or stimulated, we can make similar computations. The main difference is that the $\bar{\boldsymbol{E}}_\perp^{(+)}$ and $\bar{\boldsymbol{E}}_\perp^{(-)}$ operators must be interchanged, which yields antinormal instead of normal field correlation functions; furthermore, we must use the more general formula (D-9) instead of (D-4) since an emission process does not favor any specific polarization.

E. Photon scattering by an atom

We now consider a photon scattering process where the initial state includes an atom in state a and an incident photon, and where in the final state the atom is still in state a, but the incident photon has been replaced by another one. This is a two-photon process, since one photon disappears, and is replaced by another one.

E-1. Elastic scattering

As in § B, we shall treat classically the center of mass of the atom, supposed to be fixed at the origin of the reference frame (see however the comment on page 2094). The initial state $|\psi_{\rm in}\rangle$ of the system atom + radiation at time t_i corresponds to an atom in state a in the presence of a photon in mode i, with wave vector $\boldsymbol{k}_i$, polarization $\boldsymbol{\varepsilon}_i$ and frequency ω_i (we assume the radiation to be monochromatic):

$$|\psi_{\rm in}\rangle = |a; \boldsymbol{k}_i, \boldsymbol{\varepsilon}_i\rangle \qquad \text{with energy} \qquad E_{\rm in} = E_a + \hbar\omega_i \tag{E-1}$$

At the final time t_f, the scattering process has replaced the initial photon $\boldsymbol{k}_i, \boldsymbol{\varepsilon}_i$ by a new photon $\boldsymbol{k}_f, \boldsymbol{\varepsilon}_f$ with wave vector $\boldsymbol{k}_f$ and polarization $\boldsymbol{\varepsilon}_f$. The final state of the system is:

$$|\psi_{\rm fin}\rangle = |a; \boldsymbol{k}_f, \boldsymbol{\varepsilon}_f\rangle \qquad \text{with energy} \qquad E_{\rm fin} = E_a + \hbar\omega_f \tag{E-2}$$

Conservation of total energy requires $\omega_i = \omega_f$. This scattering thus occurs without frequency change and is called for that reason "*elastic scattering*". We shall study in § E-3 the case where the final atomic state is different from the initial atomic state. This different process is called "*Raman scattering*".

As the electric dipole interaction Hamiltonian (A-24) can only change the photon number by one unit, the system must go through an intermediate state often called a "relay state" where the atom is in state c and the radiation in a state different from its initial state. The lowest order term of the series expansion (A-12) that contributes to the scattering amplitude is of order two.

E-1-a. Two possible types of relay state

There are two possible types of relay states: those corresponding to processes we shall label (α), where the photon $\boldsymbol{k}_i, \boldsymbol{\varepsilon}_i$ is absorbed before the photon $\boldsymbol{k}_f, \boldsymbol{\varepsilon}_f$ is emitted; and those corresponding to processes labeled (β), where the photon $\boldsymbol{k}_f, \boldsymbol{\varepsilon}_f$ is emitted before the photon $\boldsymbol{k}_i, \boldsymbol{\varepsilon}_i$ is absorbed. In the first case, the relay state is the state $|\psi^{\alpha}_{\rm rel}\rangle = |c; 0\rangle$, where c is an atomic relay state and 0 is the radiation vacuum, since the photon $\boldsymbol{k}_i, \boldsymbol{\varepsilon}_i$ present in the initial state has been absorbed; the energy of this relay state is $E^{\alpha}_{\rm rel} = E_c$. In the second case, the relay state is $|\psi^{\beta}_{\rm rel}\rangle = |c; \boldsymbol{k}_i, \boldsymbol{\varepsilon}_i; \boldsymbol{k}_f, \boldsymbol{\varepsilon}_f\rangle$, since the photon $\boldsymbol{k}_f, \boldsymbol{\varepsilon}_f$ has been emitted before the photon $\boldsymbol{k}_i, \boldsymbol{\varepsilon}_i$ was absorbed: the energy of this relay state is $E^{\beta}_{\rm rel} = E_c + \hbar\omega_i + \hbar\omega_f$. Figures 2 and 3 show two different diagrams representing these same two processes.

In Figure 2, the horizontal lines represent the atomic levels; an upwards arrow represents an absorption, whereas a downwards arrow represents an emission. The advantage of this representation is to directly show the energy difference between the initial state and the relay state, equal to $E_{\rm in} - E^{\alpha}_{\rm rel} = E_a + \hbar\omega_i - E_c$ for the (α) processes, and to $E_{\rm in} - E^{\beta}_{\rm rel} = E_a - \hbar\omega_f - E_c$ for the (β) processes: this difference is simply the distance between the height of the dashed line and the height of the line representing the atomic relay state c. In particular, these two lines coincide for the (α) processes when $E_a + \hbar\omega_i = E_c$, i.e. when the absorption of the incident photon is resonant for the transition $a \leftrightarrow c$ (resonant scattering, which will be studied later on).

In Figure 3, an incoming arrow represents an absorption, whereas an outgoing one represents an emission. Reading the diagram from bottom to top, one clearly sees which atomic state and photons are present in the initial state, the relay state and the final

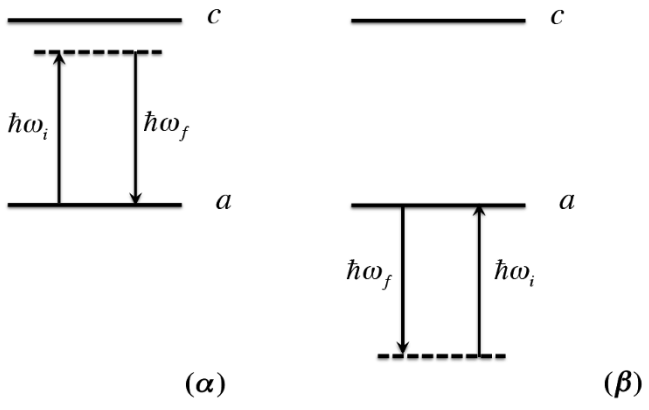


Figure 2: First diagram representation of the scattering processes labeled (α) *and* (β) *in the text; these processes have a different chronological order for the absorption of the incident photon and the emission of the scattered photon. The full horizontal lines represent the atomic levels, and the upwards arrows represent absorption processes and downwards arrows, emission processes; the horizontal dashed lines clearly show the energy differences that will appear in the denominator of the transition amplitude expression.*

state. For the (α) processes, no photons are present in the relay state, whereas both incident and scattered photons are present in that state in the (β) processes.

E-1-b. Computation of the scattering amplitude

The computation of the scattering amplitude is of the same type as the calculations already presented above. In addition, it is almost identical to the computation of the two-photon absorption amplitude explained in detail in § 1 of Complement $\mathrm{A_{XX}}$. Consequently, it will not be explicitly carried out here, but we shall merely highlight the differences with the computation of that complement. The reader interested in more details may want to read that complement before continuing with this paragraph. Relations (13) and (14) of that complement are written here:

$$\begin{aligned}\langle\psi_{\mathrm{fin}}|\bar{U}(t_f,t_i)|\psi_{\mathrm{in}}\rangle &= -2i\pi\left[A^{\alpha}_{fi}(E_{\mathrm{in}})+A^{\beta}_{fi}(E_{\mathrm{in}})\right]\delta^{(\Delta t)}(E_{\mathrm{fin}}-E_{\mathrm{in}})\\ &= -2i\pi\left[A^{\alpha}_{fi}(E_{\mathrm{in}})+A^{\beta}_{fi}(E_{\mathrm{in}})\right]\delta^{(\Delta t)}(\hbar\omega_f-\hbar\omega_i)\end{aligned} \tag{E-3}$$

where we have introduced the probability amplitudes:

$$\begin{aligned}A^{\alpha}_{fi}(E_{\mathrm{in}}) &= \sum_{|\psi^{\alpha}_{\mathrm{rel}}\rangle}\frac{\langle\psi_{\mathrm{fin}}|\boldsymbol{D}\cdot\boldsymbol{E}^{(-)}|\psi^{\alpha}_{\mathrm{rel}}\rangle\langle\psi^{\alpha}_{\mathrm{rel}}|\boldsymbol{D}\cdot\boldsymbol{E}^{(+)}|\psi_{\mathrm{in}}\rangle}{E_{\mathrm{in}}-E^{\alpha}_{\mathrm{rel}}}\\ &= \frac{\hbar\sqrt{\omega_i\omega_f}}{2\epsilon_0 L^3}\sum_c\frac{\langle a|\boldsymbol{\varepsilon}^*_f\cdot\boldsymbol{D}|c\rangle\langle c|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|a\rangle}{E_a+\hbar\omega_i-E_c}\end{aligned} \tag{E-4}$$

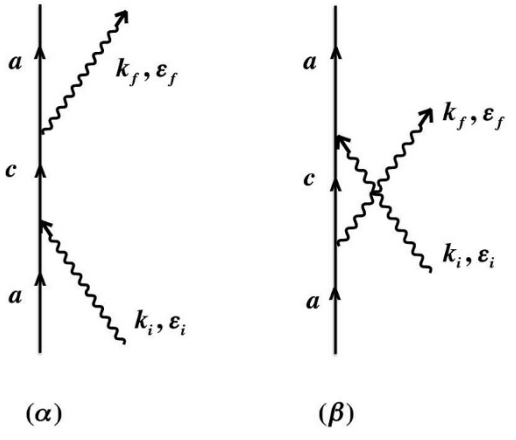


Figure 3: Another possible diagram representation of the scattering processes (α) *and* (β)*. The state of the atom is shown on the vertical full line. As for the photons, the incoming arrows (wiggly arrows pointing towards the vertical full line) represent absorption, whereas outgoing ones (wiggly arrows leaving the vertical full line) represent emissions. The bottom to top reading of each diagram shows the chronological succession of the states of the system atom + photons.*

$$A_{fi}^{\beta}(E_{\rm in}) = \sum_{|\psi_{\rm rel}^{\beta}\rangle} \frac{\langle\psi_{\rm fin}|\boldsymbol{D}\cdot\boldsymbol{E}^{(+)}|\psi_{\rm rel}^{\beta}\rangle\langle\psi_{\rm rel}^{\beta}|\boldsymbol{D}\cdot\boldsymbol{E}^{(-)}|\psi_{\rm in}\rangle}{E_{\rm in}-E_{\rm rel}^{\beta}}$$
$$= \frac{\hbar\sqrt{\omega_i\omega_f}}{2\epsilon_0 L^3}\sum_c \frac{\langle a|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|c\rangle\langle c|\boldsymbol{\varepsilon}_f^*\cdot\boldsymbol{D}|a\rangle}{E_a-\hbar\omega_f-E_c} \qquad \text{(E-5)}$$

The two amplitudes $A_{fi}^{\alpha}(E_{\rm in})$ and $A_{fi}^{\beta}(E_{\rm in})$ correspond to the two types of relay states considered above. In (E-3), the delta function expressing the energy conservation is proportional to $\delta^{(\Delta t)}(\omega_f-\omega_i)$ since $E_{\rm fin}-E_{\rm in}=\hbar(\omega_f-\omega_i)$. To write (E-4) and (E-5), we have replaced in relation (13) of Complement $\mathrm{A_{XX}}$ the interaction Hamiltonian H_I by $-\boldsymbol{D}\cdot\boldsymbol{E}^{(+)}$ or $-\boldsymbol{D}\cdot\boldsymbol{E}^{(-)}$, depending on whether it pertains to an absorption or an emission process. In the numerator of the fractions on the right-hand side of (E-4), operator $\boldsymbol{E}^{(+)}$ (which absorbs the incident photon) acts before operator $\boldsymbol{E}^{(-)}$ (which creates the scattered photon); this is to be expected for an (α) type process. The order of the two operators $\boldsymbol{E}^{(+)}$ and $\boldsymbol{E}^{(-)}$ is reversed in (E-5), as expected for a (β) type process. The $\sqrt{\omega_i\omega_f}$ coefficients on the second lines of these equalities come from the plane wave expansion (A-27) of the electric field; L is the edge of the cubic cavity used to quantize the field. We could have added a factor $\sqrt{n_i n_f}$, where n_i and n_f are the photon numbers in the initial and final states; for the sake of simplicity, we have assumed that $n_i=n_f=1$.

E-1-c. Semiclassical interpretation

Elastic scattering can also be explained by a semiclassical treatment, where the quantum treatment only applies to the atom; the incident wave is described as a classical field of frequency ω_i. This wave induces, in the atom, an oscillating dipole at the same frequency. This dipole radiates into the entire space a field oscillating at that frequency.

The semiclassical approach also enables a simple interpretation of the absorption of the incident beam as resulting from a destructive interference, in the direction of the incident field, between that field and the field scattered by the induced dipole. One can also use such a description to account for the amplification of an incident beam by an ensemble of atoms whose population has been "inverted", meaning atoms for which the population of an excited level is larger than the population of a lower energy level. The scattered field then has the opposite phase of that it would have without population inversion, so that the interference becomes constructive.

E-1-d. Rayleigh scattering

Assume that the frequency ω_i of the radiation is much smaller than all the atomic frequencies $|E_c - E_a|/\hbar$. One can then ignore ω_i and ω_f in the denominators on the second lines of (E-4) and (E-5). The only ω_i dependence of the scattering amplitudes comes from the prefactor $\sqrt{\omega_i \omega_f}$, equal to ω_i since $\omega_f = \omega_i$. The scattering cross section involves the product of the squared modulus of that amplitude, proportional to ω_i^2, by the density of the radiation's final states at frequency $\omega_f = \omega_i$, also proportional to ω_i^2. The scattering cross section therefore varies as ω_i^4, much higher for blue light than for red light.

One usually calls "Rayleigh scattering" the elastic scattering when $\omega_i \ll |E_c - E_a|/\hbar$. It explains the scattering of the visible solar light by the atmospheric oxygen and nitrogen molecules, which have much higher resonant frequencies, in the ultraviolet domain. This rapid variation with frequency of the Rayleigh scattering cross section is a reason for the sky being blue.

E-2. Resonant scattering

Assume now that the frequency ω_i of the incident photon is very close to the frequency:

$$\omega_{ca} = (E_c - E_a)/\hbar \tag{E-6}$$

of a transition between a state a and a state c having a higher energy. The absorption of the incident photon is then resonant for the transition $a \leftrightarrow c$ and the amplitude (E-4) becomes very large when the state c becomes the atomic relay state – it even diverges if the resonant condition is exactly satisfied. In that case, one can neglect all the (β) type processes; in addition, even if there are other possible atomic relay states d, e,.., one can keep only the term involving c.

To avoid the difficulties related to the divergence of (E-4) when $\omega_i = \omega_{ca}$, it is convenient to use the exact expression (A-14), which only involves three terms, instead of an infinity as in expansion (A-12) for the evolution operator. Only the last of those three terms plays a role, since it can destroy a photon and create another one. A computation similar to that leading to relation (6) of Complement $\mathrm{A_{XX}}$, but using (A-14) instead of

(A-12), then yields:

$$\langle\psi_{\rm fin}|U(t_f,t_i)|\psi_{\rm in}\rangle = \left(\frac{1}{i\hbar}\right)^2 \int_{t_i}^{t_f} {\rm d}t \int_{t_i}^{t} {\rm d}t' \; \langle\psi_{\rm fin}|U_0(t_f,t)\, H_I\, U(t,t')H_I\, U_0(t',t_i)|\psi_{\rm in}\rangle \tag{E-7}$$

Note that it is now $U(t,t')$ that appears in the middle of the matrix element in (E-7), and not $U_0(t,t')$. Starting from $|\psi_{\rm in}\rangle = |a;\boldsymbol{k}_i,\boldsymbol{\varepsilon}_i\rangle$, the first interaction Hamiltonian H_I of (E-7) brings the system to the state $|c;0\rangle$. In a similar way, if the system starts from $\langle\psi_{\rm fin}|$ the second interaction Hamiltonian of (E-7) brings it to the state $\langle c;0|$. Expression (E-7) can then be written:

$$\begin{aligned}\langle\psi_{\rm fin}|U(t_f,t_i)|\psi_{\rm in}\rangle = \left(\frac{1}{i\hbar}\right)^2 \int_{t_i}^{t_f} {\rm d}t \int_{t_i}^{t} {\rm d}t' \langle a;\boldsymbol{k}_f,\boldsymbol{\varepsilon}_f|U_0(t_f,t)\, H_I\,|c;\,0\rangle \\ \times \langle c;\,0|U(t,t')|c;\,0\rangle\langle c;\,0|H_I\, U_0(t',t_i)|a;\boldsymbol{k}_i,\boldsymbol{\varepsilon}_i\rangle\end{aligned} \tag{E-8}$$

Had we used (A-12) instead of (A-14), the central matrix element of relation (E-8) would be $\langle c;\,0|U_0(t,t')|c;\,0\rangle = \exp(-iE_c(t-t')$. In our case, using (A-14) leads to $\langle c;\,0|U(t,t')|c;\,0\rangle$ which is the probability amplitude for the system, starting from the state $|c;\,0\rangle$ at time t', to still be in that same state at time t. The calculation of that amplitude appears in the study of the radiative decay of the excited state through spontaneous emission of a photon, hence the decay of a discrete state $|c;\,0\rangle$ coupled to a continuum of final states $|a;\boldsymbol{k},\boldsymbol{\varepsilon}\rangle$. Those states represent the atom in state a in the presence of a photon with any wave vector $\boldsymbol{k}$ and polarization $\boldsymbol{\varepsilon}$. Now we showed in Complement $\rm D_{XIII}$ (§ 4) that it was possible to obtain a solution of Schrödinger's equation yielding the amplitude $\langle c;\,0|U(t,t')|c;\,0\rangle$ at long times (and not only at short times, as for the perturbative solution). This solution is written:

$$\langle c;\,0|U(t,t')|c;\,0\rangle = \exp\left[-i(E_c+\delta E_c)(t-t')/\hbar\right]\exp\left[-\Gamma_c(t-t')/2\right] \tag{E-9}$$

where δE_c is the energy shift of the state $|c;\,0\rangle$ due to its coupling with the continuum of final states[11], and Γ_c the natural width of the excited state c (the inverse of the radiative lifetime of that state). We shall assume from now on that the shift δE_c is included in the definition of the energy E_c of the state c. Starting from the more precise expression (A-14) instead of (A-12) thus leads to a very simple result: we just have to replace, in all the computations of the scattering amplitude, E_c by $E_c - i\hbar\Gamma_c/2$.

$$E_c \to E_c - i\hbar\Gamma_c/2 \tag{E-10}$$

Once this replacement has been made, we get, keeping only the amplitude (E-4) and a single relay state c, the following scattering amplitude:

$$\bar{A}_{fi}(E_{\rm in}) = \frac{\hbar\omega_i}{2\varepsilon_0 L^3}\, \frac{\langle a|\boldsymbol{\varepsilon}_f^*\cdot\boldsymbol{D}|c\rangle\langle c|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|a\rangle}{\hbar(\omega_i-\omega_{ca}+i\Gamma_c/2)} \tag{E-11}$$

This resonant scattering amplitude no longer diverges when $\omega_i = \omega_{ca}$; as ω_i is scanned around ω_{ca}, it exhibits a resonant behavior over a range equal to Γ_c.

[11] This shift is related to the "Lamb shift" of the excited atomic states.

E-3. Inelastic scattering - Raman scattering

We now consider a scattering process where, as before, an incident photon is absorbed and another emitted, but the final atomic state a' is now supposed to be different from the initial atomic state a.

E-3-a. Differences with elastic scattering

Figure 4 shows an example of such a process, called "*Raman scattering*", where the energy of the scattered photon is different from that of the incident photon[12]. The initial state of the scattering process is, as before, the state $|\psi_{\rm in}\rangle = |a; \boldsymbol{k}_i, \boldsymbol{\varepsilon}_i\rangle$, with energy $E_{\rm in} = E_a + \hbar\omega_i$; the final state, however, is now $|\psi_{\rm fin}\rangle = |a'; \boldsymbol{k}_f, \boldsymbol{\varepsilon}_f\rangle$ where $a' \neq a$, with energy $E_{\rm fin} = E_{a'} + \hbar\omega_f$.

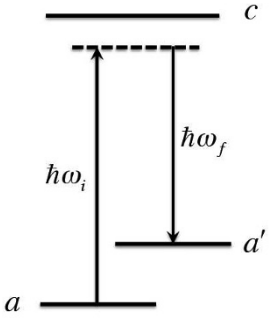


Figure 4: Raman scattering: an atom in state a absorbs an incident photon, with energy $\hbar\omega_i$; a photon $\hbar\omega_f$ is then spontaneously emitted by the atom, which ends up into a final state a' different from the initial state a.

Conservation of total energy requires:

$$E_a + \hbar\omega_i = E_{a'} + \hbar\omega_f \tag{E-12}$$

If $E_{a'} > E_a$, the Raman scattering is called "*Raman Stokes scattering*"; the energy of the scattered photon is lower than that of the incident photon. If $E_{a'} < E_a$, the Raman scattering is called "*Raman anti-Stokes scattering*"; the energy of the scattered photon is higher than that of the incident photon. As we assumed here that the mode $(\boldsymbol{k}_f, \boldsymbol{\varepsilon}_f)$ was initially empty, the scattered photon is emitted spontaneously. The process is then called "*spontaneous Raman scattering*". We shall study later the case where $(\boldsymbol{k}_f, \boldsymbol{\varepsilon}_f)$ photons are initially present, a situation resulting in "*stimulated Raman scattering*".

Equation (E-12) shows that the angular frequency of the scattered light is different from that of the incident light by a quantity $\omega_{aa'} = (E_a - E_{a'})/\hbar$, equal to the frequency of the atomic transition $a \leftrightarrow a'$. This means that Raman light spectrum provides information about the eigenfrequencies of the scattering system; this is the base of *Raman*

[12] This figure only shows a type (α) process, where the incident photon is absorbed in the first place.

spectroscopy. The systems under study are often molecules and the states a, a' are vibrational or rotational sublevels of the ground state, which means that the $\omega_{aa'}$ frequencies belong to the microwave or infrared domain. Instead of measuring directly these frequencies using spectroscopic techniques in a frequency domain where detection might be a problem, it is sometimes more convenient to illuminate the medium with an optical or ultraviolet frequency beam, and to measure the $\omega_{aa'}$ frequency via the frequency shifts of the Raman scattered light. Raman spectroscopy developed considerably once laser sources became available, yielding much higher signal intensities. The detection conditions have also been greatly improved by focusing the incident laser light on very small volumes. The analysis of the scattered light spectrum is now a very powerful tool for analyzing the chemical composition of any scattering media, since each molecule can be identified by its specific vibration-rotation eigenfrequencies.

To keep things simple, we have limited our discussion to Raman scattering by atoms or molecules in a dilute medium, where each scattering entity acts individually. Raman scattering is also used in condensed media such as liquids, crystals, surfaces, etc. and provides valuable information on the dynamics of these structures.

E-3-b. Scattering amplitude

The computation of the Raman scattering amplitude is very similar to the one leading to (E-3), (E-4) and (E-5), and yields:

$$\begin{aligned}\langle\psi_{\rm fin}|\bar{U}(t_f,t_i)|\psi_{\rm in}\rangle &= -2\pi i\left[A^{\alpha}_{fi}(E_{\rm in})+A^{\beta}_{fi}(E_{\rm in})\right]\delta^{(\Delta t)}(E_{\rm fin}-E_{\rm in})\\ &= -2\pi i\left[A^{\alpha}_{fi}(E_{\rm in})+A^{\beta}_{fi}(E_{\rm in})\right]\delta^{(\Delta t)}(E_{a'}+\hbar\omega_f-E_a-\hbar\omega_i)\end{aligned}\quad\text{(E-13)}$$

where:

$$\begin{aligned}A^{\alpha}_{fi}(E_{\rm in}) &= \sum_{|\psi^{\alpha}_{\rm rel}\rangle}\frac{\langle\psi_{\rm fin}|\boldsymbol{D}\cdot\boldsymbol{E}^{(-)}|\psi^{\alpha}_{\rm rel}\rangle\langle\psi^{\alpha}_{\rm rel}|\boldsymbol{D}\cdot\boldsymbol{E}^{(+)}|\psi_{\rm in}\rangle}{E_{\rm in}-E^{\alpha}_{\rm rel}}\\ &= \frac{\hbar\sqrt{\omega_i\omega_f}}{2\epsilon_0 L^3}\sum_c\frac{\langle a'|\boldsymbol{\varepsilon}_f^*\cdot\boldsymbol{D}|c\rangle\langle c|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|a\rangle}{E_a+\hbar\omega_i-E_c}\end{aligned}\quad\text{(E-14)}$$

$$\begin{aligned}A^{\beta}_{fi}(E_{\rm in}) &= \sum_{|\psi^{\beta}_{\rm rel}\rangle}\frac{\langle\psi_{\rm fin}|\boldsymbol{D}\cdot\boldsymbol{E}^{(+)}|\psi^{\beta}_{\rm rel}\rangle\langle\psi^{\beta}_{\rm rel}|\boldsymbol{D}\cdot\boldsymbol{E}^{(-)}|\psi_{\rm in}\rangle}{E_{\rm in}-E^{\beta}_{\rm rel}}\\ &= \frac{\hbar\sqrt{\omega_i\omega_f}}{2\epsilon_0 L^3}\sum_c\frac{\langle a'|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|c\rangle\langle c|\boldsymbol{\varepsilon}_f^*\cdot\boldsymbol{D}|a\rangle}{E_a-\hbar\omega_f-E_c}\end{aligned}\quad\text{(E-15)}$$

When the photon frequency ω_i is close to the $a\to c$ transition frequency, Raman scattering becomes resonant and amplitude (E-14) can become very large. As we did for the resonant elastic scattering, to avoid the divergence of (E-14), we just have to replace E_c by $E_c-i\Gamma_c/2$, where Γ_c is the natural width of the c state.

E-3-c. Semiclassical interpretation

As in § E-1-c, let us consider the dipole induced by the incident field on the scattering object. When that object is a molecule which vibrates and rotates, its polarizability

changes with time, and is modulated by its rotation and vibration frequencies. The dipole's oscillations induced by the incident field have an amplitude modulated at the rotation and vibration frequencies of the molecule. The Fourier spectrum of the dipole's motion contains lateral bands at frequencies shifted from the incident field frequency; these frequency shifts are equal to the molecule's rotation and vibration frequencies. This semiclassical interpretation accounts for the essential properties of the Raman scattering spectrum.

E-3-d. Stimulated Raman scattering. Raman laser

We now assume the Raman photon appears in a mode that is not initially empty, as $n_f \neq 0$ photons already occupy the mode $(\boldsymbol{k}_f, \boldsymbol{\varepsilon}_f)$. Similarly, we assume several photons (n_i for example), initially occupy the mode $(\boldsymbol{k}_i, \boldsymbol{\varepsilon}_i)$. To compute the Raman scattering amplitude $|a; n_i, n_f\rangle \rightarrow |a'; n_i - 1, n_f + 1\rangle$, we must include the factor $\sqrt{n_i(n_f + 1)}$ in expressions (E-14) and (E-15). Stimulated emission now comes into play: the factor $n_f + 1$ appearing in the probability expresses the fact that the initial presence of n_f photons in mode $(\boldsymbol{k}_f, \boldsymbol{\varepsilon}_f)$ stimulates the emission probability of a Raman photon in that mode.

Consider now (right-hand side of Fig.5) the inverse scattering process symbolized by $|a'; n_i - 1, n_f + 1\rangle \rightarrow |a; n_i, n_f\rangle$. The corresponding scattering amplitude is simply the complex conjugate of the previous one, meaning that the probability of these two processes are equal. If we start with the same number of atoms in state a and state a', the number of photons created in one of the processes is equal to the number of photons that

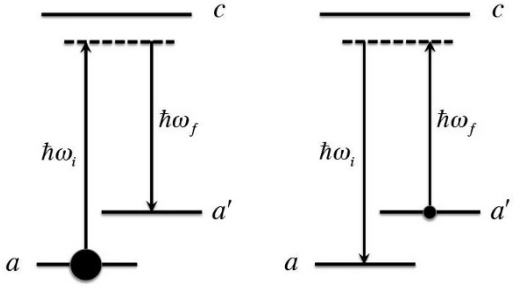


Figure 5: The left-hand side of the figure represents a stimulated Raman process where an atom in level a absorbs a photon ω_i, and ends up in state a' after the stimulated emission of an ω_f photon. The right-hand side of the figure shows the inverse process, where an atom starting from state a' absorbs an ω_f photon, and falls back in state a after the stimulated emission of an ω_i photon. These two processes have, a priori, the same probability. However, if the population of state a (shown as a large dot on the left-hand side of the figure) is higher than that of the state a' (shown as a small dot on the right-hand side of the figure), the number of processes resulting in the stimulated emission of an ω_f photon is bigger than the number of processes where that photon is absorbed. The radiation at frequency ω_f is thus amplified by stimulated emission, which allows creating a "Raman laser" at this frequency.

disappear in the inverse process. What will happen now if we start with an ensemble of atoms where the populations of states a and a' are not equal? If, for example, level a has a lower energy than level a', the relaxation mechanisms leading to thermal equilibrium tend to create a larger population in a than in a'. The number of scattering processes $|a; n_i, n_f\rangle \rightarrow |a'; n_i - 1, n_f + 1\rangle$ is then larger than the number of the inverse processes $|a'; n_i - 1, n_f + 1\rangle \rightarrow |a; n_i, n_f\rangle$, leading to an amplification of the number of ω_f photons. This amplification mechanism is the basis of *Raman laser* operation. This type of laser is different in two major ways from lasers involving a transition between an upper level, populated by a pumping process, and a lower level (with no relay level). First of all, they do not require a population inversion; the atomic media can be at thermal equilibrium, since the stimulated Raman scattering at the origin of the amplification starts from the atomic state with the largest population. They do require, however, a high intensity radiation at frequency ω_i, furnished by another laser called the "pump laser". Secondly, the ω_f frequency of the Raman laser oscillation can be scanned by changing the "pump frequency" ω_i, whereas lasers using a two-level system necessarily oscillate at a frequency very close to the atomic transition, and have thus a very small tuning range.

Conservation of total momentum

If the position $\boldsymbol{R}$ of the atomic center of mass is no longer treated classically, and placed at the origin as we have done until now, we must keep the exponential functions $\exp(i\boldsymbol{k}_i \cdot \boldsymbol{R})$ and $\exp(-i\boldsymbol{k}_f \cdot \boldsymbol{R})$ in the interaction Hamiltonian describing the absorption of an ω_i photon and the emission of an ω_f photon. The matrix element of the product of those two operators must be taken between an initial state of the center of mass, with momentum $\hbar\boldsymbol{K}_{\text{in}}$ and a final state with momentum $\hbar\boldsymbol{K}_{\text{fin}}$. This yields a $\delta(\boldsymbol{K}_{\text{fin}} - \boldsymbol{K}_{\text{in}} - \boldsymbol{k}_i + \boldsymbol{k}_f)$ function expressing the global momentum conservation in a Raman process: the momentum of the atom increases by the quantity $\hbar(\boldsymbol{k}_i - \boldsymbol{k}_f)$ during that process. It often happens that the two atomic states a and a' are two sublevels of the same electronic ground state, so that the frequency $\omega_{a'a} = (E_{a'} - E_a)/\hbar$ falls in the microwave domain; it is then much smaller than the frequencies ω_i and ω_f, which are optical frequencies. The energy conservation equation (E-12) then shows that the moduli of the two wave vectors $\boldsymbol{k}_i$ and $\boldsymbol{k}_f$ are practically the same. If the two wave vectors $\boldsymbol{k}_i$ and $\boldsymbol{k}_f$ have opposite directions, the momentum gained by the atom during a Raman process is equal to $\hbar(\boldsymbol{k}_i - \boldsymbol{k}_f) \simeq 2\hbar\boldsymbol{k}_i$. The interest of such a Raman process is to couple two states a and a', energetically very close to each other, by transferring to the atom in one of the two states a very large momentum $2\hbar\boldsymbol{k}_i$, equal to twice that of an optical photon. On the other hand, if the two states a and a' were to be coupled directly by absorption of a single photon in the microwave domain, the momentum transfer would be much smaller. This possibility of coupling two sublevels of the ground state (hence having long lifetimes) by transferring a large momentum to the atom in one of these two states, has interesting applications, in particular in atomic interferometry.

To remain concise, in this chapter we have not treated a certain number of interesting related problems. Among them are multophotonic processes, photoionization, the dressed atom method that facilites the study of light shifts, or the use of photon wave packets. All these subjects are treated in the complements of this chapter.

A_{XX} : A MULTIPHOTON PROCESS: TWO-PHOTON ABSORPTION

In a multiphoton absorption process, an atom simultaneously absorbs several photons. This complement focuses on the simplest case where two photons are absorbed, while presenting general ideas that also apply to processes involving a larger number of photons. Monochromatic and broad band excitations are successively considered. The very short time the system spends in the relay state violating energy conservation is proportional to the inverse of that energy mismatch.

B_{XX} : PHOTOIONIZATION

In a photoionization process, a photon can remove an electron from an atom, which then becomes an ion (photoelectric effect). This complement studies this process by using a quantum theory of radiation that no longer couples two discrete atomic states but rather a discrete (ground) state to a continuum of (excited) states. Two important cases are considered: a quasi-monochromatic incident radiation, and a broad band excitation. In that second case, this study provides a justification for Einstein's equation of the photoelectric effect. Lastly, we consider the case where the radiation field is so very intense that the atomic ionization no longer occurs through the absorption of one or several photons, but rather by the tunnel effect.

C_{XX} : TWO-LEVEL ATOM IN A MONOCHROMATIC FIELD. DRESSED ATOM APPROACH

The dressed atom approach is a powerful tool for describing and interpreting higher order effects that appear as a two-level atom interacts with a quasi-resonant radiation. It is valid both in the weak coupling domain (low field intensity) and the strong coupling domain (very high field intensity). An essential parameter is the ratio of the Rabi oscillation (characterizing the coupling with the field) and the natural width of the atomic levels. This general approach allows, in particular, a full understanding of the various properties of light shifts.

D_{XX} : LIGHT SHIFTS: A TOOL FOR MANIPULATING ATOMS AND FIELDS

Using light shifts has become a basic tool in atomic physics, as it allows manipulating atoms and photons. A number of applications of such methods are briefly described in this complement: laser trapping of atoms by dipole forces, mirrors for atoms, optical lattices, "Sisyphus" cooling, and one by one detection of photons in a cavity.

E_{xx} : DETECTION OF ONE- OR TWO-PHOTON WAVE PACKETS, INTERFERENCE

Just as for a massive particle, on can build wave packets for a photon by the coherent superposition of states, each having a different momentum; this leads to a description of radiation propagation in free space, and we have the possibility to model the arrival of a photon on an atom. We obtain a description of the photon absorption or scattering processes that is more realistic than that given in Chapter XX, where the incident radiation is described by a Fock state having a well defined number of photons (and hence without any spatial propagation). We introduce for the photon a function that is not its wave function, but rather yields the probability amplitude that it might be detected at a given point. Absorption and scattering of wave packets are studied, as well as the one- or two-photon detection signals; the case of two entangled photons (parametric down-conversion) is treated at the end of the complement.

Complement A_{XX}

A multiphoton process: two-photon absorption

1 Monochromatic radiation 2097
2 Non-monochromatic radiation 2101
2-a Probability amplitude, probability 2101
2-b Probability per unit time when the radiation is in a Fock state 2103
3 Discussion 2105
3-a Conservation laws 2105
3-b Case where the relay state becomes resonant for one-photon absorption 2106

In the process studied in this complement the atom absorbs, not one, but two photons of energy $\hbar\omega_i$, to go from a discrete level a to another discrete higher energy level b[1]; this process is schematized in Figure 1. For the moment we shall ignore the external degrees of freedom and suppose the atom to be infinitely heavy[2]. Conservation of total energy then requires:

$$E_b - E_a = \hbar\omega_0 = 2\hbar\omega_i \tag{1}$$

The calculation of the transition amplitude will explicit the role played by this conservation law.

1. Monochromatic radiation

Studying this process involves the computation of the transition amplitude:

$$\langle\psi_{\text{fin}}|U(t_f,t_i)|\psi_{\text{in}}\rangle \tag{2}$$

The initial state at time t_i of the system atom + radiation is:

$$|\psi_{\text{in}}\rangle = |a;n_i\rangle \qquad \text{with energy} \qquad E_{\text{in}} = E_a + n_i\hbar\omega_i \tag{3}$$

It describes an atom in state a in the presence of n_i photons in mode i, with wave vector $\boldsymbol{k}_i$, polarization $\boldsymbol{\varepsilon}_i$ and frequency ω_i (we assume the radiation is monochromatic). After the absorption process, at time t_f, the photon number is lowered by two units, going from n_i to $n_i - 2$. The final state of the system is:

$$|\psi_{\text{fin}}\rangle = |b;n_i-2\rangle \qquad \text{with energy} \qquad E_{\text{fin}} = E_b + (n_i-2)\hbar\omega_i \tag{4}$$

The atom-radiation interaction is described, as before, by the electric dipole Hamiltonian (A-30) of Chapter XX, which lowers the photon number by a single unit. This

[1] In Complement B_{XX}, we shall see how multiphoton processes can also make an atom go from a discrete state to a final state belonging to a continuum of states (photoionization).

[2] When the atom's mass is finite, there are interesting physical effects arising from the conservation of total momentum; these will be studied later on (cf. § 3; see also § 2 of Complement A_{XIX}).

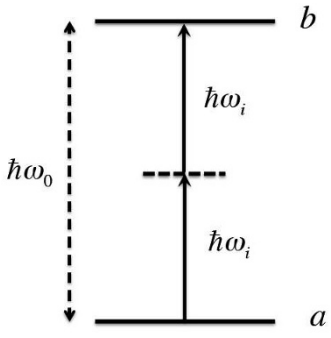


Figure 1: During a two-photon transition, the atom goes from state a to state b by absorbing two photons of energy $\hbar\omega_i$. The horizontal dashed line represents the energy half-way between the a and b levels. A third level, the relay level c, is also involved in the transition; its energy is not necessarily between those of levels a and b and it is not shown in the figure. However, we assume that the energy of that relay atomic level c is so far from the dashed line that no one-photon resonant transition can occur between a and c.

means that in the expansion (A-21) of Chapter XX for $\bar{U}(t_f, t_i)$, the lowest term that gives a non-zero contribution to the transition amplitude (2) is of order two, hence containing two operators H_I. The first brings the system atom + radiation from the initial state $|\psi_{\rm in}\rangle = |a, n_i\rangle$ to an intermediate "relay state"

$$|\psi_{\rm rel}\rangle = |c, n_i - 1\rangle \qquad \text{with energy} \qquad E_{\rm rel} = E_c + (n_i - 1)\hbar\omega_i \tag{5}$$

where c is any atomic state; the second operator H_I brings the system from this relay state $|\psi_{\rm rel}\rangle$ to the final state $|\psi_{\rm fin}\rangle = |b, n_i - 2\rangle$. One must of course sum over all accessible intermediate states c. Nevertheless, to keep the computation simple, we shall only take into account a single intermediate state (the summation of the probability amplitudes over several such states does not pose a serious problem). If we insert relation (A-21) of Chapter XX between the bra $\langle\psi_{\rm fin}|$ and the ket $|\psi_{\rm in}\rangle$, the first two terms on the right-hand side yield zero, and the third one becomes[3]:

$$\langle\psi_{\rm fin}|\bar{U}(\Delta t, 0)|\psi_{\rm in}\rangle = \left(\frac{1}{i\hbar}\right)^2 \sum_{\psi_{\rm rel}} \langle\psi_{\rm fin}|H_I|\psi_{\rm rel}\rangle\langle\psi_{\rm rel}|H_I|\psi_{\rm in}\rangle \times$$
$$\times \int_0^{\Delta t} {\rm d}t' \int_0^{t'} {\rm d}t''\ e^{iE_{\rm fin}t'/\hbar}\ e^{-iE_{\rm rel}(t'-t'')/\hbar}\ e^{-iE_{\rm in}t''/\hbar} \tag{6}$$

Let us write explicitly the argument of the exponents appearing in the integral on

[3]We assumed that the two absorbed photons were identical. If the two absorbed photons 1 and 2 were different, either by their energies, their wave vector directions, or their polarizations (albeit satisfying the conservation of energy $\hbar\omega_1 + \hbar\omega_2 = E_b - E_a = \hbar\omega_0$), this would lead to a situation similar to that of § E-1-a of Chapter XX. Two types of processes should then be considered, those where photon 1 is first absorbed, photon 2 next, and those where the photons are absorbed in the inverse order (*cf.* Figure 2).

the second line:

$$[E_b + (n_i - 2)\hbar\omega_i]\, t' - [E_c + (n_i - 1)\hbar\omega_i]\, (t' - t'') - [E_a + n_i\hbar\omega_i]\, t'' \tag{7}$$

The terms in $n_i t'$ and $n_i t''$ cancel out, leaving the expression:

$$[E_b - E_a - 2\hbar\omega_i]\, t' - [E_a - E_c + \hbar\omega_i]\, (t'' - t') \tag{8}$$

Choosing t' and $\tau = t'' - t'$ as the integral variables, the second line of (6) becomes the double integral:

$$\int_0^{\Delta t} \mathrm{d}t'\ e^{i[E_b - E_a - 2\hbar\omega_i]t'/\hbar} \int_{-t'}^{0} \mathrm{d}\tau\ e^{i[E_c - E_a - \hbar\omega_i]\tau/\hbar} \tag{9}$$

or else:

$$\int_0^{\Delta t} \mathrm{d}t' e^{i[E_b - E_a - 2\hbar\omega_i]t'/\hbar} \frac{i\hbar}{[E_a + \hbar\omega_i - E_c]} \left[1 - e^{i[E_a + \hbar\omega_i - E_c]t'/\hbar}\right] \tag{10}$$

We are dealing with situations where the frequency $\omega_i/2\pi$ of the photons is close to the two-photon resonance expressed by the conservation of energy (see relation (1)). In addition, we assume that the process is a real direct two-photon transition, and not successive one photon absorptions. In other words, we assume level a cannot absorb a photon in a resonant way and get to the intermediate level c; this means that the energy E_c of that intermediate level c must be very different from the half-way energy $(E_a + E_b)/2$ shown as the horizontal dashed line in Figure 1. It is easy to see that the first term 1 in the bracket of (10) does correspond to a two-photon resonant absorption, since its probability amplitude is written:

$$\frac{i\hbar}{[E_a + \hbar\omega_i - E_c]} \int_0^{\Delta t} \mathrm{d}t' e^{i[E_b - 2\hbar\omega_i - E_a]t'/\hbar} \sim \frac{2i\pi\hbar^2}{[E_a + \hbar\omega_i - E_c]} \delta^{(\Delta t)}\,(E_b - E_a - 2\hbar\omega_i) \tag{11}$$

(the sign $\sim$ simply means we have ignored an irrelevant phase factor), with:

$$\delta^{(\Delta t)}\,(E) = \frac{1}{\pi} \frac{\sin(E\ \Delta t/2\hbar)}{E} \tag{12}$$

The function $\delta^{(\Delta t)}$ tends towards a delta function when $\Delta t \to \infty$, as shown by relation (10) in Appendix II; it expresses an energy conservation satisfying condition (1), within $\hbar/\Delta t$. On the other hand, the second term in the bracket of (10) introduces, in the sum over t', an exponential $e^{i[E_b - E_c - \hbar\omega_i]t'/\hbar}$ that oscillates rapidly as a function of t' when condition (1) is (exactly or approximately) satisfied[4]. This term yields a non-resonant contribution, hence negligible. Its physical significance (sudden branching of the coupling between atom and field) will be discussed in comment (ii) below. We shall ignore it for the moment because of its non-resonant character. This leads us to:

$$\langle \psi_{\mathrm{fin}} | \bar{U}(\Delta t, 0) | \psi_{\mathrm{in}} \rangle = -2i\pi \frac{\langle \psi_{\mathrm{fin}} | H_I | \psi_{\mathrm{rel}} \rangle \langle \psi_{\mathrm{rel}} | H_I | \psi_{\mathrm{in}} \rangle}{E_a + \hbar\omega_i - E_c}\ \delta^{(\Delta t)}\,(E_b - E_a - 2\hbar\omega_i) \tag{13}$$

[4] In that case, $e^{i[E_b - E_c - \hbar\omega_i]t'/\hbar} \simeq e^{i[E_b + E_a - 2E_c]t'/2\hbar}$, whose exponent is necessarily large since we assumed that E_c is very different from the half-sum of E_a and E_b.

Comparison with relation (B-10) of Chapter XX shows a great similarity between the probability amplitude of a one-photon absorption process and that of a two-photon transition. We go from the first to the second by substituting the variable in the function $\delta^{(\Delta t)}$ by the one relevant for the two-photon energy conservation written in (1), and by replacing the matrix element $\langle\psi_{\text{fin}}|H_I|\psi_{\text{in}}\rangle$ by[5]:

$$\frac{\langle\psi_{\text{fin}}|H_I|\psi_{\text{rel}}\rangle\langle\psi_{\text{rel}}|H_I|\psi_{\text{in}}\rangle}{E_{\text{in}}-E_{\text{rel}}} = -\frac{\hbar\omega_i}{2\varepsilon_0 L^3}\sqrt{n_i(n_i-1)}\ \ \frac{\langle b|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|c\rangle\langle c|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|a\rangle}{E_a+\hbar\omega_i-E_c} \tag{14}$$

This means that we just have to replace, in relation (B-6) of Chapter XX for the transition amplitude, the matrix element H_I by a product of matrix elements divided by an energy difference.

Comments

(i): Characteristic time of the intermediate transition

The transit of the physical system through the intermediate relay state occurs without energy conservation, since it involves a difference $\delta E = E_a + \hbar\omega_i - E_c$ with the initial energy. Mathematically, this results in the presence, in the second time integral of (9), of an oscillating term; the larger the energy difference, the more rapid the oscillation. Once that integral is performed, we obtain the bracket appearing in (10), multiplied by a prefactor. This bracket starts from zero at time $t' = 0$, then oscillates as a function of the intermediate time t'. After a time $t' = h/\delta E$, which corresponds to one oscillation period, its average value over time equals one, precisely the value we have used for the computation of the probability amplitude.

The transit through the relay state brings in a characteristic time $\delta\tau = h/\delta E$, after which the modulus of the integral over time τ no longer increases. The larger the departure from energy conservation, the shorter that time is (this short transit through such a relay state is sometimes referred to as a "virtual transition"). The integral over time τ thus behaves completely differently from the integral over t', which, at resonance, increases linearly with time as shown from (11) and (12); this latter integral over t' may accumulate contributions over much larger times. The limitation of the times τ that actually contribute to the probability amplitude has a natural interpretation in the context of the Heisenberg time-energy uncertainly relation $\delta E \times \delta\tau \geq \hbar$.

(ii): Physical meaning of the term left out of the transition amplitude

We have left out the second term in the bracket of relation (10). Its origin is nevertheless interesting, as it arises from the sudden branching of the coupling between atom and field at time $t = 0$, as assumed in the computation. To see this, we can use a model where the interaction Hamiltonian H_I is replaced by an operator $f(t)H_I$; the time dependence of the function $f(t)$ allows introducing an adiabatic turning on of the coupling. It can be shown that the term we had ignored does disappear when turning on the interaction very slowly.

A more rigorous description can be obtained by describing the field as a wave packet propagating in space (Complement E_{XX}), and overlapping the atom only during a limited time. In that case, the interaction Hamiltonian only acts during that overlap time, even

[5] We use for H_I expression (A-24) of Chapter XX, as well as expression (A-27) for the electric field.

though the operator itself is time independent. For a wave packet with a progressive wave front, this approach shows that the term we have ignored does not even come into play.

As for the transition probability, the computation is the same as for a one-photon transition. We get for the transition probability $P^{(2)}_{a\to b}(t)$:

$$P^{(2)}_{a\to b}(t) = \left(\frac{\hbar\omega_i}{2\hbar\varepsilon_0 L^3}\right)^2 [n_i(n_i-1)] \left|\frac{\langle b|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|c\rangle\langle c|\boldsymbol{\varepsilon}_i\cdot\boldsymbol{D}|a\rangle}{E_a+\hbar\omega_i-E_c}\right|^2 \frac{4\sin^2(\omega_0-2\omega_i)\,\Delta t/2}{(\omega_0-2\omega_i)^2} \tag{15}$$

At short times, it is proportional to the square of $(\Delta t)^2$.

Finally, if several relay states are involved in the two-photon transition, one must sum expressions (13) and (14) over all the possible intermediate states $|\psi_{\rm rel}\rangle$ (taking into account the fact they each have different energies $E_{\rm rel}$); on the right-hand side of (14), this amounts to summing over all accessible intermediate atomic states c with energy E_c. Interference effects between amplitudes associated with different intermediate states c may then appear in the probability (15).

2. Non-monochromatic radiation

Consider now what happens when the initial state of the system $|\psi_{\rm in}\rangle$ contains photons of different frequencies. We are going to show that, just as for one-photon transitions, the two-photon transitions can lead to a transition probability per unit time; however, this probability involves a higher order correlation function (order 4 instead of 2).

2-a. Probability amplitude, probability

The computation of the probability amplitude is similar to that discussed in § 1; it is based, as before, on the expression of $\bar{U}$ to second order in H_I. We will carry out the calculation so as to highlight the properties of the time correlation functions of the incident electric field. The radiation initial state is $|\,\phi^R_{\rm in}\rangle$, its final state $|\,\phi^R_{\rm fin}\rangle$, and $|\,\phi^R_{\rm rel}\rangle$ is its intermediate state when the atom is in the relay state c. The two-photon transition is described by the sequence of the following states for the system atom + photon:

$$|\,a;\phi^R_{\rm in}\rangle \Rightarrow |\,c;\phi^R_{\rm rel}\rangle \Rightarrow |\,b;\phi^R_{\rm fin}\rangle \tag{16}$$

corresponding to the transition amplitude to the lowest order:

$$\langle b;\phi^R_{\rm fin}|\,\bar{U}(\Delta t,0)|\,a;\phi^R_{\rm in}\rangle = \left(\frac{1}{i\hbar}\right)^2 \int_0^{\Delta t} {\rm d}t'\langle b|\bar{\boldsymbol{D}}(t')|c\rangle\cdot\langle\phi^R_{\rm fin}|\,\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')|\,\phi^R_{\rm rel}\rangle \times \int_0^{t'} {\rm d}t''\langle c|\bar{\boldsymbol{D}}(t'')|a\rangle\cdot\langle\phi^R_{\rm rel}|\,\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t'')|\,\phi^R_{\rm in}\rangle \tag{17}$$

($\boldsymbol{R}$ is the atom's position). By analogy with relation (B-13) of Chapter XX, we set:

$$\begin{aligned}\langle c|\bar{\boldsymbol{D}}(t)|a\rangle &= \boldsymbol{\mathcal{D}}_{ca}\,e^{i(E_c-E_a)t/\hbar} \qquad \text{with:} \quad \boldsymbol{\mathcal{D}}_{ca}=\mathcal{D}_{ca}\,\mathbf{e}_{ca}\\ \langle b|\bar{\boldsymbol{D}}(t)|c\rangle &= \boldsymbol{\mathcal{D}}_{bc}\,e^{i(E_b-E_c)t/\hbar} \qquad \text{with:} \quad \boldsymbol{\mathcal{D}}_{bc}=\mathcal{D}_{bc}\,\mathbf{e}_{bc}\end{aligned} \tag{18}$$

We then get:

$$\langle b;\phi^R_{\rm fin}\,|\,\bar U(\Delta t,0)|\,a;\phi^R_{\rm in}\rangle = \\ -\frac{\mathcal{D}_{bc}\mathcal{D}_{ca}}{\hbar^2}\int_0^{\Delta t}{\rm d}t' e^{i(E_b-E_c)t'/\hbar}\langle\phi^R_{\rm fin}\,|\,\mathbf{e}_{bc}\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')|\,\phi^R_{\rm rel}\rangle \\ \times\int_0^{t'}{\rm d}t'' e^{i(E_c-E_a)t''/\hbar}\langle\phi^R_{\rm rel}\,|\mathbf{e}_{ca}\cdot\,\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t'')|\,\phi^R_{\rm in}\rangle \tag{19}$$

Expression (A-29) of Chapter XX for the electric field operator $\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t)$ is a sum of modes' contributions, each including the exponential $e^{-i\omega_i t}$ associated with its eigenfrequency. Let us focus on the contribution of mode i in $\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')$ and mode j in $\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t'')$. It involves the time integrals:

$$\int_0^{\Delta t}{\rm d}t' e^{i(E_b-E_c)t'/\hbar}e^{-i\omega_i t'}\int_0^{t'}{\rm d}t'' e^{i(E_c-E_a)t''/\hbar}e^{-i\omega_j t''} \tag{20}$$

with an exponent containing:

$$[E_b-E_c-\hbar\omega_i]\,t'+[E_c-E_a-\hbar\omega_j]\,t'' \\ =[E_b-E_a-\hbar\omega_i-\hbar\omega_j]\,t'+[E_c-E_a-\hbar\omega_j]\,(t''-t') \tag{21}$$

reminiscent of result (8) obtained for monochromatic excitation. The computation is then very similar to that of § 1, assuming that the relay state is not half-way between levels a and b, and that the frequency distribution of the incident photons does not include any of the resonance frequencies for the one-photon transitions $a\leftrightarrow c$ and $c\leftrightarrow b$. We make the usual change of variable $\tau=t''-t'$, and, in the integral over τ, we only keep the upper boundary contribution, as we did going from (10) to (13):

$$\int_{-t'}^{0}{\rm d}\tau\; e^{i[E_c-E_a-\hbar\omega_i]\tau/\hbar}\Rightarrow\frac{i\hbar}{[E_a+\hbar\omega_j-E_c]} \tag{22}$$

(we discussed in comment (ii) at the end of § 1 the significance of the lower boundary contribution, and why it is justified to ignore it). In addition, we assume the width of the frequency spectrum of the incident photons to be small compared to the one-photon resonance detuning $E_a+\hbar\omega_i-E_c$; consequently, the denominator of (22) does not vary significantly in relative value, and can be replaced by the constant value $E_a+\hbar\omega_{\rm ex}-E_c$, where $\omega_{\rm ex}/2\pi$ is the central excitation frequency. We note Δ_{ac} the distance from the resonant absorption of a photon in the intermediate state:

$$\Delta_{ac}=\frac{E_a+\hbar\omega_{\rm ex}-E_c}{\hbar} \tag{23}$$

The replacement of the integral over τ by i/Δ_{ac} yields an exponential depending on the variable t', with argument $i\,[E_b-E_a-\hbar\omega_j-\hbar\omega_i]\,t'/\hbar$. Each summation over the modes i and j with the exponential factors $e^{-i\omega_{i,j}t'}$ reconstructs the electric field $\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')$, which leads to:

$$\langle b;\phi^R_{\rm fin}\,|\,\bar U(\Delta t,0)|\,a;\phi^R_{\rm in}\rangle=-\frac{i\mathcal{D}_{bc}\mathcal{D}_{ca}}{\hbar^2\Delta_{ac}} \\ \times\int_0^{\Delta t}{\rm d}t' e^{i\omega_0 t'/\hbar}\langle\phi^R_{\rm fin}\,|\,\mathbf{e}_{bc}\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')|\,\phi^R_{\rm rel}\rangle\langle\phi^R_{\rm rel}\,|\mathbf{e}_{ca}\cdot\,\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')|\,\phi^R_{\rm in}\rangle \tag{24}$$

where ω_0 was defined as $\omega_0 = (E_b - E_a)/\hbar$. Finally, the summation over all the radiation intermediate states $|\,\phi^R_{\rm rel}\rangle$ yields a closure relation and we obtain:

$$\sum_{|\,\phi^R_{\rm rel}\rangle} \langle b;\phi^R_{\rm fin}\,|\,\bar{U}(\Delta t,0)|\,a;\phi^R_{\rm in}\rangle = -\frac{i\mathcal{D}_{bc}\mathcal{D}_{ca}}{\hbar^2\Delta_{ac}} \times \int_0^{\Delta t} {\rm d}t' e^{i\omega_0 t'/\hbar}\langle\phi^R_{\rm fin}\,|\,\left[\mathbf{e}_{bc}\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')\right]\left[\mathbf{e}_{ca}\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')\right]|\,\phi^R_{\rm in}\rangle \tag{25}$$

This result is similar to the probability amplitude for a one-photon transition, written on the second line of (B-14) in Chapter XX, provided we make the substitution:

$$\begin{aligned} &-\mathcal{D}\langle\phi^R_{\rm fin}\,|\,\mathbf{e}_d\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')|\,\phi^R_{\rm in}\rangle \\ &\Rightarrow \frac{\mathcal{D}_{bc}\mathcal{D}_{ca}}{\hbar\Delta_{ac}}\langle\phi^R_{\rm fin}\,|\,\left[\mathbf{e}_{bc}\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')\right]\left[\mathbf{e}_{ca}\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t')\right]|\,\phi^R_{\rm in}\rangle \end{aligned} \tag{26}$$

We then follow the same line of reasoning as for a one-photon transition. In equality (B-17) of Chapter XX, we must now substitute:

$$\begin{aligned} G^N_R(t',t'') \Rightarrow \langle\phi^R_{\rm in}|\left[\mathbf{e}^*_{ca}\cdot\bar{\boldsymbol{E}}^{(-)}_\perp(\boldsymbol{R},t')\right]\left[\mathbf{e}^*_{bc}\cdot\bar{\boldsymbol{E}}^{(-)}_\perp(\boldsymbol{R},t')\right] \\ \times\left[\mathbf{e}_{bc}\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t'')\right]\left[\mathbf{e}_{ca}\cdot\bar{\boldsymbol{E}}^{(+)}_\perp(\boldsymbol{R},t'')\right]|\,\phi^R_{\rm in}\rangle \end{aligned} \tag{27}$$

which is then inserted in (B-18) to yield the transition probability. This probability is given by the Fourier transform, at the angular frequency ω_0 of the atomic transition, of a correlation function of the field in the initial state, and which involves four field operators (4-point correlation function). This function is in general different from a product of correlation functions involving two field operators (those determining the absorption probability of a single photon). This means that measurements of two-photon transition probabilities yield access to characteristics of the quantum field that are different from those measured in single photon transitions.

2-b. Probability per unit time when the radiation is in a Fock state

Let us assume now that the radiation is initially in a Fock state such as that described by (B-19) in Chapter XX. In (25), we replace the positive frequency components of the electric fields by their expressions (A-27) of Chapter XX. Only the occupied modes in state $|\,\phi^R_{\rm in}\rangle$ now come into play, since each annihilation operator yields a factor equal to the square root of the mode's initial population; the other modes give a zero result. We then consider two modes, k_1 and k_2, initially occupied in the incident radiation. They yield two contributions (Figure 2) to relation (25): in one of them (term $i = k_1$ and $j = k_2$), the photon k_1 is absorbed first and brings the atom from the ground state to the relay state, then photon k_2 completes the two-photon transition and brings the atom to level b; in the other (term $i = k_2$ and $j = k_1$), the order of the two absorptions is inverted. These two contributions interfere in the probability: once the amplitude modulus is squared, four terms arise from the cross contributions of the two modes (to which we must add two non-crossed contributions $i = j = k_{1,2}$ where only one mode is involved).

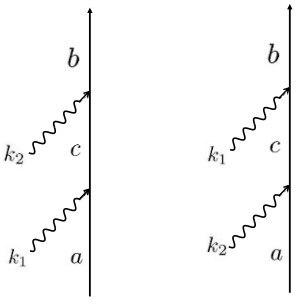


Figure 2: Two diagrams schematizing a two-photon transition with a multimode source where two modes k_1 and k_2 are initially occupied. In the left-hand side diagram, photon k_1 is absorbed first, bringing (in a non-resonant fashion) the atom from the initial sate a to the relay state c; photon k_2 then completes the (resonant) two-photon transition. In the right-hand side diagram, the order of absorption of photons k_1 and k_2 is inverted. These two diagrams describe probability amplitudes that interfere when computing the two-photon transition probability.

The same line of reasoning as in § B-2-a of Chapter XX, and summarized in Figure 1 of that chapter, can be followed here. We assume that the 4-point correlation of the field goes rapidly to zero when $t'' - t'$ is larger that a value $1/\Delta\omega$ that is small compared to Δt. One can then show that the transition probability becomes proportional to Δt, and that the two-photon transition probability per unit time can be written:

$$\begin{aligned} W^{(2)}_{a\to b} &= \frac{P^{(2)}_{a\to b}(\Delta t)}{\Delta t} \\ &= \frac{2\pi}{\hbar^2}\left(\frac{\mathcal{D}_{bc}\mathcal{D}_{ca}}{\hbar\Delta_{ac}}\right)^2 \frac{1}{4\varepsilon_0^2 L^6}\sum_{i,j}\Big[2\,\langle n_i n_j\rangle - \delta_{ij}\langle n_i(n_i+1)\rangle\Big]\,(\hbar\omega_i)\,(\hbar\omega_j) \\ &\qquad\qquad \times\ \left|(\mathbf{e}_{bc}\cdot\boldsymbol{\varepsilon}_i)\,(\mathbf{e}_{ca}\cdot\boldsymbol{\varepsilon}_j)\right|^2\ \delta(\omega_i+\omega_j-\omega_0) \end{aligned} \tag{28}$$

The delta function at then end of this expression obviously expresses total energy conservation: for the atomic transition to occur, the sum of the energies of the absorbed photons must be equal to the energy of the transition. As expected, the probability includes the photon populations that satisfy this condition. A general property is that, for $i \neq j$ (photons absorbed from two different modes), it is the average value $\langle n_i n_j\rangle$ of the product of the mode populations that appears in the two-photon transition probability, and not the product of the average values which different in general (they are nevertheless equal in the special case of a Fock state of the radiation). For $i = j$ (two photons absorbed from the same mode, as in § 1), it is the average value $\langle n_i\,(n_i-1)\rangle$ that comes into play; this value equals zero if only one photon is present in the mode, as obviously one single photon cannot induce a two-photon transition.

Comment: There also exists 3- ,..., n-photon transitions, corresponding to an energy conservation relation $E_b - E_a = n\hbar\omega_i$. The corresponding transition rates are proportional to field correlation functions of order 6, .., $2n$.

3. Discussion

Even though the transition amplitudes for one- and two-photon absorption processes are similar, the second type of processes has a number of specific features we now discuss.

3-a. Conservation laws

α. Total energy conservation

As we just saw, the function $\delta^{(\Delta t)}(E_b - E_a - 2\hbar\omega_i)$ appearing in (13) expresses the conservation of total energy. When the atom absorbs two photons to go from state a to state b, its gain in energy $E_b - E_a$ equals the sum $2\hbar\omega_i$ of the energies of the two absorbed photons.

β. Total momentum conservation

In the computation leading to the two-photon transition amplitude, the external variables have been ignored. To take them into account, we must assume the atomic center of mass is at point $\boldsymbol{R}$ and keep the exponentials $\exp(i\boldsymbol{k}_i \cdot \boldsymbol{R})$ appearing in the operators $\boldsymbol{E}^{(+)}(\boldsymbol{R})$ in the two interaction Hamiltonians H_I; as these two exponentials are multiplied by each other, they yield the operator $\exp(2i\boldsymbol{k}_i \cdot \boldsymbol{R})$. We must also include in the initial and final states the quantum numbers $\boldsymbol{K}_{\text{in}}$ and $\boldsymbol{K}_{\text{fin}}$ characterizing the initial $\hbar\boldsymbol{K}_{\text{in}}$ and final $\hbar\boldsymbol{K}_{\text{fin}}$ momenta of the center of mass. We then get in the transition amplitude an additional term:

$$\langle \boldsymbol{K}_{\text{fin}} | \exp(2i\,\boldsymbol{k}_i \cdot \boldsymbol{R}) | \boldsymbol{K}_{\text{in}} \rangle = \delta(\boldsymbol{K}_{\text{fin}} - \boldsymbol{K}_{\text{in}} - 2\boldsymbol{k}_i) \tag{29}$$

which shows that the atom's momentum increases by $2\hbar\boldsymbol{k}_i$ when it absorbs two photons.

Comment:

Imagine the atom is excited by two light beams 1 and 2, having the same frequency $\omega_i/2\pi$ but propagating in opposite directions. The previous computations must then be generalized to the case where the two photons absorbed by the atom belong, one to beam 1, and the other to beam 2. The momenta $+\hbar\boldsymbol{k}_i$ and $-\hbar\boldsymbol{k}_i$ of these two photons are then opposite and the total momentum gained by the atom during the transition is zero. As the Doppler effect and the recoil effect are linked to the variation of the atomic momentum during the transition (see Complement A_{XIX}), it follows that the two-photon absorption line does not present any Doppler broadening nor any recoil shift. Such a situation presents many advantages for high resolution spectroscopy, and is used for example in the study of the two-photon transition between the $1s$ and $2s$ states of the hydrogen atom.

γ. Conservation of total angular momentum and parity

Expression (14) appearing in the two-photon absorption amplitude is a product of two matrix elements of a component of the atomic electric dipole – which is a vector operator – and an energy denominator. In a rotation of the atom, this expression will be transformed as the product of two vectors, since the energy denominator is rotation invariant. Vectors are irreducible tensor operators (Complement G_X, exercise 8) of order $K = 1$. Consequently, expression (14) may be expanded[6] as a sum of components with total angular momentum $F = 0, 1, 2$. Using the Wigner-Eckart theorem (Complement E_X), we can show that the two-photon absorption amplitude between two levels with quantum numbers F, M and F', M' (where M and M' are the components of $\boldsymbol{F}$ on the Oz axis) is different from zero only if:

$$\begin{aligned} F - F' &= \pm 2, \pm 1, 0 \\ M - M' &= \pm 2, \pm 1, 0 \end{aligned} \tag{30}$$

In addition, the electric dipole operator appearing in (14) is an odd operator, as it is proportional to the electron position operator. Consequently, the initial and final states of the two-photon transition must have the same parity, and a parity inverse to that of the relay state.

These selection rules can be applied to the $1s \leftrightarrow 2s$ transition of the hydrogen atom, that occurs between two states having the same parity and a total angular momentum difference equal to ± 1 at most (the $1/2$ spins of the electron and the proton are taken into account). For electric dipole transitions, a two-photon transition $1s \leftrightarrow 2s$ is allowed, whereas it is forbidden for a one-photon transition.

3-b. Case where the relay state becomes resonant for one-photon absorption

In the denominator of expression (14), we have the quantity:

$$\hbar\Delta_{ac} = E_a + \hbar\omega_i - E_c \tag{31}$$

which is the difference between the energy of the atomic state a, increased by $\hbar\omega_i$, and the energy of the relay sate c. If Δ_{ac} goes to zero, we get a divergence and the expressions we have obtained become meaningless. In the computation, we did explicitly assume that the intermediate level was not resonant for a one-photon absorption, so that this divergence should not occur. Let us examine, however, what would be involved if Δ_{ac} were to go to zero. As the resonance condition for the two-photon transition is written $E_b = E_a + 2\hbar\omega_i$, the condition $\Delta_{ac} = 0$ means that the atomic relay level[7] c is exactly

[6] The product of two irreducible tensor operators K and K' can be decomposed as the product of two kets with angular momentum K and K', hence involving Clebsch-Gordan coefficients. This means that, according to the general results of Chapter X on the addition of angular momenta, a product of irreducible tensor operators can be decomposed as the sum of other irreducible tensor operators of order K'', where K'' varies between $|K - K'|$ and $K + K'$. In the particular case where $K = K' = 1$, we get three possible values $F = 0, 1, 2$.

[7] We assume here that the relay state c is discrete. If it belongs to a continuum, the sum over this relay state c in (14) becomes an integral over E_c. An adiabatic branching calculation then introduces a fraction $1/(E_a + \hbar\omega_i + i\eta - E_c)$ with $\eta \to 0$, which can then be expressed in terms of $\delta(E_a + \hbar\omega_i - E_c)$ and $\mathcal{P}(1/(E_a + \hbar\omega_i - E_c))$, where $\mathcal{P}$ is the Cauchy principal value. This calculation yields, after integration over E_c, functions of $E_{\text{in}} = E_a + 2\hbar\omega_i$ which have no reason to diverge in the vicinity of $E_{\text{in}} \simeq E_{\text{fin}}$.

half-way in energy between a and b, the initial and final atomic levels. This means that, starting from level a, a resonance would occur for both a two-photon and a one-photon process.

Is it possible to study the case where Δ_{ac} is zero or very small, while avoiding the divergence of the two-photon absorption? A method to overcome this difficulty is to note that state c has a finite lifetime τ_c, because of spontaneous emission of photons from that state. As we did before in § E-2 of Chapter XX, when studying the resonant scattering of a photon by an atom and where similar divergences appeared, we can show that it is legitimate to replace the energy E_c of state c by $E_c - i\hbar(\Gamma_c/2)$, where $\Gamma_c = 1/\tau_c$ is the natural width of level c. The denominator of the transition amplitude no longer goes to zero and the divergence disappears. The transition amplitude still varies significantly over an interval of width Γ_c when E_c varies around $E_a + \hbar\omega_i$.

Replacing E_c by $E_c - i(\hbar\Gamma_c/2)$ leads to valid results only if the matrix elements $\langle c; n_i - 1|H_I|a; n_i\rangle$ and $\langle b; n_i - 2|H_I|c; n_i - 1\rangle$, characterizing the coupling of the field with the atom for the $c \leftrightarrow a$ and $b \leftrightarrow c$ transitions, are small compared to $\hbar\Gamma_c$. If this is not true, we cannot limit the computation to the lowest field order. We must then diagonalize the Hamiltonian of the global system atom + field within the subspace $\mathcal{E}$ of the states which, in the absence of coupling, are very close to each other[8]. When no relay state is resonant, the subspace $\mathcal{E}$ is two-dimensional; it is spanned by the two states $|a; n_i\rangle$ and $|b; n_i - 2\rangle$. When one relay state c becomes resonant, we must include the state $|c; n_i - 1\rangle$ in the subspace $\mathcal{E}$ – which then becomes three-dimensional. To study the dynamics of the system, we must diagonalize the matrix:

$$\begin{pmatrix} E_a + n_i\hbar\omega_i & \langle a; n_i|H_I|c; n_i - 1\rangle & 0 \\ \langle c; n_i - 1|H_I|a; n_i\rangle & E_c + (n_i - 1)\hbar\omega_i & \langle c; n_i - 1|H_I|b; n_i - 2\rangle \\ 0 & \langle b; n_i - 2|H_I|c; n_i - 1\rangle & E_b + (n_i - 2)\hbar\omega_i \end{pmatrix} \tag{32}$$

This general treatment allows taking into account simultaneously the one- and two-photon transitions.

Concluding this complement, let us emphasize that the two-photon transitions involve a physical process different from the mere succession of two one-photon absorptions. We stressed in the discussion of § 1, and in particular in its two comments, the difference between populating the final state, which is cumulative in time and conserves the energy, and a transit through an intermediate relay state, which can only last a very short time $\Delta\tau$, limited by the non-conservation of energy. It is also noteworthy that the two-photon transition amplitude can take a form very similar to that of a one-photon transition; the only major change is the replacement of the matrix element to first order in the interaction, by a second order matrix element, divided by an energy defect factor in the relay state. These concepts can be generalized to higher order processes: similar techniques can be used to evaluate three-, four-, etc.. photon transition amplitudes.

[8]Such a description of the atom + radiation interactions is called the "dressed atom method" (see for example Chapter VI of reference [21]). In Complement C_{XX}, this method is applied to the problem of a two-level atom interacting with a strong field. The eigenstates of the total Hamiltonian restricted to the subspace $\mathcal{E}$ are called the "dressed states".

Complement B_{XX}

Photoionization

1 Brief review of the photoelectric effect 2110
1-a Interpretation in terms of photons 2110
1-b Photoionization of an atom 2111
2 Computation of photoionization rates 2112
2-a A single atom in monochromatic radiation 2112
2-b Stationary non-monochromatic radiation 2113
2-c Non-stationary and non-monochromatic radiation 2116
2-d Correlations between photoionization rates of two detector atoms 2117
3 Is a quantum treatment of radiation necessary to describe photoionization? 2118
3-a Experiments with a single photodetector atom 2118
3-b Experiments with two photodetector atoms 2119
4 Two-photon photoionization 2123
4-a Differences with the one-photon photoionization 2123
4-b Photoionization rate 2124
4-c Importance of fluctuations in the radiation intensity 2125
5 Tunnel ionization by intense laser fields 2126

All the atomic processes of absorption, emission or scattering of photons studied in Chapter XX involved transitions between two discrete states of the atom. In addition to discrete levels, atoms also have continuums of energy levels. The most directly accessible one is the simple ionization continuum, which corresponds to the loss of a single electron by the atom (ionization). This continuum starts at an energy threshold E_I above the ground state energy, and extends over all energies larger than this threshold. This energy E_I is called the "ionization energy". The aim of this complement is to study the "photoionization" process where incident radiation takes the atom from its ground state to a state belonging to the ionization continuum.

Once the atom's electron has reached the ionization continuum, it can travel an arbitrary distance from the remaining ion; it has been ejected from the atom by the incident radiation. Such a process is reminiscent of the photoelectric effect where radiation ejects an electron from a metal. This is why we shall review in § 1 a few properties of the photoelectric effect to underline its analogies with photoionization.

We shall then use quantum theory to compute, in § 2, the probability per unit time for the incident radiation to photoionize an atom. We shall assume that the incident radiation spectrum is entirely above the ionization threshold, so that no resonant absorption can bring the atom to a discrete excited state. Since the emitted electron can be amplified in a photomultiplier, the atom can play the role of a photodetector. In the case where only one photodetector D is used (§ 2-a), the computations are very similar to those exposed in Chapter XX; the only differences arise from the continuous character

of the final atomic state. Another interesting situation occurs when two detectors D1 and D2 are placed in the radiation field at points $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ (§ 2-d) and when we focus on the correlations between their signals. For example, we shall compute the probability per unit time to observe a photoionization at $\boldsymbol{R}_1$ at time t_1 and another one at $\boldsymbol{R}_2$ at time t_2.

One may wonder if a quantized radiation theory is needed to quantitatively account for the photoionization processes. Could a semiclassical theory suffice to describe the ionization of one or several quantized atoms by a classical field? In other words, can the photoelectric effect be explained "without photons" [45]? This question will be discussed in § 3.

The atom can also be photoionized by the absorption of a number of photons n, larger than one. These processes are called "*multiphoton ionization*" and play an important role in experiments using high intensity laser sources. In § 4, we shall give an idea of how to compute the rates of those processes for $n = 2$. We shall then briefly mention in § 5 another mechanism for atoms' ionization, based on a tunnel effect, and occurring when the incident radiation electric field becomes of the order of the Coulomb field between an atom's electron and the nucleus.

1. Brief review of the photoelectric effect

In 1905, Albert Einstein [43] introduced for the first time in physics the concept of "light quanta", which we now call photons. Considering the great analogy between certain statistical properties of black body radiation and those of an ideal gas of particles, Einstein proposed the idea that radiation was in fact composed of discrete quanta, each having an energy $h\nu$. In view of the successes of the wave theory of light, this return to a particle description seemed totally unrealistic for most physicists at the time. Energy quantization had indeed been introduced a few years earlier by Max Planck to account for the spectral distribution of black body radiation, but it was the exchanges of energy between matter and radiation that were quantized, not the radiation itself.

1-a. Interpretation in terms of photons

In that same 1905 article, Einstein used the concept of light quanta to give a new description of the photoelectric effect. In this process, an electron is ejected from a metal irradiated by light. Einstein postulated that the energy $h\nu$ of a light quantum from the incident beam was absorbed by an electron in the metal, hence allowing it to escape from the metal. This escape requires an energy at least equal to that of the binding energy E_L of the electron in the metal. The frequency ν of the light beam must therefore be larger than a threshold value given by $h\nu = E_L$. If $h\nu < E_L$, no electron can be ejected. If $h\nu > E_L$, the energy surplus $h\nu - E_L$ provides the electron with a kinetic energy $E_{\text{kin}} = mv^2/2$. This interpretation leads to Einstein's equation:

$$\frac{1}{2}mv^2 = h\nu - E_L \quad \text{if} \quad h\nu > E_L \tag{1}$$

giving the kinetic energy of the ejected electron as a function of ν. It means, in particular, that the kinetic energy of each electron depends only on the frequency of the light beam,

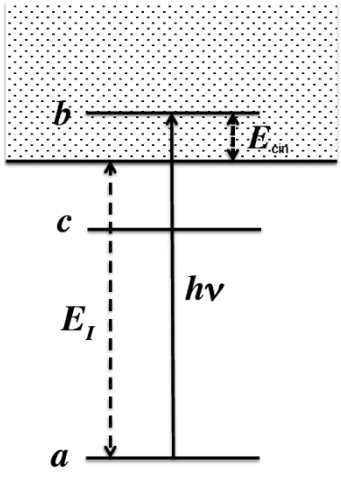


Figure 1: Photoionization of an atom. State a is the ground state, state c one of the discrete states. The continuum of states belonging to the continuous part of the spectrum (ionization continuum) starts at an energy E_I above the ground state. Energy E_I is called the "ionization energy". As the atom in state a absorbs a photon with frequency ν such that $h\nu > E_I$, it goes into a state b, which is part of the ionization continuum. The electron is ejected, and its kinetic energy $E_{\rm kin}$ (when it is far enough from the ion formed as it left the atom) is equal to the difference between $h\nu$ and E_I.

and not on its intensity[1](which, on the other hand, determines the number of ejected electrons per unit time). Equation (1) also tells us that if ν is varied and one plots the variation of $mv^2/2$ as a function of ν, one should get a straight half-line with slope h, starting from the abscissa axis at point E_L/h. All these predictions generated a certain skepticism and it was not until several years later (1913) that an experimental confirmation of the predictions of equation (1) was obtained by the work of R. Millikan and H. Fletcher on the photoelectric effect [44].

1-b. Photoionization of an atom

Figure 1 represents the ground state a of an atom, one of the states c from the discrete part of the spectrum, and the continuum part of the spectrum which starts at a distance E_I above the ground state. The origin of the energies is often chosen at the beginning of the continuum, and hence the discrete states have a negative energy. When it is in the positive energy states, belonging to the continuum and called "scattering states", the electron is no longer bound to the nucleus, although it is still attracted to it.

Consider an atom in the ground state a. A photon with energy $h\nu$ has different ways to bring energy to the atomic electron. If $h\nu < E_I$, this photon can be absorbed only if ν coincides with the frequency $\nu_{ca} = (E_c - E_a)/h$ of a transition between state a

[1]A classical theory would tend to predict that the higher the light intensity, the more energy could be furnished to the electron, thereby increasing its acceleration.

and state c belonging to the discrete spectrum. This is the absorption process between two discrete states already studied in Chapter XX. On the other hand, if $h\nu > E_I$, the atom can always absorb the photon and end up in a state within the continuum. The electron is no longer bound and can move away from the remaining ion formed once the electron has left. When the electron is far enough from the ion for their Coulomb interaction energy to be neglected, its kinetic energy is given by:

$$\frac{1}{2}mv^2 = h\nu - E_I \quad \text{if} \quad h\nu > E_I \tag{2}$$

This is the photoionization process of an atom. Equation (2) is the generalization for an atom of equation (1) introduced by Einstein for the photoelectric effect in a metal.

2. Computation of photoionization rates

Let us now see how to adapt the computations of Chapter XX to the calculation of a photoionization rate.

2-a. A single atom in monochromatic radiation

We start from expression (B-7) of Chapter XX for the probability that the system, leaving at $t = 0$ the initial state $|\psi_{\text{in}}\rangle = |\, a; n_i\rangle$ (atom in state a in the presence of n_i photons ω_i), ends up at time Δt in the final state $|\psi_{\text{fin}}\rangle = |\, b; n_i - 1\rangle$ (atom in state b with an energy $\hbar\omega_0$ above a, and one photon less in mode i) :

$$|\langle\psi_{\text{fin}}|\bar{U}(\Delta t, 0)|\psi_{\text{in}}\rangle|^2 = \frac{1}{\hbar^2}\,|\langle\psi_{\text{fin}}|H_I|\psi_{\text{in}}\rangle|^2\,\frac{4\sin^2\left[(\omega_0-\omega_i)\Delta t/2\right]}{\left[(\omega_0-\omega_i)/2\right]^2} \tag{3}$$

In Chapter XX, we used this expression for studying the case where states a and b are both discrete states. It is nevertheless still valid when b belongs to a continuum; its interpretation, however, is different. Whereas the probability of finding the atom in a discrete final state makes sense, from a physical point of viewn, when dealing with a continuum we must compute the probability of finding the atom within a non-zero energy interval. We must then sum probability (3) over states b.

As b varies, $\omega_0 = (E_b - E_a)/\hbar$ varies, and so does the matrix element of H_I. However, for large enough Δt, the variation of the H_I matrix element is much slower than that of the ratio in the right-hand side of (3). This ratio is the square of a diffraction function, whose maximum equals Δt^2 for $\omega_0 = \omega_i$ and whose width is of the order of $1/\Delta t$. The area under this function is thus of the order of $\Delta t^2 \times (1/\Delta t) = \Delta t$. Compared to functions of ω_0 with slow variations over an interval of the order of $1/\Delta t$, this function behaves, within a proportionality factor, as the product $\Delta t\,\delta(\omega_0 - \omega_i)$. It follows that the sum over b of (3) is proportional to Δt, meaning we can define a probability per unit time for the atom to reach the continuum, i.e. a photoionization rate.

The proportionality factor between a diffraction function and a delta function is given by relation (11) in Appendix II, which is written:

$$\lim_{\Delta t\to\infty}\frac{\sin^2\left[(\omega_0-\omega_i)\Delta t/2\right]}{\left[(\omega_0-\omega_i)/2\right]^2} = \pi\ \Delta t\ \delta(\frac{\omega_0-\omega_i}{2}) = 2\pi\hbar\ \Delta t\ \delta(E_b - E_a - \hbar\omega_i) \tag{4}$$

Inserting (4) into (3), summing over b, and dividing by Δt, finally yields the photoionization rate:

$$\frac{1}{\Delta t}\sum_b |\langle\psi_{\rm fin}|\bar{U}(\Delta t,0)|\psi_{\rm in}\rangle|^2 = \frac{2\pi}{\hbar}\sum_b |\langle\psi_{\rm fin}|H_I|\psi_{\rm in}\rangle|^2\delta(E_b - E_a - \hbar\omega_i)$$
$$= \frac{2\pi}{\hbar}\underbrace{|\langle\psi_{\rm fin}|H_I|\psi_{\rm in}\rangle|^2}_{E_b=E_a+\hbar\omega_i}\ \rho(E_a+\hbar\omega_i) \tag{5}$$

where $\rho(E_a+\hbar\omega_i)$ is the density of states in the continuum around the energy $E_a+\hbar\omega_i$. This expression is just a consequence of the Fermi golden rule (Chapter XIII, § C-3-b) applied to the coupling between the discrete level $|\,a;n_i\rangle$ and the continuum $|\,b;n_i-1\rangle$. It is also reminiscent of expression (C-37) of Chapter XIII which yields the transition probability per unit time between a discrete atomic state and a continuum, with an excitation induced by a classical wave described by a time-dependent sinusoidal function. This point will be discussed further in § 3.

2-b. Stationary non-monochromatic radiation

We now consider a single atom interacting with non-monochromatic radiation, described by a spectral distribution $I(\omega)$. We first assume that the field statistical properties are time-invariant. We shall consider the case of non-stationary radiation in § 2-c.

α. *Field and atomic dipole correlation functions*

In § D-1 of Chapter XX, we obtained the expression for the probability $\sum_b P^{\rm abs}_{a\to b}(\Delta t)$ for the atom to go, through the absorption of a photon, from state a to any state b different from a, after a time Δt. This probability is given by relation (D-4) of that chapter as a double integral of a sum of products of field and atomic dipole correlation functions.

We first examine the correlation functions of the atomic dipole. As in relation (B-13) of Chapter XX, we write the matrix elements of this dipole, in the Heisenberg picture (with respect to the Hamiltonian of the free atom), as:

$$\langle b|\bar{\boldsymbol{D}}(t)|a\rangle = \langle b|\boldsymbol{D}|a\rangle e^{i\omega_{ba}t} = \boldsymbol{\mathcal{D}}_{ba}e^{i\omega_{ba}t} \tag{6}$$

with:

$$\boldsymbol{\mathcal{D}}_{ba} = \mathcal{D}_{ba}\ \boldsymbol{e}_d \tag{7a}$$

where $\boldsymbol{e}_d$ is the unit vector parallel to the vector $\langle b|\boldsymbol{D}|a\rangle$. Since $\boldsymbol{e}_d^*\cdot\boldsymbol{e}_d = 1$, we have:

$$\mathcal{D}_{ba} = \boldsymbol{e}_d^*\cdot\boldsymbol{\mathcal{D}}_{ba} = \langle b|\boldsymbol{e}_d^*\cdot\boldsymbol{D}|a\rangle \tag{7b}$$

For the sake of simplicity, we shall assume the unit vector $\boldsymbol{e}_d$ to have the same direction for all the states b.

Comment:

This vector $\boldsymbol{e}_d$ is, a priori, not the same for all the states b related to a by matrix elements of $\boldsymbol{D}$. The direction of that vector actually depends on the rotation symmetry properties

of the two states[2]. One can sort the different b states by categories having the same symmetry (hence the same direction for $\boldsymbol{e}_d$) and for which the computations presented hereafter are valid. One must then add all the ionization probabilities calculated for each category.

Since the field always appears in a scalar product with $\boldsymbol{D}$, assuming that $\boldsymbol{e}_d$ has the same direction for all states b implies that only the scalar products $E_d^{(+)}$ (and $E_{d^*}^{(-)}$) of the fields with $\boldsymbol{e}_d$ (and with $\boldsymbol{e}_d^*$) appear in the correlation functions of the field.

The calculation is then very similar to that of § D-1 of Chapter XX, and we obtain:

$$\sum_b P_{a\to b}^{\text{abs}}(\Delta t) = \frac{1}{\hbar^2}\int_0^{\Delta t} \mathrm{d}t' \int_0^{\Delta t} \mathrm{d}t''\; G_A(t''-t')\; G_R^N(t''-t') \tag{8}$$

where:

$$G_A(t''-t') = \sum_b |\mathcal{D}_{ba}|^2 e^{i\omega_{ba}(t''-t')} \tag{9}$$

$$\begin{aligned} G_R^N(t''-t') &= \langle \phi_{\text{in}}^R | \left[\boldsymbol{e}_d^* \cdot \bar{\boldsymbol{E}}^{(-)}(\boldsymbol{R},t')\right] \left[\boldsymbol{e}_d \cdot \bar{\boldsymbol{E}}^{(+)}(\boldsymbol{R},t'')\right] |\phi_{\text{in}}^R\rangle \\ &= \langle \phi_{\text{in}}^R | \bar{E}_{d^*}^{(-)}(\boldsymbol{R},t') \bar{E}_d^{(+)}(\boldsymbol{R},t'') | \phi_{\text{in}}^R\rangle \end{aligned} \tag{10}$$

In this last equality, $|\phi_{\text{in}}^R\rangle$ is the radiation initial state. As this state is stationary, its properties are invariant under time translation; consequently, the correlation function depends only on the difference $t''-t'$.

The atomic correlation function (9) can also be rewritten as:

$$G_A(t''-t') = \int \mathrm{d}\omega\, \tilde{G}_A(\omega) e^{i\omega(t''-t')} \tag{11}$$

where:

$$\tilde{G}_A(\omega) = \sum_b |\mathcal{D}_{ba}|^2 \delta(\omega - \omega_{ba}) \tag{12}$$

The quantity $\tilde{G}_A(\omega)$ represents the spectral sensitivity of the "photodetector atom", that is the variation with ω of the transition intensities from the ground state a to a level b in the ionization continuum, at an energy $\hbar\omega$ above a. We shall assume here that the width $\Delta\omega_A$ of the function $\tilde{G}_A(\omega)$ is much larger than the bandwidth $\Delta\omega_R$ of the incident radiation.

$$\Delta\omega_A \gg \Delta\omega_R \tag{13}$$

[2]The ground state a is an eigenvector of the angular momentum component along the quantization axis Oz, with eigenvalue $m_a\hbar$; for the state b, the eigenvalue is $m_b\hbar$. The a to b transition thus corresponds to a variation $\Delta m = m_b - m_a$. If $\Delta m = 0$, symmetry arguments show that $\boldsymbol{e}_d$ is parallel to Oz; if $\Delta m = \pm 1$, $\boldsymbol{e}_d$ is in the plane perpendicular to Oz: $\boldsymbol{e}_d = (\boldsymbol{e}_x \mp i\boldsymbol{e}_y)/\sqrt{2}$; we note that the complex conjugate of this vector appears in the matrix element (7b).

Such a condition defines what we shall call a "broadband photodetector".

The field correlation function (9) has already been calculated in § B-2-a-β of Chapter XX assuming the radiation initial state is a Fock state $|n_1, ..., n_i, ...\rangle$, or a statistical mixture of such states with weights $p(n_1, ..., n_i, ...)$ – see equations (B-20) and (B-21) of Chapter XX. For the problem we are studying now, i.e. stationary non-monochromatic radiation, we can use the same assumption for the radiation initial state.

β. Photoionization rate

To transform equation (8), it is useful to study in more detail the $t''-t'$ dependence of $G_R^N(t''-t')$. We assume that the spectral distribution is centered around a non-zero value $\omega_{\rm ex}$, and that this distribution is entirely above the ionization threshold. Since the field $\bar{E}_d^{(+)}$ varies in $\mathrm{e}^{-i\omega t}$, and the field $\bar{E}_d^{(-)}$ in $\mathrm{e}^{i\omega t}$, we can then write:

$$G_R^N(t''-t') = e^{i\omega_{\rm ex}(t'-t'')} C_R^N(t''-t') \tag{14}$$

where $C_R^N(t''-t')$ is an "envelope" function whose Fourier transform is a function centered at $\omega = 0$ and of width $\Delta\omega_R$. This envelope function varies very slowly over time intervals short[3] compared to $1/\Delta\omega_R$. For $t'' = t'$, i.e. for $t''-t' = 0$, equation (14) leads to:

$$\begin{aligned} C_R^N(0) &= G_R^N(0) \\ &= \langle \phi_{\rm in}^R | \bar{E}_{d^*}^{(-)}(\boldsymbol{R}, t') \bar{E}_d^{(+)}(\boldsymbol{R}, t') | \phi_{\rm in}^R \rangle = I \end{aligned} \tag{15}$$

where I is the radiation intensity (which is time-independent since the radiation state is supposed to be stationary). We shall see in the next section how to generalize our results to non-stationary radiation.

Let us go back to the double integral of (8) and assume that the integration interval Δt satisfies the condition $\Delta t \gg \tau_{CA}$ where $\tau_{CA} = 1/\Delta\omega_A$ is the detector correlation time. In the plane t', t'', the function to be integrated in (8) is different from zero only in a band along the first bisector (Figure 1 of Chapter XX), of width τ_{CA} very narrow compared to Δt. If we change the integral variables t' and t'' to the variables t' and $\tau = t''-t'$, we can neglect the variation of $C_R^N(\tau)$ since, according to (13), $\tau_{CA} \ll 1/\Delta\omega_R$, and use (15) to rewrite (14) as:

$$\begin{aligned} G_R^N(t''-t') &= e^{i\omega_{\rm ex}(t'-t'')} C_R^N(t''-t') \\ &\simeq e^{i\omega_{\rm ex}(t'-t'')} C_R^N(0) = I e^{i\omega_{\rm ex}(t'-t'')} \end{aligned} \tag{16}$$

The double integral of (8) is easily performed with the new variables t' and τ. Using expression (3) for $G_R^N(t''-t')$, the integral over t' of a function that no longer depends on this variable introduces a simple factor Δt. We are then left with the integral over τ which leads to:

$$\sum_b P_{a\to b}^{\rm abs}(\Delta t) = \frac{1}{\hbar^2}\, I\, \Delta t \underbrace{\int_{-\infty}^{+\infty} \mathrm{d}\tau\; e^{-i\omega_{\rm ex}\tau} G_A(\tau)}_{=2\pi \bar{G}_A(\omega_{\rm ex})} \tag{17}$$

[3]This is not the case for $G_R^N(t''-t')$, because of the exponential $e^{i\omega_{\rm ex}(t'-t'')}$ that varies a lot over time intervals of the order of $1/\Delta\omega_R$, since $\omega_{\rm ex} \gg \Delta\omega_R$.

As this probability is proportional to Δt, we can define a photoionization rate:

$$W_{\text{phot.}} = \frac{1}{\Delta t} \sum_b P^{\text{abs}}_{a\to b}(\Delta t) = \frac{2\pi}{\hbar^2} I\, \bar{G}_A(\omega_{\text{ex}}) \tag{18}$$

This rate is proportional to the incident intensity I and to the spectral sensitivity $\tilde{G}_A(\omega_{\text{ex}})$ of the photodetector, evaluated at the radiation central frequency ω_{ex}.

2-c. Non-stationary and non-monochromatic radiation

For non-stationary radiation, the initial radiation state is no longer a Fock state or a statistical mixture of Fock states; it is rather a linear superposition of such states, creating wave packets such as those described in Complement D_{XX}. The radiation correlation function $G^N_R(t', t'')$ is no longer a function of the single variable $t'' - t'$, but depends on both t' and t''. One can still assume that the frequencies appearing in $G^N_R(t', t'')$ are centered around ω_{ex} in an interval of width $\Delta\omega_R$, which permits generalizing expressions (14) and (15) to:

$$G^N_R(t', t'') = e^{i\omega_{\text{ex}}(t'-t'')} C^N_R(t', t'') \tag{19}$$

and[4]:

$$G^N_R(t', t') = \langle \phi^R_{\text{in}} | E^{(-)}_{d^*}(\boldsymbol{R}, t') E^{(+)}_d(\boldsymbol{R}, t') | \phi^R_{\text{in}} \rangle = C^N_R(t', t') = I(\boldsymbol{R}, t') \tag{20}$$

where $I(\boldsymbol{R}, t')$ is the intensity at point $\boldsymbol{R}$ and at time t'.

Using the expansions in a_i and $a_i^\dagger$ for the field operators appearing in (10), we get an expression for $G^N_R(t', t'')$ that generalizes equation (B-20) of Chapter XX to non-stationary fields:

$$G^N_R(t', t'') = \sum_i \sum_j \frac{\hbar\sqrt{\omega_i \omega_j}}{2\varepsilon_0 L^3} (\boldsymbol{e}^*_d \cdot \varepsilon^*_j)(\boldsymbol{e}_d \cdot \varepsilon_i) \langle \phi^R_{\text{in}} | a^\dagger_j a_i | \phi^R_{\text{in}} \rangle\, e^{-i(\boldsymbol{k}_j \cdot \boldsymbol{R} - \omega_j t')} e^{i(\boldsymbol{k}_i \cdot \boldsymbol{R} - \omega_i t'')} \tag{21}$$

For fixed values of $\boldsymbol{k}_j$ and ω_j, the summation over i of (21) represents a wave packet of central frequency ω_{ex}, and whose envelope passes by a point $\boldsymbol{R}$ in a time interval of the order of $1/\Delta\omega_R$. If t' varies over a time interval $\Delta t \ll 1/\Delta\omega_R$, the variation of the envelope can be neglected. Similar conclusions are valid for a summation over j of relation (21), with fixed values for $\boldsymbol{k}_i$ and ω_i.

Let us now go back to the double integral in (8). As the phenomenon now depends on t, the integration interval will be taken between t and $t + \Delta t$ (instead of between 0 and Δt). We assume that Δt satisfies the condition :

$$\frac{1}{\Delta\omega_A} \ll \Delta t \ll \frac{1}{\Delta\omega_R} \tag{22}$$

which is possible when (13) is taken into account. Since $\Delta t \ll 1/\Delta\omega_R$, we can neglect in (19) the variation of $C^N_R(t', t'')$ when t' and t'' vary in the integration domain; we therefore replace $C^N_R(t', t'')$ by:

$$C^N_R(t, t) = \langle \phi^R_{\text{in}} | E^{(-)}_{d^*}(\boldsymbol{R}, t) E^{(+)}_d(\boldsymbol{R}, t) | \phi^R_{\text{in}} \rangle = I(\boldsymbol{R}, t) \tag{23}$$

[4]From now on, we simplify the notation by omitting the bar over operators in the Heisenberg picture, since the explicit time dependence is sufficient to indicate this point of viewn.

Using this equality in (19) we get, in the integration interval between t and $t + \Delta t$:

$$G_R^N(t', t'') \simeq e^{i\omega_{\rm ex}(t'-t'')} I(\boldsymbol{R}, t) \tag{24}$$

The computations are then quite similar to those carried out above for a stationary field: the integral over t' leads to a Δt term; the integral over $\tau = t'' - t'$ is equal to:

$$\int_{-\infty}^{+\infty} {\rm d}\tau \; e^{-i\omega_{\rm ex}\tau} G_A(\tau) = 2\pi \tilde{G}_A(\omega_{\rm ex}) \tag{25}$$

We call $W_I(\boldsymbol{R}, t)\Delta t$ the probability that a photodetector atom, placed at $\boldsymbol{R}$, will undergo a photoionization between times t and $t + \Delta t$. We get the result:

$$W_I(\boldsymbol{R}, t) = s\langle \phi_{\rm in}^R | E_{d^*}^{(-)}(\boldsymbol{R}, t) E_d^{(+)}(\boldsymbol{R}, t) | \phi_{\rm in}^R \rangle \tag{26}$$

where $s = 2\pi \bar{G}_A(\omega_{\rm ex})/\hbar^2$ is a factor characterizing the photodetector sensitivity at the radiation central frequency $\omega_{\rm ex}$. The atomic photoionization rate is thus a signal that constantly follows the time variations of the incident radiation intensity, written in (20).

2-d. Correlations between photoionization rates of two detector atoms

The previous computations can be generalized to analyze other experiments where two detector atoms are placed at $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$, and where we study correlations between the photoionizations observed on those two atoms at times t_1 and t_2. More precisely, let us call $W_{II}(\boldsymbol{R}_2, t_2; \boldsymbol{R}_1, t_1)\Delta t_1 \Delta t_2$ the probability to detect a photoionization at $\boldsymbol{R}_1$ between t_1 and $t_1 + \Delta t_1$ and another one at $\boldsymbol{R}_2$ between t_2 and $t_2 + \Delta t_2$. Computations very similar to those performed above, and which will not be explicited here (for more details, see Complement $\rm A_{II}$ of [21]), lead to:

$$W_{II}(\boldsymbol{R}_2, t_2; \boldsymbol{R}_1, t_1) = \\ s^2 \langle \phi_{\rm in}^R | E_{\boldsymbol{d}^*}^{(-)}(\boldsymbol{R}_1, t_1) E_{\boldsymbol{d}^*}^{(-)}(\boldsymbol{R}_2, t_2) E_{\boldsymbol{d}}^{(+)}(\boldsymbol{R}_2, t_2) E_{\boldsymbol{d}}^{(+)}(\boldsymbol{R}_1, t_1) | \phi_{\rm in}^R \rangle \tag{27}$$

It is easy to understand why two operators $\boldsymbol{E}_\perp^{(+)}$ preceded by two operators $\boldsymbol{E}_\perp^{(-)}$ appear in (27). The double photoionization rate W_{II} is computed from a probability that is the modulus squared of a probability amplitude for a photon to be absorbed at $\boldsymbol{R}_1, t_1$ and another one at $\boldsymbol{R}_2, t_2$. This amplitude must contain a product of two operators $\boldsymbol{E}_\perp^{(+)}$. Its conjugate must contain two operators $\boldsymbol{E}_\perp^{(-)}$ arranged in inverse order. We therefore should find in (27) two operators $\boldsymbol{E}_\perp^{(-)}$ followed by two operators $\boldsymbol{E}_\perp^{(+)}$ with different orders of $\boldsymbol{R}_1, t_1$ and $\boldsymbol{R}_2, t_2$.

There is a great analogy between the simple and double photoionization rates W_I and W_{II} given by equations (25) and (27) and the correlation functions G_1 and G_2 studied in Chapter XVI. The functions G_1 and G_2 give the probability densities of finding a particle at $\boldsymbol{r}_1, t_1$ for G_1, or a particle at $\boldsymbol{r}_1, t_1$ and another one at $\boldsymbol{r}_2, t_2$ for G_2. Note, however, that G_1 and G_2 give the probability of finding one or two particles at specific points, whereas we are now dealing with the probability of photoionization of atoms placed at specific points. The field operators appearing in (26) and (27) are the positive or negative frequency components of the electric field, since these are the operators describing the emission or absorption of photons.

3. Is a quantum treatment of radiation necessary to describe photoionization?

In a semiclassical treatment of photoionization, the radiation field is described as a classical field, while the atom follows a quantum treatment. The atom-field coupling is then a time-dependent perturbation that can induce transitions between a discrete atomic state, such as the ground sate a, and a state b, part of the ionization continuum. Does such a treatment yield the same results as those obtained in a quantum treatment? We are going to show that, while this is often the case, this is not always true.

3-a. Experiments with a single photodetector atom

In the simple case of an oscillating monochromatic classical field, with frequency ω_i, the transition probability per unit time takes a form that is reminiscent of the Fermi golden rule, valid for a constant perturbation coupling a discrete state to a continuum[5].

In the more general case where the classical field is non-monochromatic but stationary, one can follow the same line of computations that led to equation (8). This shows that the transition probability from a discrete state a to any state of the continuum can still be expressed as the integral of the product of two correlation functions: one, $G_A(t''-t')$, for the atomic dipole, another, $G_R(t''-t')$, for the radiation. In both cases, the field appears in the transition probability only via a correlation function. For the classical case, the quantum average value (10) must be replaced by the product of the negative and positive frequency components of the classical field:

$$\mathcal{G}_R^N(t''-t') = \mathcal{E}_{\boldsymbol{d}^*}^{(-)}(\boldsymbol{R},t')\ \mathcal{E}_{\boldsymbol{d}}^{(+)}(\boldsymbol{R},t'') \tag{28}$$

Note that in this relation, in order to distinguish the quantum fields from the classical fields, these latter fields are written with curly letters; the subscript d means that the field has been projected onto the polarization unit vector defined in (7a).

Expression (28) has been obtained for a perfectly known classical field. Another possibility is that the classical field is only known in a probabilistic sense, as is the case with a classical statistical mixture of fields, with given probabilities. The transition probability (8), where the correlation function $\bar{G}_R^N(t''-t')$ is replaced by $\mathcal{G}_R^N(t''-t')$, must then be averaged over all the states of the statistical mixture, which amounts to replacing (28) by:

$$\mathcal{G}_R^N(t''-t') = \overline{\mathcal{E}_{\boldsymbol{d}^*}^{(-)}(\boldsymbol{R},t')\ \mathcal{E}_{\boldsymbol{d}}^{(+)}(\boldsymbol{R},t'')} \tag{29}$$

where the bar above the product of the two fields symbolizes the statistical average. It seems that for all signals involving one single photodetector atom, the quantum predictions are identical to those of a semiclassical theory, using a classical field with the same correlation function as the quantum field. In particular, for a stationary field, the Fourier transform of the field correlation function is simply the spectral distribution $I_R(\omega)$ of that field. For a quantum field, this property was established in Chapter XX – see relation (B-20). For a classical field, this property is a consequence of the Wiener-Khintchine theorem. It follows that the photoionization probability of an atom is the same, whether it is computed with a stationary quantum or classical field, as long as they both have the same spectral distribution.

[5] See for example § C-3 of Chapter XIII, and relation (C-37) in particular.

Comment

The equivalence of the predictions of the two theories is also valid for a non-stationary field. The semiclassical theory predicts that the photoionization rate at time t' is proportional to the classical field intensity $\overline{\mathcal{E}_{\boldsymbol{d}^*}^{(-)}(\boldsymbol{R},t')\mathcal{E}_{\boldsymbol{d}}^{(+)}(\boldsymbol{R},t')}$, which now depends on t' since the field is no longer stationary. We shall see in Complement $\mathrm{D_{XX}}$ that a similar result is obtained in quantum theory: the photoionization probability at time t' of an atom receiving a one-photon wave packet is again given by the modulus squared of a function, which can be considered to be the photon wave function, evaluated at $t = t'$.

3-b. Experiments with two photodetector atoms

The same line of computations that led to equation (27) can be followed for a classical field. It leads to an expression similar to (27), where the field quantum correlation function is replaced by the statistical average of the product of two negative frequency components and two positive frequency components of the classical field:

$$\overline{\mathcal{E}_{\boldsymbol{d}^*}^{(-)}(\boldsymbol{R}_1,t_1)\mathcal{E}_{\boldsymbol{d}^*}^{(-)}(\boldsymbol{R}_2,t_2)\mathcal{E}_{d}^{(+)}(\boldsymbol{R}_2,t_2)\mathcal{E}_{d}^{(+)}(\boldsymbol{R}_1,t_1)} \tag{30}$$

As classical fields commute with each other, expression (30) can be rewritten as:

$$\overline{\mathcal{I}(\boldsymbol{R}_1,t_1)\mathcal{I}(\boldsymbol{R}_2,t_2)} \tag{31}$$

where:

$$\mathcal{I}(\boldsymbol{R}_1,t_1) = \mathcal{E}_{\boldsymbol{d}^*}^{(-)}(\boldsymbol{R}_1,t_1)\mathcal{E}_{d}^{(+)}(\boldsymbol{R}_1,t_1) \tag{32}$$

is the classical field intensity at point $\boldsymbol{R}_1$ at time t_1 and with a similar equation for $\mathcal{I}(\boldsymbol{R}_2,t_2)$. Correlations between the photoionization rates of the two photodetector atoms thus involve, in the semiclassical theory, the product of the intensities arriving on both photodetectors. The correlation function of the field amplitude is replaced by the correlation function of the intensity.

α. *Situations where a semiclassical treatment is adequate*

A first situation where quantum and semiclassical predictions agree is the case where the field state is a coherent state $|\{\alpha_i\}\rangle$ described by the set of classical normal variables $\{\alpha_i\}$ (Chapitre XIX, § B-3-b). Each mode i is in a coherent state $|\alpha_i\rangle$, meaning the state $|\{\alpha_i\}\rangle$ is an eigenket of the operator $\boldsymbol{E}^{(+)}(\boldsymbol{R},t)$ with eigenvalue $\mathcal{E}^{(+)}(\{\alpha_i\},\boldsymbol{R},t)$ equal to the classical field corresponding to the set of classical normal variables $\{\alpha_i\}$. In a similar way, the bra $\langle\{\alpha_i\}|$ is an eigenbra of $\boldsymbol{E}^{(-)}(\boldsymbol{R},t)$ with eigenvalue $\mathcal{E}^{(-)}(\{\alpha_i\},\boldsymbol{R},t)$. For a coherent state of the field, the rate W_{II} written in (27) therefore becomes: égal à:

$$\begin{aligned}
&W_{II}(\boldsymbol{R}_2,t_2;\boldsymbol{R}_1,t_1)\\
&= s^2\langle\{\alpha_i\}|\ E_{\perp\boldsymbol{d}^*}^{(-)}(\boldsymbol{R}_1,t_1)E_{\perp\boldsymbol{d}^*}^{(-)}(\boldsymbol{R}_2,t_2)E_{\perp\boldsymbol{d}}^{(+)}(\boldsymbol{R}_2,t_2)E_{\perp\boldsymbol{d}}^{(+)}(\boldsymbol{R}_1,t_1)\ |\{\alpha_i\}\rangle\\
&= s^2\mathcal{E}^{(-)}(\{\alpha_i\},\boldsymbol{R}_1,t_1)\ \mathcal{E}^{(-)}(\{\alpha_i\},\boldsymbol{R}_2,t_2)\ \mathcal{E}^{(+)}(\{\alpha_i\},\boldsymbol{R}_2,t_2)\ \mathcal{E}^{(+)}(\{\alpha_i\},\boldsymbol{R}_1,t_1)
\end{aligned} \tag{33}$$

The quantum result for W_{II} does coincide with the semiclassical prediction. The same conclusion holds when the state of the quantum field is a statistical mixture of coherent states $|\{\alpha_i\}\rangle$ with statistical weights $P(\{\alpha_i\})$.

Another situation where the quantum and semiclassical predictions agree is the case of a thermal field. In the quantum description, Wick's theorem (Complement C_{XVI}) allows expressing the four-point correlation function appearing in W_{II} as the sum of products of two-point correlation functions. In a similar way, in the semiclassical theory, the thermal field is a Gaussian random field, and here again, the classical four-point correlation function is the sum of products of two-point correlation functions. Provided we use the same two-point correlation functions in both theories, their predictions agree.

Comment

The interferometric analysis of the electric field of the light emitted by the stars (to measure their angular diameter) is confronted with the problem of atmospheric fluctuations, which introduce a random phase shift between the two arms of the interferometer. The analysis of intensity correlations is much less sensitive to these fluctuations. As different parts of the stars emit incoherent waves, the total field received from the star is Gaussian, and the result we just established shows that analyzing intensity correlations allows obtaining two-point correlation functions and hence the same information as the one contained in a field correlation measurement. The validity of such a method, based on intensity correlations, was experimentally demonstrated in 1956 by Robert Hanbury Brown and Richard Twiss [46].

β. Situations requiring a quantum radiation treatment

An example of a situation where a quantum treatment of the radiation becomes essential is shown in Figure (2). A one-photon wave packet is emitted by atom A; this wave packet is described by ϕ (cf. Complement D_{XX}). It then goes through a beamsplitter LS that divides it into two wave packets: a transmitted wave packet ϕ_T and a reflected wave packet ϕ_R, which then arrive on two detectors D_1 and D_2.

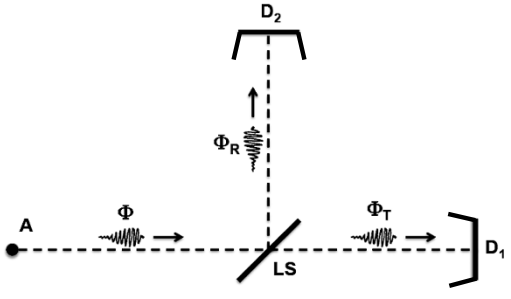


Figure 2: Atom A emits a photon described by a wave packet ϕ. This wave packet goes through a beamsplitter that divides it into a transmitted wave packet ϕ_T and a reflected wave packet ϕ_R, which then arrive on two detectors D_1 and D_2. The quantum and semiclassical predictions concerning the correlations between the signals detected on D_1 and D_2 are significantly different (see text).

In a quantum description of the radiation, the radiation state ψ after crossing the beamsplitter is still a one-photon state, described by the linear superposition of the two one-photon wave packets ϕ_T and ϕ_R, that is:

$$|\psi\rangle = c_T|\phi_T\rangle + c_R|\phi_R\rangle \tag{34}$$

In the expression (27) for the rate W_{II} giving the probability of a photoionization at time t_1 of detector D_1 at $\boldsymbol{R}_1$, and a photoionization at time t_2 of detector D_2 at $\boldsymbol{R}_2$, appears the squared norm of the ket:

$$E^{(+)}_{\perp \boldsymbol{d}}(\boldsymbol{R}_2, t_2)\ E^{(+)}_{\perp \boldsymbol{d}}(\boldsymbol{R}_1, t_1)\ |\psi\rangle \tag{35}$$

where ψ is given by (34). The first operator $E^{(+)}_{\perp \boldsymbol{d}}(\boldsymbol{R}_1, t_1)$, which destroys a photon, yields the vacuum $|0\rangle$ when it acts on state $|\psi\rangle$ that contains only one photon. The second destruction operator $E^{(+)}_{\perp \boldsymbol{d}}(\boldsymbol{R}_2, t_2)$ will then yield 0 when acting on the vacuum. The two detectors D_1 and D_2 cannot both undergo a photoionization. This result was, a priori, obvious: a single photon cannot produce two photoionizations.

If, on the other hand, the radiation emitted by the atom is described classically, the two wave packets ϕ_T and ϕ_R are classical wave packets which can ionize the detectors D_1 and D_2 they encounter.

Single photon sources are not easy to fabricate. An experiment close to the situation in Figure 2 is described in reference [47]. Instead of the atom A in Figure 2, it uses as a light source atoms emitting pairs of photons in a radiative cascade: the atom emits a photon of frequency ν_α going from a state α to a state β, then a photon ν_β going from state β to a state γ. If we call τ_β the radiative lifetime of state β, photon ν_β is emitted after photon ν_α in a time window having a width of the order of τ_β. Imagine we add to the experimental set-up of Figure 2 a third detector (not shown in the figure) that detects the ν_α photon and can trigger a departure time: after each detection of a ν_α photon, detectors D_1 and D_2 are activated, but for a short time interval of the order of τ_β. The probability of detecting a single photon ν_β during that time window is much higher than in a time window of the same length, but not triggered by the detection of a ν_α photon. This trigger method provides an equivalent of a single photon source, and was used to observe that a single ν_β photon could not simultaneously excite both detectors D_1 and D_2.

γ. *Resonance fluorescence of a single atom. Photon antibunching*

Another experiment clearly shows the need for a quantum description of radiation: the study of the second order correlation function W_{II} of the fluorescent light emitted by a single atom or ion, and excited by a resonant laser beam.

Imagine the emitting object A in Figure 2 is a single trapped ion[6]. Submitted to the resonant laser excitation, the ion emits a series of photons which enter the set-up of Figure 2. The distances between the beamsplitter and the detectors D_1 and D_2 are equal, so that the two wave packets associated with each photon arrive at the same time on D_1 and D_2.

[6]Ion trapping is now a well mastered technique. The results presented in Figure 3 have been obtained on a single $^{24}Mg^+$ ion [48]. The first experimental evidence for photon antibunching in the fluorescent light from a single atom were obtained on a sodium atomic beam at very low intensity, with an observation volume small enough for the probability of its containing more than one atom to be negligible [49].

For a continuous laser excitation with a constant intensity, the statistical properties of the fluorescent light are invariant under time translation; consequently, the quantum correlation function $W_{II}(\boldsymbol{R}_2, t_2; \boldsymbol{R}_1, t_1)$ characterizing the photoionizations detected on D_1 and D_2 depends only on $\tau = t_2 - t_1$. It shall be noted $g^{(2)}(\tau)$. Figure 3 plots the variations of $g^{(2)}(\tau)$ as a function of τ, for increasing values (from bottom to top) of the laser intensity. This figure shows that $g^{(2)}(\tau)$ is zero for $\tau = 0$; in other words, the detected photons are "antibunched" in time (one cannot detect simultaneously a photon at D_1 and another one at D_2).

The quantum interpretation of that result is as follows. Each photon emitted by the ion is detected either by D_1, or by D_2. Right after the emission of a photon, the ion is "projected" into the ground state a of the $a - b$ transition excited by the laser. This means it cannot immediately emit another photon as it must first be re-excited by the laser, and that takes a certain time. This is why $g^{(2)}(\tau)$ is zero for $\tau = 0$. Actually, after it emits a photon, the atom starting from g will oscillate between the state a and the excited state b at the Rabi frequency characterizing the atom-laser coupling, and proportional to the laser field amplitude. The Rabi oscillations explain the oscillations of $g^{(2)}(\tau)$ that appear in Figure 3 at frequencies higher and higher as the laser intensity increases.

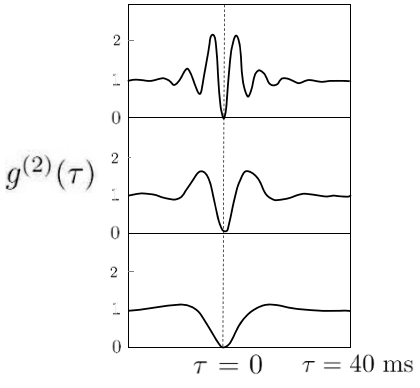


Figure 3: Intensity correlations in the resonance fluorescence of a single ion excited by a laser. The figure shows the time correlations $g^{(2)}(\tau)$ between the signals from the two detectors D_1 and D_2 as a function of the delay τ between two detections. The three curves correspond to increasing intensities (from bottom to top) of the laser beam exciting the resonant fluorescence of the trapped ion. It shows that $g^{(2)}(\tau)$ is zero for $\tau = 0$ and, for small positive values of τ, it increases with τ (figure adapted from [48]).

Let us examine now the predictions of a theory that classically treats the field emitted by the ion. We established above that the correlations between the photoionization rates of the two detectors are described by the correlation function $\overline{\mathcal{I}(t)\mathcal{I}(t+\tau)}$ of

the classical intensity $\mathcal{I}(t)$. For a stationary field, this classical correlation function $g_{\rm cl}^{(2)}$ depends only on τ:

$$\overline{\mathcal{I}(t)\mathcal{I}(t+\tau)} = g_{\rm cl}^{(2)}(\tau) \tag{36}$$

In addition, writing that $\overline{(\mathcal{I}(t)-\mathcal{I}(t+\tau))^2} \geqslant 0$, we get:

$$\overline{(\mathcal{I}(t))^2} + \overline{(\mathcal{I}(t+\tau))^2} \geqslant 2\ \overline{\mathcal{I}(t)\mathcal{I}(t+\tau)}$$

that is, taking into account the field stationarity and relation (36):

$$g_{\rm cl}^{(2)}(\tau=0) \geqslant g_{\rm cl}^{(2)}(\tau) \tag{37}$$

The semiclassical theory therefore predicts that $g_{\rm cl}^{(2)}(\tau)$ should not be an increasing function of τ in the vicinity of $\tau = 0$. This is contradicted by the experimental results shown in Figure 3, and hence proves that the fluorescent light emitted by a single ion excited by a resonant laser beam cannot be described as a classical field.

The radiation quantum theory is thus essential to account for all photoionization experimental results. This remains true even though the simple photoelectric effect observed on a single photodetector can be described by a semiclassical theory (without photons).

4. Two-photon photoionization

4-a. Differences with the one-photon photoionization

We now consider a two-photon absorption process similar to those studied in Complement A_{XX}, but where the final state b of the two-photon absorption process is now part of the atomic ionization continuum. This continuum starts at an energy E_I (ionization energy) above the energy of the ground state a (Fig. 4). This process is called two-photon photoionization.

The photoionization process transforms the atom into an ion and an electron, which moves away. When the distance between the electron and the ion is large enough, their Coulomb interaction energy becomes negligible and the electron energy is just its kinetic energy. Total energy conservation tells us that this kinetic energy is equal to:

$$E_{\rm kin} = 2\hbar\omega_i - E_I \tag{38}$$

If we plot the variations of $E_{\rm kin}$ as a function of ω_i, we get a straight half-line with slope $2\hbar$, which starts from the abscissa axis at point $E_I/2\hbar$. This result is a generalization of the photoelectric law established in 1905 by Einstein.

The previous result clearly shows that it is not necessary for the incident photon energy $\hbar\omega_i$ to be larger than the ionization energy for the atom to undergo photoionization. Figure 4 shows that $\hbar\omega_i$ is lower than E_I whereas $2\hbar\omega_i$ is larger than E_I. This result can be generalized: if $p\hbar\omega_i < E_I$, with $p = 1, 2, \ldots q-1$, but if $q\hbar\omega_i > E_I$, we have a q-photon photoionization. The kinetic energy of the photoelectron, once it is far enough from the ion, is equal to $E_{\rm kin} = q\hbar\omega_i - E_I$.

4-b. Photoionization rate

We first assume the radiation is monochromatic, and use expressions (13) and (14) of Complement A_{XX} for the two-photon absorption probability amplitude, whose modulus squared yields the probability. As the final state b now belongs to a continuum, we must sum this probability over b and use Fermi's golden rule to compute the photoionization probability per unit time. Since the modulus squared of equation (14) is proportional to $n_i(n_i - 1)$, where n_i is the number of incident photons, the photoionization rate increases as the square of the incident radiation intensity (for $n_i \gg 1$). In a similar way, it can be shown that a q-photon ionization increases as the qth power of the incident radiation (for $n_i \gg 1$).

Consider now the case of non-monochromatic stationary radiation. As in section § 2-b, we assume the radiation spectral density is centered around a frequency ω_{ex} with a width $\Delta\omega_R$ much smaller than the spectral bandwidth $\Delta\omega_A$ of the detector. The field correlation function $G_R^N(t' - t'')$ that appears in the two-photon absorption probability is still given by relation (27) of Complement A_{XX}. The two operators $\bar{\boldsymbol{E}}_\perp^{(+)}(\boldsymbol{R}, t')$ appearing in this equation have a predominantly exponential $e^{-i\omega_{ex}t'}$ time dependence. Similarly, the two operators $\bar{\boldsymbol{E}}_\perp^{(-)}(\boldsymbol{R}, t'')$ have a predominantly exponential $e^{+i\omega_{ex}t''}$ time dependence. With the same reasoning as in § 2-b-β, which led to (14), we set:

$$G_R^N(t'' - t') = e^{2i\omega_{ex}(t'-t'')} C_R^N(t'' - t') \tag{39}$$

where $C_R^N(t'' - t')$ is an "envelope" function with a much slower dependence in $t'' - t'$, and on a time scale of the order of $1/\Delta\omega_R$. In the double integral of (8), the correlation

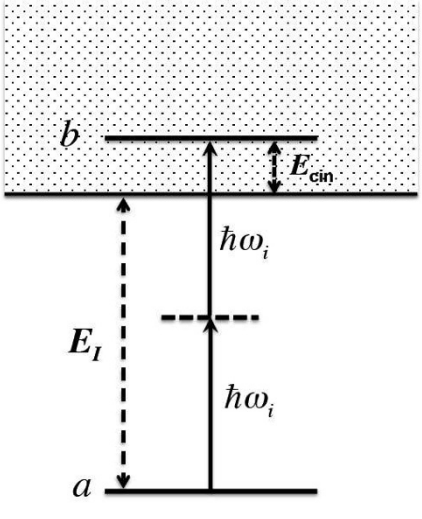


Figure 4: Two-photon photoionization. The atom goes from state a to state b, which is part of the ionization continuum, through the absorption of two photons, with energy $\hbar\omega_i$. The unbound electron produced at the end of that process leaves the atom with a kinetic energy equal, when the electron is far enough from the atom, to $E_{kin} = 2\hbar\omega_i - E_I$.

function $G_A(t''-t')$ is different from zero only for $|t''-t'| < \tau_{CA}$ where τ_{CA} is the correlation time of the atomic dipole, much shorter than $1/\Delta\omega_R$. We can therefore, as in § 2-b-β, take $t''=t'$ in the envelope function $C_R^N(t''-t')$ defined in relation (39) which yields G_R^N. We obtain for the field correlation function appearing in the two-photon ionization rate:

$$G_R^N(t''-t') = e^{2i\omega_{\rm ex}(t'-t'')}\times \langle\phi_{\rm in}^R|\left[\boldsymbol{e}_{ca}^*\cdot\boldsymbol{E}_\perp^{(-)}(\boldsymbol{R},t')\right]\left[\boldsymbol{e}_{bc}^*\cdot\boldsymbol{E}_\perp^{(-)}(\boldsymbol{R},t')\right] \left[\boldsymbol{e}_{bc}\cdot\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R},t')\right]\left[\boldsymbol{e}_{ca}\cdot\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R},t')\right]|\phi_{\rm in}^R\rangle \tag{40}$$

Note that the average value appearing in this equation is independent of t' since the radiation is supposed to be stationary.

4-c. Importance of fluctuations in the radiation intensity

Even when the radiation is monochromatic, i.e. when only a single mode i is populated, its intensity can take different values, spread out around an average value; the only case where the radiation intensity is well defined is when the radiation is in a Fock state $|n_i\rangle$. If we only consider stationary monochromatic radiation, the most general state is a statistical mixture of Fock states $|n_i\rangle$ with weight $p(n_i)$.

As an example, if the mode i is in thermal equilibrium at temperature T, the probability of its containing n_i photons is:

$$p(n_i) = \frac{1}{Z}\, e^{-\frac{(n_i+1/2)\hbar\omega_i}{k_B T}} \tag{41}$$

where k_B is the Boltzmann constant and Z the partition function given by:

$$Z = \sum_{n_i} e^{-\frac{(n_i+1/2)\hbar\omega_i}{k_B T}} \tag{42}$$

From these two equations one can easily compute the average value $\langle n\rangle$ of the number of photons n in this mode, as well as the average value $\langle n^2\rangle$ of n^2, for radiation at thermal equilibrium (see demonstration below). In particular, we can show that:

$$\langle n^2\rangle = \langle n\rangle + 2\langle n\rangle^2 \tag{43}$$

According to the previous results, the two-photon photoionization rate is proportional to $\langle n(n-1)\rangle = \langle n^2\rangle - \langle n\rangle$. If the radiation has a well-defined intensity, i.e. if it is in a Fock state $|n\rangle$, we have $\langle n^2\rangle = \langle n\rangle^2$ and the photoionization rate is proportional to $\langle n\rangle^2$ if $\langle n\rangle \gg 1$. On the other hand, if the radiation is in thermal equilibrium with the same average value $\langle n\rangle$, the photoionization rate is, according to (43), proportional to $2\langle n\rangle^2$, i.e. twice as large as for a state with no dispersion in intensity but same average value $\langle n\rangle$. An intensity fluctuation, keeping the average value $\langle n\rangle$ constant, considerably increases the photoionization rate. This result is to be expected for a nonlinear phenomenon: the values of n above the average value $\langle n\rangle$ contribute much more than the values below.

***Demonstration of relation* (43)**

This calculation has already been explained in §§ 2-b and 3-b of Complement B_{XV}. We briefly recall its principle. We note $x = \hbar\omega_i/k_BT$ and eliminate the irrelevant factor $e^{-x/2}$ from the partition function by setting:

$$Z(x) = e^{-x/2}K(x) \tag{44}$$

The function $K(x)$ is easily computed, since:

$$K(x) = 1 + e^{-x} + e^{-2x} + ... + e^{-nx} + ... = \frac{1}{1 - e^{-x}} \tag{45}$$

We also know that:

$$\begin{aligned}\langle n\rangle &= \sum_i n_i\ p(n_i) = \frac{1}{K(x)}\sum_i n_i\ e^{-n_i x} \\ &= \frac{-1}{K(x)}\frac{\mathrm{d}K(x)}{\mathrm{d}x}\end{aligned} \tag{46}$$

This leads to:

$$\langle n\rangle = \frac{1}{e^x - 1} \tag{47}$$

A similar calculation permits computing $\langle n^2\rangle$ using the second derivative $\mathrm{d}^2K(x)/\mathrm{d}x^2$, which leads to relation (43), the equivalent of relation (42b) of Complement B_{XV}.

If the radiation is no longer monochromatic, but still stationary and at thermal equilibrium, we can use the field correlation function (40) and Wick's theorem (Complement C_{XVI}) to rewrite this function as the sum of products of second order correlation functions:

$$\begin{aligned}G_R^N(t'' - t') =& e^{2i\omega_{\mathrm{ex}}(t'-t'')}\times \\ &\Big\{\langle\phi_{\mathrm{in}}^R|\left[\boldsymbol{e}_{ca}^*\cdot\boldsymbol{E}_\perp^{(-)}(\boldsymbol{R},t')\right]\left[\boldsymbol{e}_{ca}\cdot\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R},t')\right]|\phi_{\mathrm{in}}^R\rangle \\ &\quad\times\langle\phi_{\mathrm{in}}^R|\left[\boldsymbol{e}_{bc}^*\cdot\boldsymbol{E}_\perp^{(-)}(\boldsymbol{R},t')\right]\left[\boldsymbol{e}_{bc}\cdot\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R},t')\right]|\phi_{\mathrm{in}}^R\rangle \\ &+\langle\phi_{\mathrm{in}}^R|\left[\boldsymbol{e}_{ca}^*\cdot\boldsymbol{E}_\perp^{(-)}(\boldsymbol{R},t')\right]\left[\boldsymbol{e}_{bc}\cdot\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R},t')\right]|\phi_{\mathrm{in}}^R\rangle \\ &\quad\times\langle\phi_{\mathrm{in}}^R|\left[\boldsymbol{e}_{bc}^*\cdot\boldsymbol{E}_\perp^{(-)}(\boldsymbol{R},t')\right]\left[\boldsymbol{e}_{ca}\cdot\boldsymbol{E}_\perp^{(+)}(\boldsymbol{R},t')\right]|\phi_{\mathrm{in}}^R\rangle\Big\}\end{aligned} \tag{48}$$

We thus get a sum of two terms, each proportional to the square of an average intensity.

5. Tunnel ionization by intense laser fields

As high power lasers became available, studies of multi-photon ionization processes led to the discovery of many new physical phenomena. In particular, when the instantaneous laser field becomes of the order of the Coulomb field binding the electron to the nucleus, ionization no longer results from a multi-photon ionization process, but from a tunnel effect. The laser field, yielding a potential that varies linearly as a function of the electron-nucleus distance, lowers the Coulomb potential sufficiently to allow the electron to escape

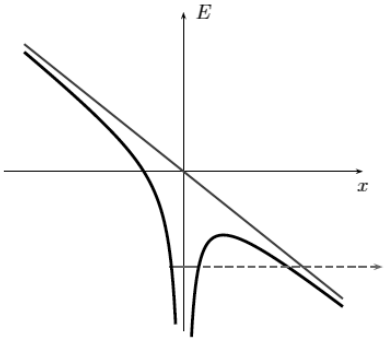


Figure 5: Effective potential seen by an electron undergoing tunnel ionization. The electron leaves the ion by tunneling through the potential barrier, sum of the ionic Coulomb potential and the linear potential $-qEx$ associated with the laser electric field, assumed to be linearly polarized along the x axis.

via a tunnel effect (see Fig.5). Once the electron has left the ion, it is accelerated by the laser field. As the oscillating laser field changes sign, the acceleration produced by the laser field is inverted, the electron comes back toward the ion and emits, as it passes close to the ion, a "bremsstrahlung" radiation (braking radiation). It can be shown that the frequency of this radiation is an odd harmonic of the laser frequency. The order of this harmonic is very high and can reach several hundred. As the fraction of the period of the laser field where the electron can escape by the tunnel effect is very small, the electron wave packet that leaves the ion has a very short time extension. The bremsstrahlung radiation it emits when it comes back to the ion also extends over a very short time, expressed in tens of attoseconds (one attosecond is equal to 10^{-18} sec). The interested reader can find an up to date review of these developments in Chapters 10 and 27 of reference [24].

Complement C_{XX}

Two-level atom in a monochromatic field. Dressed-atom method

1 Brief description of the dressed-atom method 2130
- 1-a State energies of the atom + photon system in the absence of coupling 2131
- 1-b Coupling matrix elements 2131
- 1-c Outline of the dressed-atom method 2133
- 1-d Physical meaning of photon number 2135
- 1-e Effects of spontaneous emission 2135

2 Weak coupling domain 2137
- 2-a Eigenvalues and eigenvectors of the effective Hamiltonian . . 2137
- 2-b Light shifts and radiative broadening 2138
- 2-c Dependence on incident intensity and detuning 2139
- 2-d Semiclassical interpretation in the weak coupling domain . . 2139
- 2-e Some extensions 2140

3 Strong coupling domain 2141
- 3-a Eigenvalues and eigenvectors of the effective Hamiltonian . . 2141
- 3-b Variation of dressed state energies with detuning 2142
- 3-c Fluorescence triplet 2144
- 3-d Temporal correlations between fluorescent photons 2145

4 Modifications of the field. Dispersion and absorption 2147
- 4-a Atom in a cavity 2147
- 4-b Frequency shift of the field in the presence of the atom 2148
- 4-c Field absorption 2149

Introduction.

The probability amplitude for an atom, subject to monochromatic radiation, to absorb a photon and go from a discrete state a to another discrete state b was calculated in § B of Chapter XX. We used, however, a perturbative treatment limited to lowest order with respect of the interaction Hamiltonian. The predictions of such an approximate calculation are, a priori, only valid for times that are sufficiently short for the higher order corrections to remain negligible. This complement presents another approach to atom-photon interactions, called the "dressed-atom approach", which does not have those limitations. It considers the atom and the mode of the quantum field it interacts with as a single quantum system. As this unified system is described by a time-independent Hamiltonian[1], one can study its energy diagram to obtain very useful information.

[1] This would obviously not be possible in a classical description of the radiation (even with a quantum treatment of the atom), since the field varies sinusoidally with time, and its coupling Hamiltonian with the atom is time-dependent.

This will allow us to improve the results of Chapter XX. The dressed-atom method yields a non-perturbative description of the physical processes under study, and hence remains valid even for intense fields. It provides new insights into several important physical phenomena: atoms' behavior in an intense electromagnetic field, spectral distribution of the light spontaneously emitted by an atom in such an intense field, time correlations between the emitted photons, origin of the forces exerted on an atom by radiation with a space-varying intensity.

As is usual in the literature, and as we already did in Complement C_{XIX}, we shall note g the atomic ground state and e the atomic excited state (instead of using our previous notation of a and b for the two atomic levels). For the same reason, rather than keeping the notation ω_i of Chapter XX for the angular frequency of the exciting beam, we shall use ω_L, that refers more explicitly to a laser beam, which can have a very high intensity.

We start in § 1 with a brief description of the dressed-atom method[2]. We assume the frequency ω_L to be close to resonance with the atomic frequency, noted $\omega_0 = (E_e - E_g)/\hbar$, but far enough from all the other atomic transition frequencies. The radiation-atom interaction will be characterized by a frequency called the "Rabi frequency". It is the equivalent, in this radiation quantum treatment, of the precession frequency of a spin turning around a classical radiofrequency field in a magnetic resonance experiment (Complement F_{IV}). Establishing the energy diagram of the unified atom-photon system will enable us to study both the weak coupling regime (Rabi frequency small compared to the natural width Γ of state e or to the detuning $|\omega_L - \omega_0|$ between the field and atomic frequencies) and the strong coupling regime (Rabi frequency large compared to the natural width and to the detuning).

The weak coupling regime is studied in § 2. We show that the ground state's energy undergoes a "light shift", by an amount that is proportional to the field intensity, and whose value as a function of the detuning follows a Lorentzian dispersion curve. The ground state also undergoes a radiative broadening, proportional to the field intensity, and which can be interpreted as a probability per unit time of leaving the ground state through the absorption of a photon.

We focus in § 3 on the strong coupling domain, where the Rabi oscillation between sates g and e appears, although damped because of the radiative instability of e. The energy diagram of the dressed-atom allows interpreting phenomena that are specific to the strong coupling regime, such as the fluorescence triplet and the temporal correlations between the photons emitted in the lateral components of that triplet.

The atom-field coupling perturbs, not only the atom, but also the field. We show in § 4 that the real and imaginary parts of the refractive index, associated with the atom's presence, are actually perturbations of the field, just as the light shifts and radiative broadening are perturbations of the atom.

1. Brief description of the dressed-atom method

Let us call "laser mode L" the quantum field mode that is populated by photons with frequency ω_L. In the absence of coupling between the atom A and this laser mode, the Hamiltonian of the total system $A + L$ is equal to $H_A + H_L$, where H_A is the atomic Hamiltonian, and H_L the Hamiltonian of the radiation in laser mode L. The energy levels

[2] A more detailed description can be found in Chapter VI of [21].

of $H_A + H_L$ are labeled by two quantum numbers: e or g for the atomic internal state[3], N for the photon number in the only radiation mode that is not empty and contains photons of frequency ω_L.

1-a. State energies of the atom + photon system in the absence of coupling

Consider the two states $|g, N\rangle$ and $|e, N-1\rangle$, whose energies with respect to the photon vacuum state are:

$$E_{g,N} = E_g + N\hbar\omega_L \qquad E_{e,N-1} = E_e + (N-1)\hbar\omega_L \tag{1}$$

Their energy difference is:

$$\begin{aligned} E_{g,N} - E_{e,N-1} &= \hbar\omega_L - (E_e - E_g) \\ &= \hbar(\omega_L - \omega_0) = \hbar\delta \end{aligned} \tag{2}$$

where:

$$\delta = \omega_L - \omega_0 \tag{3}$$

is the frequency detuning between the field frequency ω_L and the atomic transition frequency $\omega_0 = (E_e - E_g)/\hbar$. For a resonant field, i.e. when $\omega_L = \omega_0$, the two states are degenerate. Here we consider very small detunings (in absolute value) compared to ω_0:

$$|\delta| \ll \omega_0 \tag{4}$$

Consequently, even if the field is not exactly resonant, the two states $|g, N\rangle$ and $|e, N-1\rangle$ can be grouped in a two-dimensional multiplicity $\mathcal{E}(N)$:

$$\mathcal{E}(N) = \{|g, N\rangle, |e, N-1\rangle\} \tag{5}$$

which is far from all the other states of the system atom + field. There is an infinity of other multiplicities, for values of N going from 1 to infinity. As an example, Figure 1 shows the three multiplicities $\mathcal{E}(N-1)$, $\mathcal{E}(N)$ and $\mathcal{E}(N+1)$; there are $N-1$ others with lower energies, and an infinity with higher energies. Each multiplicity is separated from the next one by the distance $\hbar\omega_L$, and the spitting between the two levels inside one multiplicity equals $\hbar\delta$. The multiplicity $\mathcal{E}(1)$ corresponding to $N = 1$, includes the two states $|g, 1\rangle$ and $|e, 0\rangle$; on the other hand, the state $|g, 0\rangle$ is isolated.

1-b. Coupling matrix elements

The coupling H_I between the atom and the mode ω_L is proportional to the product of the atomic dipole moment $\boldsymbol{D}$ and the ω_L mode component of the radiation electric field $\boldsymbol{E}$. We can choose the origin of the coordinates so as to be able to write $e^{i\boldsymbol{k}_L\cdot\boldsymbol{R}} = -i$ where $\boldsymbol{R}$ is the position of the atomic center of mass. We then get for H_I:

$$H_I = -\sqrt{\frac{\hbar\omega_L}{2\varepsilon_0 L^3}}\ \boldsymbol{D}\cdot(\boldsymbol{\varepsilon}_L a_L + \boldsymbol{\varepsilon}_L^* a_L^\dagger) \tag{6}$$

[3]The atomic external degrees of freedom are treated classically by assuming that the atom is fixed at point $\boldsymbol{R}$.

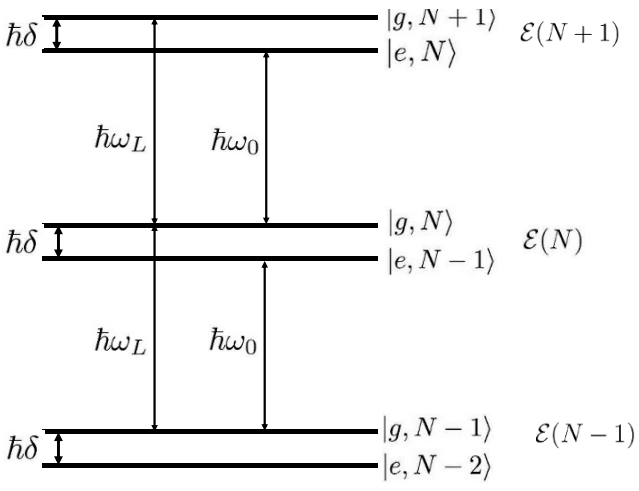


Figure 1: Energy levels of the system atom + photon in the absence of coupling. Only three adjacent multiplicities $\mathcal{E}(N-1)$, $\mathcal{E}(N)$ and $\mathcal{E}(N+1)$ are shown in the figure; many more exist below or above, corresponding respectively to smaller or larger values of N. Each vertical arrow links a pair of states having an energy difference indicated next to it.

The only non-zero matrix elements of the odd operator $\boldsymbol{D}$ are those between e and g. The annihilation and creation operators a_L and $a_L^\dagger$ change N by ∓ 1. It follows that $|g, N\rangle$ is coupled to $|e, N-1\rangle$ and $|e, N+1\rangle$, whereas $|e, N-1\rangle$ is coupled to $|g, N\rangle$ and $|g, N-2\rangle$. The two states $|g, N\rangle$ and $|e, N-1\rangle$ of the multiplicity $\mathcal{E}(N)$ are thus coupled to each other by the matrix element:

$$\langle e, N-1| H_I |g, N\rangle = \frac{\hbar\,\Omega_N}{2}, \tag{7}$$

where Ω_N is the "Rabi frequency" defined as:

$$\Omega_N = \sqrt{N}\ \ \Omega_1 \tag{8}$$

with:

$$\Omega_1 = \frac{2}{\hbar}\sqrt{\frac{\hbar\omega_L}{2\varepsilon_0 L^3}}\ \ \boldsymbol{\varepsilon}_L \cdot \langle e|\, \boldsymbol{D}\, |g\rangle \tag{9}$$

We assume Ω_1 is real and positive. If this is not the case, it suffices to change the relative phase[4] of the kets $|g\rangle$ and $|e\rangle$, which modifies the phase of the matrix element $\langle e|\, \boldsymbol{D}\, |g\rangle$; a suitable choice of that phase will make Ω_1 real and positive.

As operator H_I changes both the photon number by one unit and the atomic internal state, it does not have matrix elements between the kets of multiplicity $\mathcal{E}(N)$ and those of multiplicities $\mathcal{E}(N \pm 1)$. On the other hand, it can couple these kets with those of multiplicities $\mathcal{E}(N \pm 2)$ and, to higher order, to those of multiplicities even

[4]Such a phase change will affect the non-diagonal elements of the density matrix, leaving the physical predictions unchanged.

further away. However, the distances (in energy) between these multiplicities and $\mathcal{E}(N)$ are of the order of $2\hbar\omega_L$ (or of a multiple of that energy), whereas we have assumed that the interaction matrix elements are very small compared to $\hbar\omega_L$. The multiplicities other than $\mathcal{E}(N)$ have therefore energies too different from those of $\mathcal{E}(N)$ to play a significant role. We shall ignore their non-resonant coupling, which has a negligible effect for a quasi-resonant excitation.

1-c. Outline of the dressed-atom method

At the beginning of section § 1, we have described the quantum states of the system $A + L$ (atom + laser mode) in the absence of coupling; we showed they can be grouped into multiplicities $\mathcal{E}(N)$, with $N = 0, 1, 2, \ldots$ well separated from each other when condition (4) is satisfied. As an example, Figure 1 shows that the multiplicity $\mathcal{E}(N)$ for an atom with two levels g and e includes the states $|g, N\rangle$ and $|e, N-1\rangle$ separated by an energy $\hbar\delta$. Relation (7) tells us that these two states are coupled by an operator H_I describing the interaction between A and L, and that the corresponding matrix elements equal $\hbar\Omega_N/2$. We also discussed why, for a quasi-resonant excitation, the couplings between different multiplicities were negligible, so that one can separately study each multiplicity $\mathcal{E}(N)$.

α. *Dressed states and energies*

The first step in the dressed-atom approach is to study the energy levels of the system $A + L$ inside a multiplicity $\mathcal{E}(N)$, taking into account the coupling H_I restricted to $\mathcal{E}(N)$. We must diagonalize the Hamiltonian $H_A + H_L + H_I$ inside this sub-space. For a two-level atom, the restriction of that Hamiltonian to $\mathcal{E}(N)$, noted $H_{AL}^{(N)}$, is represented in the base of the kets written in (5) by a Hermitian 2×2 matrix equal to:

$$\begin{aligned}\left(H_{AL}^{(N)}\right) &= \begin{pmatrix} E_g + N\hbar\omega_L & \hbar\Omega_N/2 \\ \hbar\Omega_N/2 & E_e + (N-1)\hbar\omega_L \end{pmatrix} \\ &= [E_e + (N-1)\hbar\omega_L]\,\mathbb{1} + \hbar \begin{pmatrix} \delta & \Omega_N/2 \\ \Omega_N/2 & 0 \end{pmatrix} \end{aligned} \tag{10}$$

where $\mathbb{1}$ is the identity operator in $\mathcal{E}(N)$.

Because of the coupling created by the non-diagonal elements $\Omega_N/2$, the states $|g, N\rangle$ and $|e, N-1\rangle$ whose non-perturbed energies are separated by $\hbar\delta$ (left-hand side of Figure 2) are transformed into two states $|\psi_+(N)\rangle$ and $|\psi_-(N)\rangle$ with energies $\hbar\lambda_\pm$ (right-hand side of the figure):

$$\hbar\lambda_\pm = [E_e + (N-1)\hbar\omega_L] + \frac{\hbar\delta}{2} \pm \frac{\hbar}{2}\sqrt{\Omega_N^2 + \delta^2} \tag{11}$$

The new states $|\psi_+(N)\rangle$ and $|\psi_-(N)\rangle$ are linear superpositions of the initial states; they are called "dressed states", and their respective energies $\hbar\lambda_\pm$ the "dressed energies". This complement will show that a great number of interesting physical phenomena occurring when an atom is coupled to a laser mode can be interpreted in terms of these dressed states and their energies.

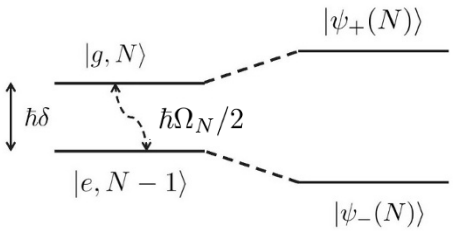


Figure 2: Energies of the states of the system $A+L$ within $\mathcal{E}(N)$, in the absence of coupling (left-hand side of the figure), and in the presence of a coupling of intensity $\hbar\Omega_N/2$ between the two initial states (right-hand side of the figure).

β. Rabi oscillation

Let us consider first a particularly simple application of the dressed-atom method. Imagine, for example, that the system is, at time $t=0$, in the state $|g,N\rangle$:

$$|\psi(t=0)\rangle = |\psi_{\text{in}}\rangle = |g,N\rangle \tag{12}$$

and let us try to find the probability that it will be found at a later time t in the state:

$$|\psi_{\text{fin}}\rangle = |e,N-1\rangle \tag{13}$$

We are dealing with the evolution of a system with two levels coupled by a static perturbation $\hbar\Omega_N/2$. This problem was studied in detail in § C of Chapter IV. We must first expand the initial state on the states $|\psi_\pm(N)\rangle$, and multiply each of them by an exponential $e^{-i\lambda_\pm t}$ whose argument is proportional to its energy $\hbar\lambda_\pm$:

$$|\psi(t)\rangle = e^{-i\lambda_+ t}\,\langle\psi_+(N)\,|g,N\rangle\quad|\psi_+(N)\rangle + e^{-i\lambda_- t}\,\langle\psi_-(N)\,|g,N\rangle\;\;|\psi_-(N)\rangle \tag{14}$$

The probability amplitude of finding the system in state $|e,N-1\rangle$ is then:

$$\begin{aligned}\langle e,N-1\,|\psi(t)\rangle \propto\; & e^{-i\lambda t/2}\,\langle\psi_+(N)\,|g,N\rangle\;\;\langle e,N-1\,|\psi_+(N)\rangle \\ & + e^{i\lambda t/2}\,\langle\psi_-(N)\,|g,N\rangle\;\;\langle e,N-1\,|\psi_-(N)\rangle\end{aligned} \tag{15}$$

where we introduced the Bohr frequency:

$$\lambda = \lambda_+ - \lambda_- = \sqrt{\Omega_N^2+\delta^2} \tag{16}$$

In (15), the sign $\propto$ simply means that we omitted a global phase factor $e^{-i(\lambda_+ +\lambda_-)t/2}$, with no physical significance. The probability of finding the system at time t in the final state (13) is therefore:

$$\begin{aligned}\mathcal{P}(t) &= |\langle e,N-1\,|\psi(t)\rangle|^2 \\ &= \left|C_+^g\right|^2\left|C_+^e\right|^2 + \left|C_-^g\right|^2\left|C_-^e\right|^2 + C_+^g\left(C_+^e\right)^*\left(C_-^g\right)^* C_-^e\;e^{-i\lambda t} + \text{c.c.}\end{aligned} \tag{17}$$

where the C are the scalar products:

$$C_{\pm}^{g} = \langle \psi_{\pm}(N) \, | g, N \rangle \qquad \text{and} \qquad C_{\pm}^{e} = \langle \psi_{\pm}(N) \, | e, N-1 \rangle \tag{18}$$

and c.c. means the complex conjugate.

We see that the probability $\mathcal{P}(t)$ is an oscillating function of time, with a frequency λ that is the only Bohr frequency $\sqrt{\Omega_N^2 + \delta^2}$ of the system within the multiplicity $\mathcal{E}_N$. This frequency can obviously be expanded to all orders of the perturbation Ω_N^2, but the result we obtained is not perturbative. The oscillation we found concerns the total system formed by a two-level atom placed in a monochromatic radiation field that can be intense and resonant: starting from state $|g, N\rangle$, the atom absorbs a photon and goes to state $|e, N-1\rangle$; it then comes back to state $|g, N\rangle$ by stimulated emission of a photon, and so on.

1-d. Physical meaning of photon number

The situation is different depending on whether the atom is placed in a real cavity or in free space.

If the atom is placed in a real cavity, as in some experiments, the field modes are the cavity eigenmodes. Such a situation will be discussed in § 4; the photon number N then has a perfectly clear physical meaning. The volume L^3, appearing in the modal expansion of the fields, and which is found in expressions (6) and (9) above, is simply the volume of the cavity containing the photons.

If the atom is in free space, the volume L^3, introduced to obtain discrete modes, is simply used in the computation, without any precise physical meaning. On the other hand, the energy density in the vicinity of the atom, proportional to $N\hbar\omega_L/L^3$, does have a physical meaning. Provided we keep N/L^3 constant, we can change N and L^3 arbitrarily without changing the coupling between the atom and the field; this is because the coupling is characterized by the Rabi frequency Ω_N, which depends on $\sqrt{N/L^3}$. In that case, the photon number N does not have an intrinsic physical meaning.

Imagine, for example, that the field is in a coherent state (Complement $\mathrm{G_V}$). The values of N are then distributed around an average value $\langle N \rangle$ in an interval of width $\Delta N = \sqrt{\langle N \rangle}$, very small in relative value compared to $\langle N \rangle$, but very large in absolute value. If both N and L^3 go to infinity, keeping the ratio N/L^3 constant, the Rabi frequency Ω_N will barely change in relative value even when N varies over a large interval around $\langle N \rangle$. The frequency Ω_N can thus be replaced in (10) by a constant Ω_R (which does not depend on N):

$$\Omega_N \simeq \Omega_R \tag{19}$$

This will be done in what follows and in § 3.

1-e. Effects of spontaneous emission

We have ignored, until now, all the field modes others than the laser mode. However, when the atom is in the excited state e, it can spontaneously emit a photon in another mode. This means that, in addition to the atom A and the laser mode L, we must take into account the system R including all the modes that, initially, did not contain any photons. As R is a very large system (sometimes called "reservoir" for that

reason), the coupling effects between $A + L$ and R must be described by a so-called "master equation"; this equation describes the evolution of the density operator σ_{A+L} of $A+L$ under the effect of the coupling with R (see part D of Chapter VI in [21]). Though we shall not introduce this master equation here, we shall merely discuss the physical interpretation of the results it leads to.

As the frequency spectrum of the reservoir R has a width $\Delta\omega$ of the order of the optical frequency, its associated correlation time $\tau_c \simeq 1/\Delta\omega$ is much shorter than all the other characteristic times of the problem. It is, in particular, much shorter than the radiative lifetime τ_R:

$$\tau_R = \frac{1}{\Gamma} \tag{20}$$

where Γ is the natural width of the excited state e; it is also shorter than the inverse of the Rabi frequency, which yields the characteristic time of the coupling to the laser mode. This means that, when the system $A + L$ is in the state $|e, N-1\rangle$ and a spontaneous emission occurs, it lasts for a time interval too short for the atom to have sufficient time to couple with L. The system then goes quasi-instantaneously from state $|e, N-1\rangle$, which belongs to $\mathcal{E}(N)$, to state $|g, N-1\rangle$ in the lower multiplicity $\mathcal{E}(N-1)$ – see Figure 1. As a consequence (see § C-3 of Chapter III in [21], as well as Complement D_{XIII}), the evolution within $\mathcal{E}(N)$ can still be described by the same equations as above, provided we simply add an imaginary term to the energy E_e of the excited state:

$$E_e \Longrightarrow E_e - i\hbar\frac{\Gamma}{2} \tag{21}$$

This means that, to describe the evolution of the system $A+L$ within[5] $\mathcal{E}(N)$ while taking into account spontaneous emission processes, we must replace the Hamiltonian $H_{AL}^{(N)}$ written in (10) by the effective non-Hermitian Hamiltonian:

$$\left(H_{\text{eff}}^{(N)}\right) = [E_e + (N-1)\hbar\omega_L]\,\mathbb{1} + \hbar\begin{pmatrix} \delta & \Omega_R/2 \\ \Omega_R/2 & -i\Gamma/2 \end{pmatrix} \tag{22}$$

Because of the imaginary term $-i\Gamma/2$ appearing in the matrix, the two eigenvalues of $H_{\text{eff}}^{(N)}$ also have an imaginary part: the two dressed states are now unstable as a result of spontaneous photon emission, which can occur in any of these states.

Within a constant factor, given by the term proportional to $\mathbb{1}$ on the right-hand side of (22), the eigenvalues $\hbar\lambda_\pm$ of H_{eff} are obtained by diagonalizing the 2-dimensional matrix that follows on the right-hand side of (22). The exact solution is written[6]:

$$\lambda_\pm = -i\frac{\Gamma}{4} + \frac{\delta}{2} \pm \frac{1}{2}\sqrt{\Omega_R^2 + \left(\delta + i\frac{\Gamma}{2}\right)^2} \tag{23}$$

We now discuss the physical meaning of these results in two limiting cases.

[5] The coupling with the reservoir induces other important effects leading to transitions between different multiplicities. This is what happens, for example, in the fluorescence phenomenon studied in § 3-c.

[6] For brevity, we use a slighlty incorrect mathematical notation, since the square root sign must in principle be applied on a real and positive number. What we mean with the square root sign written on the right-hand side of (23), is either one of the two complex numbers whose square is equal to the complex number under the root sign.

2. Weak coupling domain

We start with the weak coupling domain, which is more directly related with the results of Chapter XX.

2-a. Eigenvalues and eigenvectors of the effective Hamiltonian

Consider first the case where the non-diagonal coupling $\hbar\Omega_R/2$ between the two non-perturbed states of $\mathcal{E}(N)$ is small compared to the differences between the energies of these two states (including the imaginary term associated with the natural width of e). As this difference is complex, we must take its modulus:

$$\Omega_R \ll \left|\delta + i\frac{\Gamma}{2}\right| \tag{24}$$

This inequality is satisfied if:

$$\Omega_R \ll \Gamma \quad \text{or} \quad \Omega_{\mathrm{R}} \ll |\delta| \tag{25}$$

The weak coupling domain is thus obtained for low light intensities, or large frequency detunings.

For weak coupling, we can apply perturbation theory to obtain the energy corrections for the states $|g, N\rangle$ and $|e, N-1\rangle$, to order 2 in Ω_R. Starting from (22), we obtain in this way the correction $\delta E_{g,N}$ to the energy of state $|g, N\rangle$:

$$\delta E_{g,N} = \hbar\frac{(\Omega_R/2)^2}{\delta + i\Gamma/2} = \hbar\delta_g - i\hbar\frac{\gamma_g}{2} \tag{26}$$

where:

$$\delta_g = \frac{\delta}{4\delta^2 + \Gamma^2}\Omega_R^2 \qquad \text{and} \qquad \gamma_g = \frac{\Gamma}{4\delta^2 + \Gamma^2}\Omega_R^2 \tag{27}$$

A similar calculation yields for the correction to the energy of state $|e, N-1\rangle$:

$$\delta E_{e,N-1} = -\hbar\delta_g + i\hbar\frac{\gamma_g}{2} \tag{28}$$

We can write the approximate eigenvalues of the effective Hamiltonian (22) in the form:

$$\begin{aligned}\lambda_+ &\simeq \delta + \delta_g - i\frac{\gamma_g}{2} + ..\\ \lambda_- &\simeq -\delta_g - i\frac{\Gamma}{2} + i\frac{\gamma_g}{2} + ..\end{aligned} \tag{29}$$

which coincides with an expansion in powers of Ω_R/Γ of the exact result (23).

Perturbation theory also allows computing the eigenstates of H_{eff} to first order in Ω_R. The state $|\overline{g, N}\rangle$, which tends towards $|g, N\rangle$ when Ω_R goes to zero, is written:

$$|\overline{g, N}\rangle = |g, N\rangle + \frac{(\Omega_R/2)}{\delta + i\Gamma/2}|e, N-1\rangle \tag{30}$$

This means that state $|g, N\rangle$ is "contaminated" by state $|e, N-1\rangle$. A similar computation for state $|\overline{e, N-1}\rangle$, which tends towards $|e, N-1\rangle$ when Ω goes to zero, yields:

$$|\overline{e, N-1}\rangle = |e, N-1\rangle - \frac{(\Omega_R/2)}{\delta + i\Gamma/2}|g, N\rangle \tag{31}$$

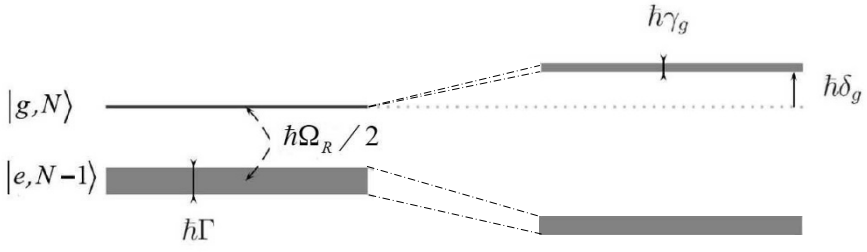


Figure 3: Non-perturbed states (left-hand side of the picture) and perturbed states (right-hand side) in the $\mathcal{E}(N)$ multiplicity. The coupling, characterized by the Rabi frequency Ω_R, shifts state $|g,N\rangle$ by a quantity $\hbar\delta_g$ (representing the light shift of the ground sate g); its wave function is "contaminated" by the unstable wave function of state $|e,N-1\rangle$, meaning that the ground state also becomes unstable as shown by its radiative broadening $\hbar\gamma_g$. State $|e,N-1\rangle$ is shifted in the opposite way, compared to $|g,N\rangle$; its width is reduced from $\hbar\Gamma$ to $\hbar(\Gamma-\gamma_g)$.

2-b. Light shifts and radiative broadening

The real parts of $\delta E_{g,N}$ and $\delta E_{e,N-1}$ represent shifts in the energy levels induced by the coupling with the light and called for that reason "light shifts". The imaginary part of $\delta E_{g,N}$ represents a radiative broadening of state $|g,N\rangle$, which becomes unstable under the coupling effect. The imaginary part of $\delta E_{e,N-1}$ describes a reduction of the radiative broadening $\hbar\Gamma$ of state $|e,N-1\rangle$.

Figure 3 shows, in its left-hand side, the non-perturbed states $|g,N\rangle$ and $|e,N-1\rangle$ in the $\mathcal{E}(N)$ multiplicity. They are separated by the gap $\hbar\delta$; if δ is positive, state $|g,N\rangle$ is above state $|e,N-1\rangle$; conversely, if δ is negative, state $|e,N-1\rangle$ is now above state $|g,N\rangle$. The thickness of the line representing state $|e,N-1\rangle$ symbolizes its natural width $\hbar\Gamma$. The right-hand side of the figure represents the states perturbed by the interaction with light. The two states $|g,N\rangle$ and $|e,N-1\rangle$ repel each other, meaning that they undergo light shifts of opposite signs. The $\hbar\delta_g$ shift of $|g,N\rangle$ is positive if $|g,N\rangle$ is above $|e,N\rangle$, i.e. if δ is positive; it is negative if δ is negative. The stable $|g,N\rangle$ state is also "contaminated" by the unstable state $|e,N-1\rangle$, which makes it unstable as shown by its radiative broadening $\hbar\gamma_g$. An atom in state g cannot stay there indefinitely: it will leave that state with a probability per unit time equal to γ_g, which can be interpreted as the photon absorption rate of an atom in state g. Conversely, due to the "contamination" of the unstable state $|e,N-1\rangle$ by the stable state $|g,N\rangle$, state $|e,N-1\rangle$ becomes less unstable and its width is reduced from $\hbar\Gamma$ to $\hbar(\Gamma-\gamma_g)$.

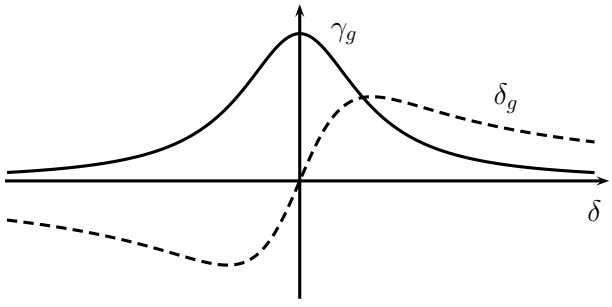


Figure 4: Plots of the light shift $\hbar\delta_g$ (dashed line curve) and of the radiative broadening $\hbar\gamma_g$ (solid line curve) as a function of the detuning δ between the laser frequency and the atomic frequency.

2-c. Dependence on incident intensity and detuning

The shifts $\hbar\delta_g$ and radiative broadenings $\hbar\gamma_g$ given by equation (27) are all proportional to Ω_R^2, hence to the number of incident photons, meaning to the light intensity. Their variations with the detuning δ follow respectively Lorentzian dispersion and absorption curves (Fig. 4).

When the detuning is very large (in absolute value) compared to the natural width of e ($|\delta| \gg \Gamma$), we can neglect Γ^2 compared to $4\delta^2$ in the denominators of expressions (27), which yields:

$$|\delta_g| = \frac{\Omega_R^2}{4|\delta|} \qquad \gamma_g = \frac{\Omega_R^2 \Gamma}{4\delta^2} \tag{32}$$

This leads to:

$$\gamma_g = |\delta_g| \frac{\Gamma}{|\delta|} \ll |\delta_g| \tag{33}$$

For large detunings, the light shifts are thus much larger than the radiative broadenings.

2-d. Semiclassical interpretation in the weak coupling domain

In this weak coupling domain, the atom responds linearly to the incident field; the results we just discussed can be interpreted semiclassically, in terms of a dipole induced by the incident field (see for example [50]). This dipole has a component in phase with the field and a quadrature component, related to the field by a dynamic polarizability $\alpha(\omega)$.

The quadrature component absorbs energy from the field. It varies with the detuning δ as an absorption curve; it is responsible for the absorption rate associated with the radiative broadening γ_g. The in-phase component of the dipole yields a polarization energy. Its variation with the detuning follows a dispersion curve. It is responsible for the light shift, just as the Stark shift results from the interaction of a static electric field with the static dipole it induces. This is why this light shift is often called a "dynamic Stark effect".

2-e. Some extensions

We now discuss some direct, important extensions of the previous study.

α. Non-monochromatic incident radiation

Imagine the radiation state is now a Fock state $|n_1, n_2, ..., n_i, ..\rangle$, or a statistical mixture of such states; the radiation spectral distribution is then described by the function $I(\omega_L)$. To second order perturbation theory (weak coupling domain), the processes that come into play in the light shifts and radiative broadenings are stimulated absorption and re-emission of photons. When several modes contain photons and the radiation state is a Fock state, the photon must be re-emitted by stimulated emission in the same mode it was absorbed from (otherwise the matrix element describing the second order coupling would be zero). This means that the effects of the different field modes can be added independently; we then get for δ_g and γ_g:

$$\delta_g \propto \int \mathrm{d}\omega_L |\mathcal{D}_{eg}|^2 I(\omega_L) \ \frac{\delta}{4\delta^2 + \Gamma^2}$$
$$\gamma_g \propto \int \mathrm{d}\omega_L |\mathcal{D}_{eg}|^2 I(\omega_L) \ \frac{\Gamma}{4\delta^2 + \Gamma^2} \tag{34}$$

β. Degenerate ground state

Assume the ground state g has a non-zero angular momentum J_g and therefore contains several Zeeman sublevels m_g; one can then show [51] that the sublevels of g having a well-defined light shift and radiative broadening are obtained by diagonalizing the Hermitian matrix whose elements are:

$$\sum_{m_e} \langle m_g | \boldsymbol{\varepsilon}_L^* \cdot \boldsymbol{D} | m_e \rangle \langle m_e | \boldsymbol{\varepsilon}_L \cdot \boldsymbol{D} | m_g' \rangle \tag{35}$$

where the states m_e are the sublevels of e. The eigenstates $|g_\alpha\rangle$ of this matrix, with eigenvalues μ_α, undergo light shifts proportional to $\mu_\alpha \delta_g$ and radiative broadenings proportional to $\mu_\alpha \gamma_g$ (where δ_g and γ_g are the shifts and broadenings for a two-level atom).

Reference [52] studies the symmetry properties of matrix (35), and discusses the equivalence between the light shifts and the effect of fictitious magnetic and electric fields acting on the ground multiplicity of the atom. We shall simply focus here on the simple case of a $J_g = 1/2 \leftrightarrow J_e = 1/2$ transition such as, for example, the hyperfine component $F = 1/2 \leftrightarrow F = 1/2$ of the $6^1S_0 \leftrightarrow 6^3P_1$ transition of the 199-isotope of mercury (λ_L= 253.7 nm). It is on such a transition that light shifts were observed for the first time [53].

The left-hand side of Figure 5 shows the components σ_+ and σ_- of this transition, that link respectively $m_g = -1/2$ to $m_e = +1/2$ and $m_g = +1/2$ to $m_e = -1/2$. If the beam polarization is σ_+, level $m_g = -1/2$ has a non-zero and well defined light shift, since the absorption and re-emission of a σ_+ photon can link sublevel $m_g = -1/2$ only to itself. On the other hand, level $m_g = +1/2$ is not shifted because there is no σ_+ optical transition starting from $m_g = +1/2$. We get opposite conclusions for a σ_- polarization of the light beam: the light shift of sublevel $m_g = +1/2$ is well-defined and sublevel $m_g = -1/2$ is not shifted. Now, by symmetry, the Clebsch-Gordan coefficients (Complement B_X) for the σ_+ and σ_- transitions are equal; the light shifts have the same

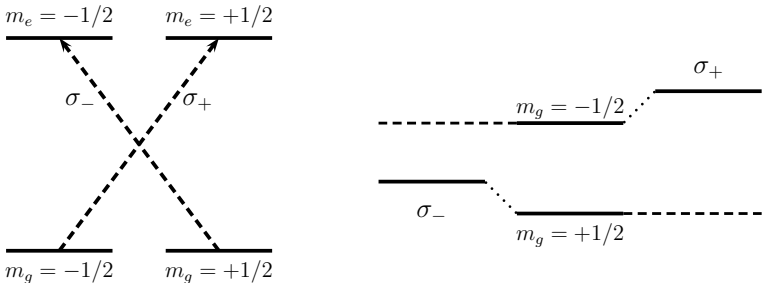


Figure 5: The left-hand side of the figure represents the $J_g = 1/2 \to J_e = 1/2$ transition, and the light beam polarizations that can induce transitions between its Zeeman sublevels. The diagram on the right-hand side plots in its center the ground state energy levels in the absence of any light beam (Zeeman levels in a static magnetic field); the two lateral extensions depict their light shifts by a non-resonant light beam with polarization σ_+ (on the right), or σ_- (on the left). In the first case, the light selectively shifts the sublevel $m_g = -1/2$, in the second case, it shifts the sublevel $m_g = +1/2$. This is why, depending on whether the beam polarization is σ_+ or σ_-, the variation in the gap between the two Zeeman sublevels changes sign.

value for a σ_+ excitation of sublevel $m_g = -1/2$ and for a σ_- excitation with same intensity of sublevel $m_g = +1/2$.

In the presence of a static magnetic field, there is an energy gap between the two atomic sublevels $m_g = \pm 1/2$ (Zeeman effect). The right-hand side of Figure 5 shows that a non-resonant light excitation changes this gap by the same amount, but in the opposite directions[7] depending on whether it has a σ_+ or σ_- polarization. The ground state magnetic resonance line, detected by optical methods using a resonant beam (Complement C_{XIX}, § 2-b), is thus shifted when a second non-resonant beam is applied; this shift has opposite directions, depending on whether that beam has a σ_+ or σ_- polarization. As relaxation times can be very long in the ground state, its magnetic resonance line is very narrow, which allows detecting very small light shifts, of the order of a few Hz. This is how the existence of light shifts were demonstrated in 1961, when laser sources were not yet available in laboratories [53]. With laser sources, one routinely observes shifts of the order of 10^6 Hz, and even more.

3. Strong coupling domain

We now examine how the previous results are modified in the strong coupling regime.

3-a. Eigenvalues and eigenvectors of the effective Hamiltonian

A strong coupling regime means that the non-diagonal element $\hbar\Omega_R/2$ of the effective Hamiltonian written in (22) is large compared to the difference between two diagonal elements:

$$\Omega_R \gg |\delta| \quad \text{and} \quad \Omega_R \gg \Gamma \tag{36}$$

[7] We assume the detuning δ is large compared to the Zeeman splitting.

For the sake of simplicity, we shall only consider the resonant case ($\delta = 0$). Equation (23), which yields the eigenvalues of the 2×2 matrix of (22) for any value of Ω_R, then becomes:

$$\lambda_{\pm} = -i\frac{\Gamma}{4} \pm \frac{1}{2}\sqrt{\Omega_R^2 - \frac{\Gamma^2}{4}} \tag{37}$$

where, as we did above, we use the concise notation for the square root of a number that is not always positive (see note 6).

As long as $\Omega_R < \Gamma/2$, the last term on the right-hand side of (37) is purely imaginary. The same is true for the two eigenvalues $\lambda_{\pm}$, which are equal to:

$$\lambda_{\pm} = -i\left[\frac{\Gamma}{4} \mp \frac{1}{2}\sqrt{\frac{\Gamma^2}{4} - \Omega_R^2}\right] \tag{38}$$

If, in addition, $\Omega_R \ll \Gamma$, a limited expansion of (38) in powers of Ω_R/Γ yields $\lambda_+ = -i\gamma_g/2$ and $\lambda_- = -i(\Gamma - \gamma_g)/2$; as expected, we confirm the results of the previous § for the weak coupling regime. As Ω_R increases, while remaining lower than $\Gamma/2$, the eigenvalue λ_+ increases whereas λ_- decreases, but their sum $(\lambda_+ + \lambda_-)$ remains constant and equal to $-i\Gamma/2$. When Ω_R reaches the value $\Gamma/2$, both eigenvalues $\lambda_{\pm}$ are equal to $-i\Gamma/4$.

As soon as Ω_R goes beyond $\Gamma/2$, the last term in (37) becomes real. The two eigenvalues $\lambda_{\pm}$ have opposite real parts and the same imaginary part, equal to $-i\Gamma/4$; the two dressed levels now always have the same width $\Gamma/4$. As the coupling becomes strong ($\Omega_R \gg \Gamma$), the energies are equal to:

$$\lambda_{\pm} \simeq -i\frac{\Gamma}{4} \pm \frac{\Omega_R}{2} \tag{39}$$

and the eigenvectors tend toward symmetric and antisymmetric linear combinations of $|g, N\rangle$ and $|e, N-1\rangle$:

$$|\psi_{\pm}(N)\rangle \to \frac{1}{\sqrt{2}}\left[|g, N\rangle \pm |e, N-1\rangle\right] \tag{40}$$

Such states can no longer be considered to be a result of a light mutual contamination of the non-perturbed states of multiplicity $\mathcal{E}(\mathcal{N})$. They are actually *entangled*, and hence impossible to consider as products of an atomic state and a field state. These states of the global atom + field system are commonly called *atom-field dressed states*.

3-b. Variation of dressed state energies with detuning

The solid lines in Figure 6 show the energies of the dressed states $|\psi_{\pm}(N)\rangle$ as a function of $\hbar\omega_L$. The energies are defined with respect to the energy of the non-perturbed state $|e, N-1\rangle$, chosen to be equal to $\hbar\omega_0$, and represented by a horizontal line with ordinate $\hbar\omega_0$. Compared to this energy, state $|g, N\rangle$ has an energy equal to $\hbar\omega_L$ which varies, as a function of $\hbar\omega_L$, as a straight line of slope unity. This line intersects the horizontal line representing the energy of $|e, N-1\rangle$ at a point with abscissa $\hbar\omega_0$.

At resonance ($\omega_L = \omega_0$), the two dressed states $|\psi_{\pm}(N)\rangle$ are separated by an energy $\hbar\Omega_R$, since we assume we are in the strong coupling domain where $\Omega_R \gg \Gamma$. Let

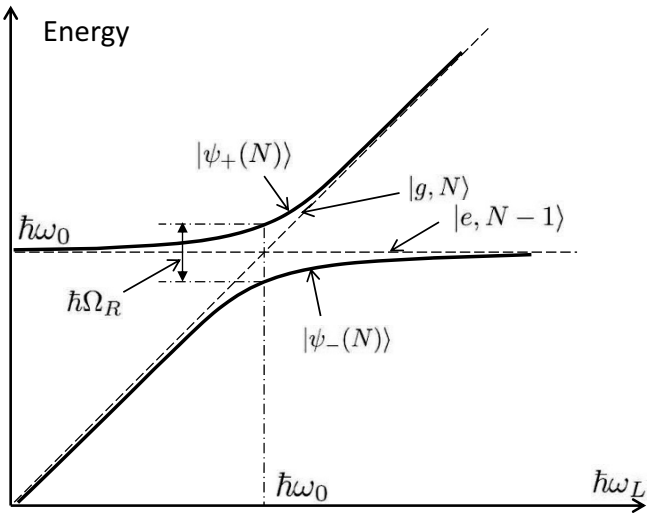


Figure 6: Energies of the dressed states $|\psi_\pm(N)\rangle$ (solid lines) and of the non-perturbed states $|g,N\rangle$ and $|e,N-1\rangle$ (dashed lines) as a function of $\hbar\omega_L$. The energies are defined with respect to the energy of the non-perturbed state $|e,N-1\rangle$, chosen to be equal to $\hbar\omega_0$. As $\hbar\omega_L$ varies, the energies of the dressed states follow a hyperbola whose asymptotes are the straight lines representing the energies of the non-perturbed states (anticrossing).

us first completely neglect Γ. Leaving resonance, and as the detuning becomes larger and larger (in absolute value), one finally reaches regions where $|\delta|$ is larger than Ω_R, which corresponds to a weak coupling regime. Varying the detuning, one then continuously goes from a strong coupling to a weak coupling region. The energies of the dressed levels $|\psi_\pm(N)\rangle$ follow a hyperbola whose asymptotes are the energies of the non-perturbed states $|g,N\rangle$ and $|e,N-1\rangle$ (Figure 6). As they come close to their asymptotes, the dressed states become very close to the non-perturbed corresponding states, and the distance between the hyperbola and its asymptote is simply the light shift δ_g defined in (27).

To take into account the natural width Γ of the excited level e, one should add a width to the dressed levels shown in Figure 6. Far away from the anticrossing center, close to the asymptotes, the width would be $\Gamma-\gamma_g$ for the dressed states that are close to the horizontal asymptote, and γ_g for the dressed states that are close to the asymptote with slope one. Following one hyperbola branch continuously, the width will progressively change from one of these values to the other, and take the value $\Gamma/4$ at the center of the anticrossing.

Another interesting phenomenon occurs when the system continuously follows one of the hyperbola branches. Imagine it follows the lower branch, from left to right, for instance because the excitation frequency is slowly varied. If the transit is slow enough to neglect any non-adiabatic transition to the other dressed state, i.e. to the other hyperbola branch, one continuously goes from state $|g,N\rangle$ to state $|e,N-1\rangle$. This is

another convenient way to go from g to e: instead of applying a resonant field during the time necessary for the Rabi oscillation to bring the system from g to e (π pulse), one slowly scans the field frequency through resonance, from a lower to a higher value. Note however that this scanning cannot be too slow, since it must occur on a time scale that is too rapid for the dissipative processes to be able to change the atomic internal state. Such a transit is often referred to as an "*adiabatic fast passage*", as it must be slow enough to remain adiabatic and fast enough to avoid any dissipation during the transit time. The dressed-atom approach allows clearly specifying the conditions for transferring the atom from one level to another.

3-c. Fluorescence triplet

With the dressed-atom approach, we can also simply explain the spectrum of the lines spontaneously emitted by an atom subjected to intense radiation. When studying elastic scattering in § E-1 of Chapter XX, we showed that, when the exciting radiation had an intensity low enough to allow a perturbation treatment, the radiation emitted spontaneously by the atom had the same frequency as the exciting radiation. We now show that the situation is different in the case of an intense excitation radiation: new frequencies appear in the light emitted by the atom[8].

We assume the exciting radiation to be resonant and intense, so that the two dressed states $|\psi_{\pm}(N)\rangle$ of multiplicity $\mathcal{E}(N)$ are separated by an energy interval $\hbar\Omega_R$ (Figure 7). These two states are linear superpositions of the states $|g, N\rangle$ and $|e, N-1\rangle$; consequently, they both have a non-zero projection onto $|e, N-1\rangle$. Similarly, the two states $|\psi_{\pm}(N-1)\rangle$ of multiplicity $\mathcal{E}(N-1)$ are linear superpositions of the states $|g, N-1\rangle$ and $|e, N-2\rangle$; they both have a non-zero projection onto $|g, N-1\rangle$. The lines emitted spontaneously by the atom are those that link two energy levels between which the atomic dipole operator $\boldsymbol{D}$ has a non-zero matrix element. Since $\boldsymbol{D}$ does not change the photon number N and can link e to g, the matrix element $\langle g, N-1|\boldsymbol{D}|e, N-1\rangle$ is non-zero; each of the two states $|\psi_{\pm}(N)\rangle$ can be linked via $\boldsymbol{D}$ to each of the two states $|\psi_{\pm}(N-1)\rangle$. The four radiative transitions represented by the curly arrows in Figure 7 are thus possible and correspond to three distinct frequencies: frequency $\omega_L + \Omega_R$ for the $|\psi_+(N)\rangle \rightarrow |\psi_-(N-1)\rangle$ transition; frequency ω_L for both the $|\psi_+(N)\rangle \rightarrow |\psi_+(N-1)\rangle$ and the $|\psi_-(N)\rangle \rightarrow |\psi_-(N-1)\rangle$ transitions; frequency $\omega_L - \Omega_R$ for the $|\psi_-(N)\rangle \rightarrow |\psi_+(N-1)\rangle$ transition. We get a frequency triplet for the spontaneously emitted light, which was first predicted by Mollow [54] by using a semiclassical treatment.

Autler-Townes doublet

Imagine that one of the two atomic states we considered until now, for example g, is connected via an allowed transition to a third state c, meaning that $\langle c|\boldsymbol{D}|g\rangle$ is non-zero. Let us also assume that the radiation frequency ω_L that is resonant for the $g \rightarrow e$ transition, is completely off-resonance for the $g \rightarrow c$ transition; consequently, it does not perturb state c, so that the sates $|c, N\rangle$ can be considered as eigenstates of the total Hamiltonian, even if they are slightly shifted. The two states $|\psi_{\pm}(N)\rangle$, which both have a non-zero projection onto $|g, N\rangle$, can therefore be connected via $\boldsymbol{D}$ to $|c, N\rangle$. This means that, because of the presence of an intense radiation exciting the $g \rightarrow e$ transition, the transition $g \rightarrow c$ is split into two lines separated by $\hbar\Omega_R$, called the "Autler-Townes doublet" [55].

[8] This is not Raman scattering as we assume there are no other atomic states except e and g.

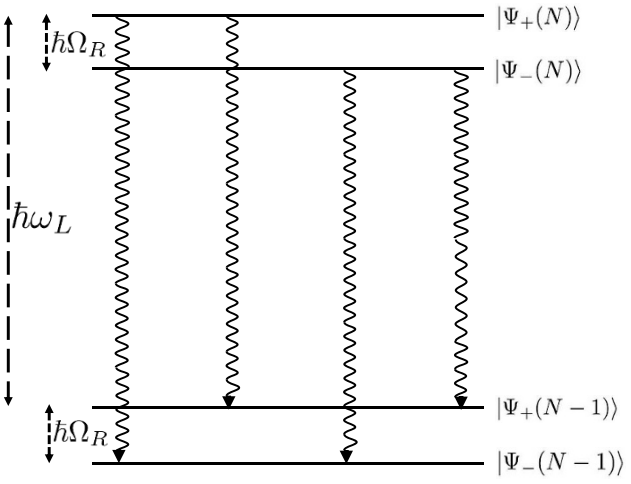


Figure 7: Radiative transitions between one of the two states $|\psi_\pm(N)\rangle$ of multiplicity $\mathcal{E}(N)$ to one of the two states $|\psi_\pm(N-1)\rangle$ of multiplicity $\mathcal{E}(N-1)$, whose energy is lower than that of $\mathcal{E}(N)$ by the quantity $\hbar\omega_L$. We assume the exciting radiation to be exactly at resonance, so that the energy interval between the two states $|\psi_\pm\rangle$ of each multiplicity is equal to $\hbar\Omega_R$.

3-d. Temporal correlations between fluorescent photons

We now study the characteristics of the radiation spontaneously emitted by an atom that interacts continuously with the electromagnetic field of a laser.

α. *Radiative cascade of the dressed atom*

We saw that a dressed atom, spontaneously emitting a photon, goes from the multiplicity $\mathcal{E}(N)$ to the one just below, $\mathcal{E}(N-1)$, located at an energy distance $\hbar\omega_L$. We shall not study here the precise evolution of the physical system as it leaves multiplicity $\mathcal{E}(N)$, which requires the master equation, already mentioned in § 1-e. Our discussion will remain qualitative, but the interested reader will find a more detailed approach in § D of Chapter VI in [21].

Once it reaches $\mathcal{E}(N-1)$, the atom can spontaneously emit a new photon, which brings the dressed atom to $\mathcal{E}(N-2)$, and so on. The series of photons spontaneously emitted by the atom in continuous interaction with the laser radiation can be viewed as a "radiative cascade" of the dressed atom descending its energy diagram.

This image of a radiative cascade permits studying the time correlations between the photons emitted by the atom. As we shall see, the observed correlations depend on the spectral resolution of the photodetectors used.

β. High spectral resolution photodetection

With a high enough spectral resolution, one can observe the time-correlations between photons emitted in the two side-bands of the triplet. Suppose we place filters in front of the photodetectors, so that each can receive only one of the components of the fluorescent triplet, centered at frequencies ω_L, $\omega_L + \Omega_R$ and $\omega_L - \Omega_R$. This means that the spectral resolution of the apparatus is better than the splitting frequency Ω_R of this triplet, but it does not imply that it is lower than the natural width Γ of each components. If we call $\delta\omega$ this spectral resolution, we then have:

$$\Gamma < \delta\omega < \Omega_R \tag{41}$$

Imagine that at a given time, a detector registers a photon emitted for example in the lateral band centered at $\omega_L + \Omega_R$, as the system undergoes the transition $|\Psi_+(N)\rangle \rightarrow |\Psi_-(N-1)\rangle$ (curly arrow on the left-hand side of the Figure 7). The next photon is emitted as the system, starting from $|\Psi_-(N-1)\rangle$, undergoes either the $|\Psi_-(N-1)\rangle \rightarrow |\Psi_-(N-2)\rangle$ transition, emitting a photon of frequency ω_L, or the $|\Psi_-(N-1)\rangle \rightarrow |\Psi_+(N-2)\rangle$ transition, emitting a photon of frequency $\omega_L - \Omega_R$. This means that a second photon with the same frequency $\omega_L + \Omega_R$ as the first one, cannot be emitted right after the first one.

If the frequency of that second photon is ω_L, the system ends up in state $|\Psi_-(N-2)\rangle$; from that state, it can emit either a third photon with frequency ω_L, or a third photon with a lower frequency $\omega_L - \Omega_R$. If, on the other hand, the frequency of that second photon is $\omega_L - \Omega_R$, the system ends up in state $|\Psi_+(N-2)\rangle$; from that state, it can emit either a photon with a higher frequency $\omega_L + \Omega_R$, or a photon of frequency ω_L. As opposed to the second photon, the third photon may thus have the same frequency $\omega_L + \Omega_R$ as the first one. Following the same line of reasoning one can argue that, if the first photon has a frequency $\omega_L - \Omega_R$, this cannot be the case for the second photon; one must wait until the third photon to eventually obtain the same frequency $\omega_L - \Omega_R$. This means that, if photons with only the two extreme frequencies $\omega_L \pm \Omega_R$ are selectively observed, the detected emission processes will necessarily alternate in time (but these events may be separated by any number of photon emissions at the central frequency ω_L).

Comment:

Taking a Fourier transform to return to the time domain, relation (41) imposes a limit to the temporal resolution δt of the detection system: $\delta t \geq 1/\Omega_R$. This means that it is not possible to measure the exact time at which a photon is emitted with a precision of the order of the Rabi precession period.

γ. Photodetection with high temporal resolution

We now study the opposite case where the detectors have a temporal resolution better than the Rabi precession period. This allows a precise determination of the time at which the photon is emitted, but one can no longer distinguish the frequencies of the three triplet components.

We have seen above (§ 1-e) that the elementary spontaneous emission processes have a correlation time that is very short compared to all the other characteristic times

of the problem (because of the large spectral width $\Delta\omega$ of the empty modes' reservoir). A spontaneous emission from $\mathcal{E}(N)$ thus corresponds to a very short "quantum jump" taking the system from state $|e, N-1\rangle$ in $\mathcal{E}(N)$ into state $|g, N-1\rangle$ in $\mathcal{E}(N-1)$. Once the atom has reached this second state, it cannot emit a second photon right away, since no spontaneous emission can occur from a ground state g. A certain time must elapse for the atom-laser interaction to bring the system from state $|g, N-1\rangle$ to the state $|e, N-2\rangle$ it is coupled with, and from which another photon can be spontaneously emitted. The system then falls back to state $|g, N-2\rangle$, and the previous process repeats itself (with an N value lowered by one unit). It therefore becomes clear why one observes a "temporal antibunching" of the photons emitted by a single atom, as they must be separated by a time interval at least of the order of $1/\Omega_R$; this antibunching was already referred to in § 3-b-γ of Complement B_{XX}.

4. Modifications of the field. Dispersion and absorption

We now study how the field is modified by its interaction with the atom.

4-a. Atom in a cavity

The atom-field interaction does not solely perturb the atom; it also changes the field. In order to study this effect, it is convenient to imagine the atom being placed in a real cavity assumed to be perfect, meaning that its losses can be ignored (they occur on a time scale much longer than all the other relevant times of the experiment). As opposed to what we did before, we shall keep the N dependence of the Rabi frequencies Ω_N given by equations (8) and (9), since in a cavity the photon number N has a physical meaning (§ 1-d).

Figure 8 shows the first multiplicities $\mathcal{E}(N)$ of the system atom + field for low and increasing values of the photon number N, starting at $N = 0$. Multiplicity $\mathcal{E}(0)$

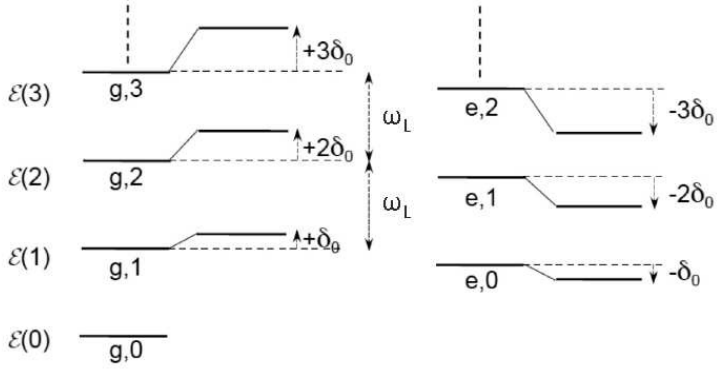


Figure 8: Energy levels of the system atom + field for low values of the photon number N (in angular frequency units, meaning the energies are divided by $\hbar$). States $|g, N\rangle$ and $|e, N-1\rangle$ of $\mathcal{E}(N)$ undergo opposite shifts, proportional to N. State $|g, 0\rangle$ is not shifted.

contains a single state $|g,0\rangle$. Multiplicity $\mathcal{E}(1)$ contains the two states $|g,1\rangle$ and $|e,0\rangle$. Multiplicity $\mathcal{E}(2)$ contains the two states $|g,2\rangle$ and $|e,1\rangle$, and so on.

We shall assume we are in the weak coupling regime, so that we can use the perturbative results of § 2 for the light shifts of the different energy levels. State $|g,0\rangle$ is not shifted as it is not coupled to any other state[9]. States $|g,1\rangle$ and $|e,0\rangle$ of $\mathcal{E}(1)$ undergo opposite light shifts, respectively $+\hbar\delta_g$ and $-\hbar\delta_g$, where δ_g is given by equation (27), where we have replaced Ω_R by Ω_1, the Rabi frequency for $N=1$ (see (8)). Setting:

$$\delta_0 = \Omega_1^2 \frac{\delta}{4\delta^2+\Gamma^2} \tag{42}$$

the light shifts of states $|g,1\rangle$ and $|e,0\rangle$ are, respectively, $+\hbar\delta_0$ and $-\hbar\delta_0$. According to (8), the squares of the Rabi frequencies Ω_N characterizing the atom-field coupling in the multiplicities $\mathcal{E}(N)$, are proportional to N; this means that states $|g,2\rangle$ and $|e,1\rangle$ of $\mathcal{E}(2)$ undergo light shifts respectively equal to $+2\hbar\delta_0$ and $-2\hbar\delta_0$. More generally, states $|g,N\rangle$ and $|e,N-1\rangle$ of $\mathcal{E}(N)$ undergo shifts respectively equal to $+N\hbar\delta_0$ and $-N\hbar\delta_0$.

A similar reasoning can be applied to the radiative broadening. It shows that the radiative broadenings of states $|g,N\rangle$ and $|e,N-1\rangle$ of $\mathcal{E}(N)$ are respectively equal[10] to $+N\gamma_0$ and $\Gamma - N\gamma_0$, where γ_0 is given, according to (27), by:

$$\gamma_0 = \Omega_1^2 \frac{\Gamma}{4\delta^2+\Gamma^2} \tag{43}$$

4-b. Frequency shift of the field in the presence of the atom

Consider the left column in Figure 8. The gap between the perturbed energies of states $|g,1\rangle$ and $|g,0\rangle$ is equal to $\hbar(\omega_L+\delta_0)$; the gap between the perturbed energies of states $|g,2\rangle$ and $|g,1\rangle$ is equal to $\hbar(\omega_L+2\delta_0-\delta_0)=\hbar(\omega_L+\delta_0)$, and so on. As the light shifts of the states $|g,N\rangle$ are proportional to $N\delta_0$, increasing linearly with N, the perturbed levels in the left column of Figure 8 have a constant gap between them, even in the presence of the coupling; the distance between consecutive levels simply goes from $\hbar\omega_L$ to $\hbar(\omega_L+\delta_0)$. In other words, the presence in the cavity of an atom in its ground state g changes the field frequency from ω_L to $\omega_L+\delta_0$. As the light shifts of the levels in the right column of Figure 8 have an opposite sign, a similar argument shows that the presence in the cavity of an atom in the excited state e changes the field frequency from ω_L to $\omega_L-\delta_0$.

The atom-field interaction thus shifts the field frequency inside the cavity by a quantity that changes sign, depending on whether the atom is in the internal state g or e. Let us assume this interaction lasts a time T, as will be the case if an atom is introduced into the cavity and takes that time to traverse it. Compared to the free oscillation in the absence of the atom, the field oscillation will be out of phase by an amount ϕ; this phase shift is equal to $\phi=+\delta_0 T$ if the atom is in state g, and to $\phi=-\delta_0 T$ if the atom is in state e.

This change in the field frequency is a phenomenon similar to that described by the real part of the refractive index: a light beam going through an atomic media changes

[9] There is actually a coupling between state $|g,0\rangle$ and state $|e,1\rangle$, but is highly non-resonant; we shall ignore it since, as stipulated above, our computation is to zeroth order in Ω_R/ω_L.

[10] State $|g,0\rangle$ does not undergo any radiative broadening.

its propagation velocity without any changes in its frequency. In a cavity, the field wavelength cannot change as it is fixed by the boundary conditions on the cavity walls, and hence by the cavity size. The phase shift compared to the free evolution cannot accumulate in space but must accumulate in time (resulting in a frequency change of the field). Note that if one varies the ω_L frequency of the field around the atomic frequency ω_0, the sign change of the light shift δ_0 is reminiscent of the sign change of the real part of the refractive index in the vicinity of an atomic resonance[11].

4-c. Field absorption

Consider an atom in its ground state g in the presence, at time $t = 0$, of an ω_L mode of the field in a quasi-classical coherent state $|\alpha\rangle$ (Complement G_V). The state of the total system reads:

$$|\psi(0)\rangle = |g\rangle \otimes |\alpha\rangle = \sum_{N=0}^{\infty} \frac{\alpha^N}{\sqrt{N!}}\, e^{-|\alpha|^2/2}\, |g, N\rangle. \tag{44}$$

The time evolution in the presence of coupling changes the energies of the states $|g, N\rangle$ to:

$$\tilde{E}_N = N\hbar(\omega_L + \delta_0 - i\gamma_0/2). \tag{45}$$

and the state of the system at time t becomes:

$$\begin{aligned} |\psi(t)\rangle &= \sum_{N=0}^{\infty} \frac{\alpha^N}{\sqrt{N!}} e^{-|\alpha|^2/2} \exp[-iN(\omega_L + \delta_0 - i\gamma_0/2)t]\ |g, N\rangle \\ &\propto |g\rangle \otimes |\alpha \exp[-i(\omega_L + \delta_0 - i\gamma_0/2)t]\rangle. \end{aligned} \tag{46}$$

The atom is still in the presence of a quasi-classical coherent state. However, compared to the free field evolution of that state in the absence of coupling, the atom-field interaction has introduced a phase shift $\delta_0 t$ (as already discussed above) as well as a decrease in amplitude $e^{-\gamma_0 t/2}$, resulting in an attenuation of the amplitude of the field. This is reminiscent of the radiation absorption described by the imaginary part of the refractive index.

To sum up, we showed that the atom-field coupling produces light shifts and radiative broadening of the atomic levels, corresponding to the well-know field dispersion and absorption phenomena in optics.

Conclusion.

In conclusion, we showed for many various situations that the dressed-atom approach brings strong clarifications while keeping the calculations simple. Considering the atom and the field mode with which it interacts as a quantum system described by a time-independent Hamiltonian allows introducing true energy levels for the global system; this leads to a new, broad overview of the stimulated absorption and emission of photons.

As an example, this approach makes it very clear how the atom-photon coupling changes the energy diagram of the dressed atom at high field intensity; this leads to a

[11] This effect is sometimes referred to as "*anomalous dispersion*."

very simple interpretation of the new frequencies appearing in the atomic fluorescence spectrum in the strong coupling domain. As the energy diagram of the dressed atom is a succession of multiplicities separated by an energy equal to a photon energy, the spontaneous emission of a photon is viewed in this approach as a quantum jump of the dressed atom from one multiplicity to the one just below (radiative cascade). This approach allows a simple calculation of the delay function yielding the distribution of the time intervals between two successive quantum jumps; this permits studying the time correlations between fluorescent photons. Let us also mention that this delay function allows simulating the temporal evolution of an atom, hence obtaining individual quantum trajectories, which can be used to get an averaged atomic evolution. Several experimental applications of the dressed-atom method are presented in the next complement.

Complement D_{XX}

Light shifts: a tool for manipulating atoms and fields

1 Dipole forces and laser trapping 2151
2 Mirrors for atoms 2153
3 Optical lattices 2153
4 Sub-Doppler cooling. Sisyphus effect 2155
4-a Laser configurations with space-dependent polarization . . . 2156
4-b Atomic transition 2156
4-c Light shifts 2156
4-d Optical pumping 2157
4-e Sisyphus effect 2158
5 Non-destructive detection of a photon 2159

Light shifts studied in § 2-b of Complement C_{XX} exhibit a number of important properties leading to numerous applications; these will be briefly discussed in this complement.

As these light shifts are proportional to the laser intensity, their magnitude can be space-dependent if the laser intensity is not homogeneous in space. These shifts can be used to create either potential wells (§ 1) to trap atoms once they are cold enough (laser trapping), or potential barriers (§ 2) reflecting atoms (mirrors for atoms). A particularly interesting example involves periodic optical potential wells created at the nodes and antinodes of a laser standing wave in an off-resonant condition (§ 3). This situation is reminiscent of that encountered by electrons trapped in the periodic potential of a crystal lattice. Neutral atoms trapped in optical lattices can thus serve as models for condensed matter problems.

For low enough values of the detuning δ between the laser frequency and the atomic frequency, and if the ground state has several Zeeman sublevels, non-dissipative effects, such as light shifts, can coexist with dissipative effects, such as optical pumping between Zeeman sublevels. We explain in § 4 how correlations between these two types of effects can lead to new cooling mechanisms, such as Sisyphus cooling, allowing the atoms to reach temperatures much lower than with Doppler cooling.

Finally, we show in § 5 how the light shifts undergone by an atom crossing a highly detuned cavity allows determining the number of photons present in the cavity, by performing measurements on the atoms at the cavity exit, without absorbing any of the cavity photons.

1. Dipole forces and laser trapping

When the light intensity varies in space, as with a focalized laser beam or a standing wave, the light shifts also become space-dependent. If the detuning δ between the laser frequency and the atomic frequency is large compared to the natural width Γ of the excited level, it is then justified to ignore the dissipation due to spontaneous emission, on

the characteristic time scales of the experiment. The light shift $\hbar\delta_g(\boldsymbol{R})$ of ground state g depends, as does the light intensity, on the position $\boldsymbol{R}$ of the atomic center of mass; it can therefore be considered as a potential energy $V(\boldsymbol{R}) = \hbar\delta_g(\boldsymbol{R})$ that affects the atomic motion. This potential has the same sign as the light shift, and hence depends on the sign of the frequency detuning δ.

The potential $V(\mathbf{r})$ gives rise to a force:

$$\mathbf{F}_{\text{dip}}(\mathbf{R}) = -\boldsymbol{\nabla}_{\mathbf{R}} V(\mathbf{R}) \tag{1}$$

called the "dipole force", or sometimes the "reactive force" (§ 11-4 in [24]). It is different from the radiation pressure forces studied in § 1-d of Complement A_{XIX}, which come from momentum exchanges as the atom absorbs photons that are spontaneously reemitted. The dipole forces introduced here arise from the spatial variations of the light shifts undergone by the dressed-atom levels. One could say they are caused by the redistribution of photons between the different plane waves composing the laser beam[1]: the atom absorbs a photon from one plane wave and re-emits it, by stimulated emission, in another plane wave; this process changes the atom's momentum, and hence giving rise to a force.

Comment:

As is the case for light shifts, the intensity of the dipole forces, as a function of the frequency detuning δ between the laser frequency and the atomic frequency, follows a dispersion curve. In addition, the light shifts of the two dressed levels in multiplicity $\mathcal{E}(N)$ have an opposite sign for a given detuning δ; the dipole force thus changes sign from one dressed state $|\Psi_+(N)\rangle$ to its associated state $|\Psi_-(N)\rangle$. When the detuning δ is not too large, and if spontaneous emission processes can occur, the dressed-atom radiative cascade can lead to sign changes of the dipole force, as the atom goes from states $|\Psi_\pm(N)\rangle$ to $|\Psi_\mp(N-1)\rangle$; this is the origin of the fluctuations of the dipole forces.

An important application of dipole forces is the implementation of laser traps. Consider first a laser beam detuned toward the red ($\omega_L < \omega_0$) and focalized at point O. The light shift, zero outside the laser beam, is negative inside the laser beam; it increases in absolute value as one gets closer to the focal point, where it reaches its maximum value. This creates a potential well that could trap a neutral atom; this will indeed happen if the atom's kinetic energy, of the order of $k_B T$, is lower than the depth U_0 of the potential well. This is why these laser traps have been built only since the 1980's, once atomic cooling techniques (Complement A_{XIX}, § 2) allowed slowing down atoms to temperatures of the order of a microkelvin [56].

Comment:

The trapping forces involved in laser traps are of the order of an atomic dipole multiplied by the gradient of a laser field. They are much weaker than the forces exerted by a static electric field on a charged particle. This explains why laser traps for neutral atoms are much shallower than ion or electron traps. There exist, however, other types of traps for

[1] A single plane wave does not have an intensity gradient, and cannot exert a dipole force. These forces, due to intensity gradients, require the presence of several plane waves with different wave vectors.

neutral atoms, using different physical mechanisms (for a short review, see for example § 2-c of Complement A_{XIX}, and Chapter 14 of reference [24]).

2. Mirrors for atoms

A laser detuned toward the blue ($\omega_L > \omega_0$) gives rise to repulsive potentials. Imagine for example that the laser wave propagates within a bloc of glass, and undergoes total internal reflection at the boundary between the glass and the vacuum (Figure 1-a). An evanescent wave appears outside the glass, with an amplitude decaying exponentially in a direction perpendicular to the boundary, becoming negligible over a distance of the order of the laser wavelength. This evanescent wave creates a potential barrier of height U_0, which reflects atoms[2] arriving with an energy $E < U_0$ (Figure 1-b). This set-up can be used as a mirror for neutral atoms [57].

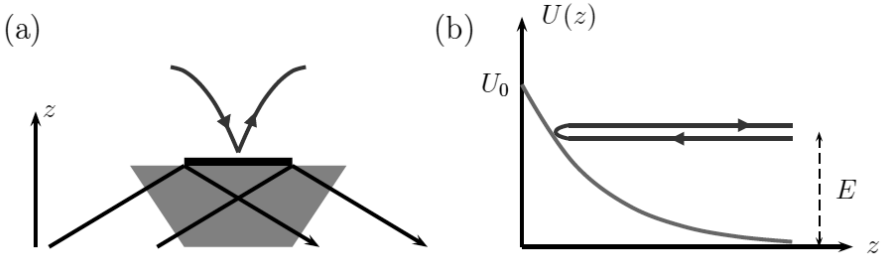


Figure 1: (a) A laser beam traveling within a block of glass (shaded in grey in the figure) undergoes total internal reflection at the boundary. Outside the glass, an evanescent wave appears. (b) If the laser is detuned toward the blue ($\omega_L > \omega_0$), this evanescent wave creates a potential barrier of height U_0. Atoms falling on this barrier with an energy E lower than U_0 are reflected by the barrier and turn around.

3. Optical lattices

When laser beams form a standing wave, the light intensity is modulated in space, with a periodicity $\lambda/2$: the intensity is zero at the nodes, and maximal at the antinodes. This creates periodic potential wells, located at the antinodes of the wave for a negative detuning ($\omega_L < \omega_0$), and at the nodes for a positive detuning ($\omega_L > \omega_0$). Figure 2 shows a two-dimensional optical lattice created by two standing waves, along two orthogonal axes[3].

The study of optical lattices is interesting for several reasons, in particular because the motion of a neutral atom in an optical lattice is reminiscent of that of an electron in

[2] Atoms falling on a solid surface would stick to it, rather than being reflected.

[3] The frequencies ω_{L1} and ω_{L2} of the two standing waves are in general sufficiently far apart for the interference terms between the two waves to have a negligible effect on the atom's motion; the potentials created by the two waves can then be independently added. This requires $|\omega_{L1} - \omega_{L2}|$ to be large compared to all the characteristic frequencies of the atom's motion, such as it's vibrational motion inside a well; in that case, the interference terms oscillate too fast to have a significant effect.

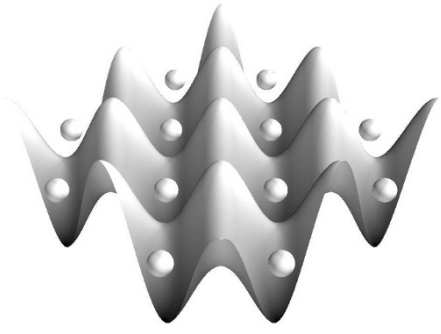


Figure 2: Schematic representation of a two-dimensional optical lattice: placed in a superposition of two standing laser waves along two orthogonal axes, the atom is subjected to a potential periodic in space, represented by the undulating surface in the figure. These periodic potential wells, located at the antinodes of the standing waves for $\omega_L < \omega_0$, and at the nodes for $\omega_L > \omega_0$, form an optical lattice. The spheres above the surface indicate the positions where the atoms can be trapped.

a crystal lattice. Granted the order of magnitudes involved are quite different, since the spatial period of an optical lattice is of the order of a micron, whereas the period of a crystal lattice is of the order of a fraction of a nanometer. Nevertheless, optical lattices offer a large number of possibilities not available to crystalline lattices:

- One can easily change the intensity of the laser waves forming the standing waves, hence modifying the depth of the potential wells; this allows controlling the tunnel effect between adjacent wells. This method was used to explore the transition between a deep well regime where the atoms are localized at the bottom of the wells, and a shallow well regime where the atoms' wave functions are delocalized over the entire lattice [58].

- One can abruptly switch off the trapping laser beams (which obviously cannot be done for a crystal lattice) and study the resulting behavior of the liberated atoms. Studying the expansion velocity of the clouds of atoms yields information on their velocity distribution, and hence on their temperature (time-of-flight method). Studying their spatial distribution and the possible appearance of a diffraction pattern allows determining whether the matter waves trapped in distinct potential wells of the optical lattice were coherent or not.

- One can use two different frequencies ω_{L1} and ω_{L2} for the two laser waves counterpropagating to form the one-dimensional standing laser wave. This leads to a "standing" laser wave, moving with constant velocity if $\omega_{L1} - \omega_{L2}$ is fixed, or with an acceleration if $\omega_{L1} - \omega_{L2}$ increases linearly with time. In this latter case, the atom experiences a constant inertial force in the rest frame of the standing wave; its motion is then similar to that of an electron in a crystal lattice periodic potential, subjected in addition to a static electric field. The motion of such a particle, in a periodic lattice and subjected to a constant force, is predicted to be oscillatory,

following the so-called *Bloch oscillations*; the experimental observation of such oscillations is facilitated in an optical lattice, as the atom's relaxation time can be much longer than the oscillations' period [59].

Cold atoms trapped in an optical lattice are a model system for "simulating" a number of situations encountered in solid state physics. Cold atom studies involve interactions between atoms much weaker than the Coulomb interactions between electrons. Furthermore, they can be controlled thanks to resonance effects occurring as atoms collide with each other.

Note finally that optical lattices are a good example highlighting the importance of light shifts. One may wonder if it might not be simpler to shift atomic levels by Zeeman or Stark effects in static magnetic or electric fields, rather than using an off-resonance light beam to produce a light shift. The advantage of the light shifts is that they can be used to form potentials varying over very short distances, of the order of an optical wavelength, which is much more difficult to attain with static fields.

4. Sub-Doppler cooling. Sisyphus effect

We described, in § 2-b of Complement A_{XIX}, a cooling mechanism for the atoms, based on the Doppler effect, and called for that reason "Doppler cooling". We computed the friction and diffusion coefficients associated with that mechanism and showed that the lowest temperature T_D that could be reached by Doppler cooling was of the order of $\hbar\Gamma/k_B$ (where Γ is the natural width of the atoms' excited states, and k_B the Boltzmann constant). Actually, the first measurements of the temperatures reached by laser cooling, and based on the time-of-flight method [60], showed that temperatures much lower than T_D could be obtained; furthermore, their dependence on the detuning δ between the laser beams' frequency and the atomic frequency did not follow the prediction of the Doppler cooling theory. This implied the existence of other cooling mechanisms for the atoms, leading to temperatures lower than the Doppler limit T_D; as expected, these mechanisms were called "sub-Doppler" mechanisms. One of them, called the "Sisyphus effect", will be described in this section.

The theory of Doppler laser cooling, exposed in § 2-b of Complement A_{XIX}, does not take into account several important characteristics of laser cooling experiments.

- In most experiments performed in three-dimensional space, the polarization of the laser field cannot be uniform. The spatial variations of this polarization should not be ignored.

- The atoms under study have several sublevels, in the lower g state and in the excited e state. The two-level atom approximation of § 2-b in Complement A_{XIX} is therefore not sufficient.

- As there are several sublevels in the lower state g, one should include the effects of the optical pumping between these sublevels, effects whose characteristic time constants (pumping times) are longer than the lifetime $1/\Gamma$ of the excited state.

- As the detuning δ between the laser beams' frequency and the atomic frequency is different from zero, one must take into account the light shifts of the lower level g, which can take on different values for the different sublevels.

Before describing the Sisyphus effect, we first show on a simple example how these different effects come into play.

4-a. Laser configurations with space-dependent polarization

Laser configurations with a space-dependent polarization do not necessarily involve three pairs of laser beams counterpropagating along the Ox, Oy and Oz axes. They can be achieved in one-dimension, and are easier to study, as long as the two counterpropagating laser waves have different polarizations. As an example, Figure 3 represents two laser waves propagating in opposite directions along the Oz axis and having linear orthogonal polarizations $\mathbf{e}_x$ and $\mathbf{e}_y$. The polarization of the total field changes from right-hand circularly polarized (σ_+ with respect to the quantization axis Oz) to left-hand circularly polarized (σ_-) in planes separated by a distance $\lambda/4$, and is linear at $\pm 45°$ of the Ox and Oy axes, half-way between these planes.

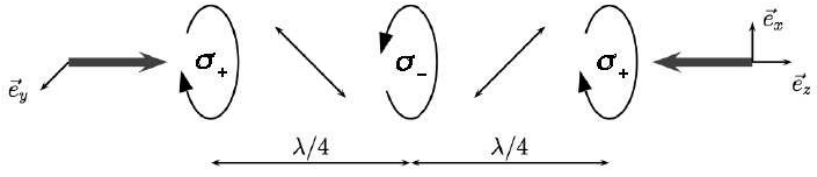


Figure 3: Laser configuration with a space-dependent polarization: two laser waves propagate in opposite directions along the Oz axis, having linear orthogonal polarizations $\mathbf{e}_x$ *and* $\mathbf{e}_y$.

4-b. Atomic transition

Many of the laser cooling experiments use transitions between a lower state g with angular momentum J_g and an excited state e with an angular momentum equal to $J_e = J_g + 1$. Here, we shall consider the simplest possible case $J_g = 1/2$, where the lower state contains only 2 sublevels $g_{\pm 1/2}$. We then have $J_e = 3/2$ and the excited state has 4 sublevels $e_{\pm 1/2}$ and $e_{\pm 3/2}$.

4-c. Light shifts

Consider first a point in space where the laser field polarization is σ_+ (with respect to the quantization axis Oz). We saw in § 1-b of Complement C_{XIX} that photons with a σ_+ (σ_-) polarization have a spin angular momentum $+\hbar$ ($-\hbar$) along the Oz axis. Conservation of total angular momentum in the photon absorption process leads to the selection rule $m_e - m_g = +1(-1)$ for the absorption of a $\sigma_+(\sigma_-)$ photon, where m_e and m_g are the magnetic quantum number of the states involved in the transition. Figure 4 represents the 2 transitions $g_{-1/2} \leftrightarrow e_{+1/2}$ and $g_{+1/2} \leftrightarrow e_{+3/2}$ ($m_e - m_g = +1$) that can be excited by the laser field. The numbers 1/3 and 1 shown next to these 2 transitions are the squares of the Clebsch-Gordan coefficients of these transitions (Complement B_X); they indicate that the $g_{+1/2} \rightarrow e_{+3/2}$ transition is 3 times more intense than the $g_{-1/2} \leftrightarrow$

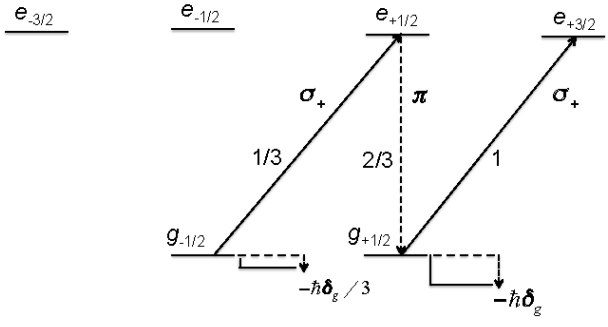


Figure 4: Transition $1/2 \leftrightarrow 3/2$. *The oblique upwards arrows show the transitions excited at a point where the laser field polarization is* σ_+, *the vertical downward arrow indicates the spontaneous emission transition from sublevel* $e_{+1/2}$ *toward sublevel* $g_{+1/2}$. *The light shifts of states* $g_{-1/2}$ *and* $g_{+1/2}$ *are noted* $-\hbar\delta_g/3$ *and* $-\hbar\delta_g$. *At a point where the laser polarization becomes* σ_- *instead of* σ_+, *the shifts of the two sublevels are interchanged, by symmetry.*

$e_{+1/2}$ transition. As the detuning δ between the laser beams' frequency and the atomic frequency is negative in a laser cooling experiment, both states $g_{\pm 1/2}$ have a negative light shift, but with a modulus 3 times larger for state $g_{+1/2}$ than for state $g_{-1/2}$. These light shifts are written in the figure as $-\hbar\delta_g$ and $-\hbar\delta_g/3$, where δ_g is positive.

At a point in space where the polarization is σ_-, the previous results are interchanged. It is now the $g_{-1/2} \leftrightarrow e_{-3/2}$ transition that is 3 times more intense than the $g_{+1/2} \leftrightarrow e_{-1/2}$ transition, yielding light shifts equal to $-\hbar\delta_g$ and $-\hbar\delta_g/3$ for the states $g_{-1/2}$ and $g_{+1/2}$, respectively.

Finally, at a point where the polarization is linear, the two light shifts are identical for symmetry reasons, and proportional to the square of the Clebsch-Gordan coefficient (equal to 2/3), indicated in the figure for the $g_{+1/2} \leftrightarrow e_{+1/2}$ transition. Consequently, they are both equal to $-2\hbar\delta_g/3$.

This means that as one moves along z axis, the positions of the 2 Zeeman sublevels $g_{-1/2}$ and $g_{+1/2}$ oscillate, with opposing phases, between the values $-\hbar\delta_g$ and $-\hbar\delta_g/3$ (taking the energy of the unperturbed ground state equal to zero).

4-d. Optical pumping

Let us focus on a point where the laser field polarization is σ_+ and there is an atom in state $g_{+1/2}$. The atom can absorb a σ_+ photon and end up in state $e_{+3/2}$. From this state, it can only fall, by spontaneous emission, back to its initial state $g_{+1/2}$; optical pumping (§ 1-b of Complement C_{XIX}) does not lead, in this case, to any population change. On the other hand, if the atom is initially in state $g_{-1/2}$ and absorbs a photon σ_+ that brings it to state $e_{+1/2}$, it can then fall back, by spontaneous emission, into state $g_{+1/2}$; optical pumping takes place from the least shifted sublevel $g_{-1/2}$ towards the most shifted sublevel $g_{+1/2}$. A comparable situation is found at a point where the laser field polarization is σ_-. Optical pumping can only occur from the least shifted sublevel $g_{+1/2}$ toward the most shifted sublevel $g_{-1/2}$. As for a point where the laser field polarization

is linear, since the Clebsch-Gordan coefficients of the $g_{-1/2} \leftrightarrow e_{-1/2}$ and $g_{+1/2} \leftrightarrow e_{+1/2}$ transitions are equal, as are those of the $e_{-1/2} \leftrightarrow g_{+1/2}$ and $e_{+1/2} \leftrightarrow g_{-1/2}$ transitions, optical pumping cannot favor one of the the populations of the 2 sublevels $g_{-1/2}$ and $g_{+1/2}$. To sum up, optical pumping can only transfer population from the least shifted sublevel to the most shifted sublevel, with a maximum efficiency at points where the laser field polarization is circular.

4-e. Sisyphus effect

We now show how the correlations between the light shifts and the optical pumping effects studied in the last two sections can reduce the atom's kinetic energy, and hence cool it down.

Figure 5 shows, for an atom moving along the z axis, the energies of its 2 sublevels $g_{-1/2}$ and $g_{+1/2}$, shifted by the light. Let us assume the atom starts from the bottom of a potential valley, at a point where the laser field polarization is σ_+, and is initially in its most shifted state $g_{+1/2}$. As it moves toward the right, it climbs a potential well, and looses some kinetic energy. If the optical pumping time is long enough, it will have time to reach the top of the hill, where the laser field polarization is σ_-; it then has a high probability to undergo an optical pumping cycle and be transferred to the most shifted sublevel, which is now sublevel $g_{-1/2}$. The whole cycle we just described can repeat itself, and each time the kinetic energy of the atom is lowered by a quantity of the order of the maximum energy difference between the two sublevels in Figure 5, equal to $(2/3)\hbar\delta_g$. The atom is facing a situation similar to that of the hero of Greek mythology, Sisyphus: it must endlessly climb a potential hill since it is sent back to the bottom as soon as it reaches the top, hence the name Sisyphus effect given to this mechanism.

The temperature reached by such a mechanism can be estimated by a simple

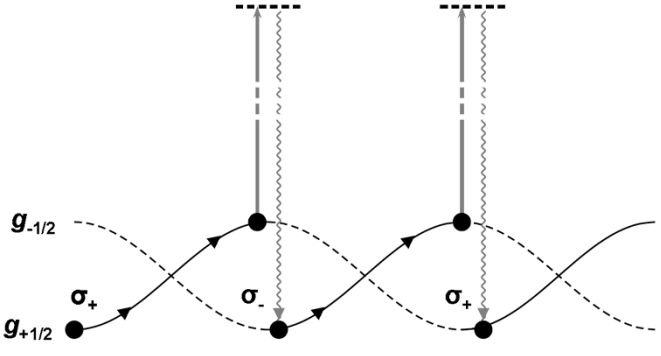


Figure 5: Principle of Sisyphus cooling: an atom in state $g_{+1/2}$ moving from a point where the laser field polarization is σ_+ must climb a potential hill of height $2\hbar\delta_g/3$, which decreases its kinetic energy. When it reaches the top of the hill, where the laser field polarization is σ_-, it has a strong probability to fall back, by spontaneous emission, to the state $g_{-1/2}$. As the cycle repeats itself, the atom is for ever climbing potential hills, like the hero Sisyphus in Greek mythology. Its kinetic energy diminishes constantly.

argument. The atom's kinetic energy decreases during each Sisyphus cycle, until it is low enough for the atom to be trapped at the bottom of a potential well. At that point its kinetic energy must be of the order of $\hbar\delta_g$. The ultimate temperature T_S that can be reached by Sisyphus cooling is expected to be $T_S \simeq \hbar\delta_g/k_B$. In laser cooling experiments, the laser intensities are generally low and the atomic transitions are not saturated; consequently, $\hbar\delta_g \ll \hbar\Gamma$ and hence $T_S \ll T_D$. This explains why the measured temperatures can be two orders of magnitude lower than the Doppler temperature, and reach values of the order of $10^{-6}K$, opening the way to numerous applications.

All these qualitative predictions have been confirmed by more quantitative models, see ([61]) and ([62]). Experiments have confirmed the theoretical predictions, in particular those concerning the dependence of T_S on the various experimental parameters, such as laser intensity and detuning [63].

5. Non-destructive detection of a photon

Consider now an experiment where the atoms of a beam cross, one after the other, a cavity containing radiation whose quantum state is described by a Fock state; the number of photons in the cavity mode is fixed, equal for example to 0 (radiation vacuum) or 1 (single photon). The atoms are prepared in a coherent superposition of the two states g and e:

$$|\psi_{\text{in}}\rangle = \frac{1}{\sqrt{2}}(|g\rangle + |e\rangle) \tag{2}$$

While each atom interacts with the photons, its levels undergo light shifts resulting in different phases for the two atomic states as the atom crosses the cavity; note, however, that if the detuning δ between the laser frequency and the atomic frequency is sufficiently large, no photon will be absorbed or emitted. As the atom exits the cavity, the radiation state is the same initial Fock state, whereas the atomic state is modified by this phase factor. The final atomic state can be written (within a global phase factor of no physical significance):

$$|\psi_{\text{fin}}\rangle = \frac{1}{\sqrt{2}}(|g\rangle + e^{-i\phi}|e\rangle) \tag{3}$$

The phase ϕ is simply the integral over time of the energy difference between the dressed-atom levels that come into play as the atom crosses the cavity. It is given by the energy diagram of the dressed-atom.

Figure 8 of Complement C_{XX} shows that the gap between states $|e\,,0\rangle$ and $|g\,,0\rangle$ is reduced from $\hbar\omega_0$ to $\hbar(\omega_0 - \delta_0)$ by the light shifts. For a cavity with no photons ($N = 0$), when the atom exits the cavity, the coherence between its states e and g has been dephased by:

$$\phi_0 = -\int \delta_0(t')\ \mathrm{d}t' \tag{4}$$

where $\delta_0(t')$ is obtained by replacing in (42) of Complement C_{XX} the Rabi frequency Ω_1 by a function of time that accounts for the motion of the atom in the cavity mode, where it is subjected to a time-dependent light intensity; remember that we assumed the

detuning δ between the atomic frequency and the laser frequency to be large enough for no real photon absorption by the atom to occur[4]. If now the cavity contains one photon ($N = 1$), Figure 8 indicates that the gap between states $|e\,,1\rangle$ and $|g\,,1\rangle$ is reduced by the light shifts to $\hbar(\omega_L - 2\delta_0 - \delta_0)$, i.e. to $\hbar(\omega_L - 3\delta_0)$. When the atom exits the cavity, the coherence between its states e and g is now shifted by three times the amount obtained in (4). This means that an atom, traversing the cavity in a superposition of states g and e, keeps in the phase of that coherent superposition a trace of the number of photons present in the cavity; this occurs without any photon absorption (since the detuning δ is too large). To sum up, if $N = 0$, the state of the atom at the cavity exit is:

$$|\psi_{\text{fin}}(N = 0)\rangle = \frac{1}{\sqrt{2}}(|g\rangle + e^{-i\phi_0}|e\rangle) \tag{5}$$

whereas if $N = 1$, this state is:

$$|\psi_{\text{fin}}(N = 1)\rangle = \frac{1}{\sqrt{2}}(|g\rangle + e^{-i\phi_1}|e\rangle) \tag{6}$$

with $\phi_1 = 3\phi_0$.

How can we make use of this trace left on the atom by the possible presence of a photon in the cavity, and determine if this cavity contains zero or one photon? The time taken by the atom to cross the cavity can be adjusted by changing the atom's speed. Imagine that this time is tuned so that $\phi_0 - \phi_1 = \pi$; this means that the two states (5) and (6) are now orthogonal. As the atom exits the cavity, we can apply to it a $\pi/2$ laser pulse adjusted to transform $|\psi_f(N = 0)\rangle$ into $|g\rangle$. That same pulse will transform $|\psi_f(N = 1)\rangle$ into the state orthogonal to $|g\rangle$, that is to $|e\rangle$. This means that measuring the atomic state after this $\pi/2$ laser pulse allows concluding that $N = 0$ if the atom is found in state g, and that $N = 1$ if the atom is found in state e. The measurement can be repeated several times by sending a stream of atoms, one after the other, and applying to each of them the same procedure; one can measure several times in a row the same value N, which proves the number of photons in the cavity did not change during the measurements. As opposed to photoionization where a photon is absorbed giving rise to a photoelectron (Complement A_{XX}), this method is non-destructive: the presence of the photon is detected without it being absorbed. This experiment, generalized to the case where the photon number is larger than one, is described in more detail in reference [64].

Conclusion.

For a long time, light shifts have been considered as an interesting physical phenomenon without specific applications, and even as an undesired perturbation for high resolution spectroscopy, since they modify the atomic transition frequencies one is trying to measure with the highest possible precision. These shifts must be taken into account to extract from the measurements the non-perturbed frequencies of atomic transitions; most of the

[4]We also assume that the field variation encountered by the atom as it crosses the cavity is slow enough for non-adiabatic transitions from e to g, or from g to e, to be highly improbable. We also suppose that the natural width Γ of the excited state e, and the time T the atom takes to cross the cavity, are small enough for $\Gamma T \ll 1$, meaning spontaneous emission from state e does not have time to occur while the atom crosses the cavity.

time, several measurements at different light intensities must be performed to extrapolate the results to zero light intensity.

This complement clearly shows how much the situation has changed, by presenting the large variety of experimental methods using light shifts of atomic energy levels, and their great number of applications. These methods were implemented more than 20 years after these shifts were theoretically predicted and experimentally demonstrated; this illustrates the long term practical impact of fundamental research. These methods allow acting both on the internal and external atomic variables; they also permit using atoms as a very sensitive non-destructive probe for the properties of a field composed of only a few photons. These methods made it possible to trap atoms in a standing laser wave, or to obtain periodic lattices of neutral atoms trapped in such a wave. It also led to laser cooling methods that allowed reaching temperatures previously totally inaccessible for atomic gases, millions of times lower than the lowest temperatures found in the interstellar or intergalactic space of the Universe.

Complement E_{XX}

Detection of one- or two-photon wave packets, interference

1 One-photon wave packet, photodetection probability 2165
- 1-a Photoionization of a broadband detector 2165
- 1-b Detection probability amplitude 2166
- 1-c Temporal variation of the signal 2167

2 One- or two-photon interference signals 2167
- 2-a How should one compute photon interference? 2168
- 2-b Interference signal for a one-photon wave packet in two modes 2168
- 2-c Interference signals for a product of two one-photon wave packets 2170

3 Absorption amplitude of a photon by an atom 2174
- 3-a Computation of the amplitude 2174
- 3-b Properties of that amplitude 2175

4 Scattering of a wave packet 2176
- 4-a Absorption amplitude by atom B of the photon scattered by atom A 2176
- 4-b Wave packet scattered by atom A 2177

5 Example of wave packets with two entangled photons . . . 2181
- 5-a Parametric down-conversion 2181
- 5-b Temporal correlations between the two photons generated in parametric down-conversion 2183

Introduction

In Chapter XX the initial and final states of the atom + photon(s) system were chosen as states that, in the absence of interaction, had a well defined energy; before the interaction, such states do not evolve in time, as if the photon were not propagating in space. As an example, in the scattering process of a photon by an atom, the chosen initial radiation state is a photon with momentum $\hbar\boldsymbol{k}_i$ and energy $\hbar\omega_i = \hbar ck_i$, which spreads over the entire space; similarly, the final state is also a photon with momentum $\hbar\boldsymbol{k}_f$ and energy $\hbar\omega_f = \hbar ck_f$. The interaction was "turned on" at time t_i, which allowed computing the probability amplitude for the atom + photon(s) system to go from one state to the other between t_i and t_f. This is clearly a phenomenological approach: what actually happens is that the interaction operator remains constant but only comes into play to change the state vector when the atom is in the presence of radiation. A more realistic description of the process should involve the propagation of wave packets, with the incident radiation being described by a wave packet initially very far away from the atom, but going towards it. Their interaction then gives rise to a scattered wave packet moving away towards infinity, while the incident wave packet, modified by the interaction, also continues on its way.

Note however that introducing a wave packet for a photon cannot be done by the standard method used for a massive particle. As already pointed out at the end of § B-2 in Chapter XIX, a photon does not have a position operator. One cannot obtain its spatial wave function by projecting its state vector onto the eigenvectors of that operator, and then squaring this wave function's modulus to get the probability of finding the photon in any given region of space. One could then imagine using the spatial variations of the electric and magnetic fields to infer the photon localization. But for radiation states with exactly one, two, etc. photons, the average value of theses fields at any point in space is zero (it is the sum of zero average value creation and annihilation operators in each mode). Consequently, for a single photon, this average value cannot be directly used for building a wave packet localized in space. This is why we shall use another approach: we shall assume the photon interacts with detectors, well localized in space, and compute the probability of its detection by these apparatus. This will lead us to introduce an amplitude for the photon detection (by photoionization) at a given point, which presents close analogies with the spatial wave function of a massive particle in non-relativistic quantum mechanics.

We start in § 1 by exposing the general idea of this approach; we introduce a function $\mathcal{E}(\boldsymbol{r}, t)$ that allows localizing a single photon in space through its probability of being absorbed by a broadband detector. This leads to the concept of wave packets, even though the average value of the electric field remains zero throughout the entire space. In the perturbative computations, one can also introduce initial and final radiation states that are wave packets, described by linear superpositions of photon states with different momenta and energies.

In § 2, we show how the detection amplitude $\mathcal{E}(\boldsymbol{r}, t)$ allows studying light interference phenomena involving one or two photons. These phenomena are interpreted in terms of interference between the transition amplitudes associated with different paths leading the quantum field from a given initial state to a given final state. Starting first (§ 2-a) with a general discussion of the interference signals, we then focus in § 2-b on the interference involving one photon being simultaneously in two modes of the field. Finally, we examine interference involving two photons in the simple case where the system is described by the product of two one-photon wave packets (§ 2-c).

In § 3, we replace the broadband detector by an atom with discrete energy levels. Without having to assume that the coupling between the atom and the field is turned on abruptly (which is hard to justify from a physical point of view), a number of results of Chapter XX are confirmed with, in addition, the possibility of studying the temporal aspect of the absorption phenomenon. In § 4, we extend this method to study the scattering of photons by an atom. Here again, we will confirm the results of Chapter XX, while enriching our understanding of the temporal aspects of the physical process.

Finally, in § 5, we consider "real" two-photon wave packets that are entangled wave packets. Parametric down-conversion is an example of a situation leading to strong temporal correlations between two entangled photons. Such correlations are impossible to understand in terms of a classical treatment of the radiation.

In this entire complement, we have limited our studies to one- or two-photon wave packets, but the computations can be extended to wave packets containing a larger number of photons[1].

[1] Or even an undetermined number, as in a coherent state.

1. One-photon wave packet, photodetection probability

A one-photon wave packet is described by a state $|\psi\rangle$, where the photon number is precisely equal to one (eigenstate of the photon number operator, with eigenvalue equal to 1). We build this wave packet[2] as a linear superposition of states with different momenta $\hbar\boldsymbol{k}$:

$$|\psi\rangle = \int \mathrm{d}^3k\, c(\boldsymbol{k})\, a^\dagger_{\boldsymbol{k}}|0\rangle = \int \mathrm{d}^3k\, c(\boldsymbol{k})\, |\boldsymbol{k}\rangle \tag{1}$$

State $|\psi\rangle$ is not stationary (it is not an energy eigenstate). It is of the type considered in § B-3-c of Chapter XIX, an eigenket of operator $\hat{N}$ (total number of photons) with an eigenvalue equal to one. It is assumed to be normalized:

$$\int \mathrm{d}^3k\, |c(\boldsymbol{k})|^2 = 1, \tag{2}$$

which allows interpreting $|c(\boldsymbol{k})|^2$ as the probability density for the photon momentum to be equal to $\hbar\boldsymbol{k}$.

1-a. Photoionization of a broadband detector

Imagine we place an atom playing the role of a photodetector at point $\boldsymbol{r}$ in the radiation field described by state (1). According to relation (26) of Complement $\mathrm{B_{XX}}$, the probability $W_I(\boldsymbol{r},t)\mathrm{d}t$ for observing a photoelectron emission between times t and $t+\mathrm{d}t$ is given (to the interaction's lowest order) by:

$$W_I(\boldsymbol{r},t) = s\langle\psi|E^{(-)}(\boldsymbol{r},t)E^{(+)}(\boldsymbol{r},t)|\psi\rangle \tag{3}$$

where s is a constant depending on the photodetector sensitivity. $E^{(-)}(\boldsymbol{r},t)$ and $E^{(+)}(\boldsymbol{r},t)$ are the negative and positive frequency components of the electric field operator appearing in its plane wave expansion as given by (A-7) in Chapter XIX:

$$E^{(+)}(\boldsymbol{r},t) = \left[E^{(-)}(\boldsymbol{r},t)\right]^\dagger = i\int \frac{\mathrm{d}^3k}{(2\pi)^{3/2}}\,\sqrt{\frac{\hbar\omega_k}{2\epsilon_0}}\; a(\boldsymbol{k})\, e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega_k t)} \tag{4}$$

with:

$$\omega_k = c\times|\mathbf{k}| = c\times k \tag{5}$$

where c is the speed of light. Remember that expression (3) was established in the interaction representation, where the state vector $|\psi\rangle$ evolves only under the effect of the atom-radiation interaction; the operators evolve freely only under the effect of the atomic or radiation Hamiltonians (i.e. without mutual interaction). However, as we are performing a computation to lowest order, we can consider in (3) that $|\psi\rangle$ is actually

[2]For the sake of simplicity, we ignore in this complement the degrees of freedom of the radiation polarization, which do not play a significant role in the effects under study. This amounts to assuming that all the vectors $\boldsymbol{k}$ appearing in (1) have almost the same directions and that they are all associated with the same polarization $\boldsymbol{\varepsilon}$.

constant. The annihilation operator $E^{(+)}(\boldsymbol{r},t)$ in (3) acting on the one-photon state $|\psi\rangle$ yields the vacuum, and we can rewrite (3) as:

$$W_I(\boldsymbol{r},t) = s\langle\psi|E^{(-)}(\boldsymbol{r},t)|0\rangle\langle 0|E^{(+)}(\boldsymbol{r},t)|\psi\rangle = s|\langle 0|E^{(+)}(\boldsymbol{r},t)|\psi\rangle|^2 \tag{6}$$

that is:

$$W_I(\boldsymbol{r},t) = s\left|\int \frac{\mathrm{d}^3k}{(2\pi)^{3/2}}\,\sqrt{\frac{\hbar\omega_k}{2\epsilon_0}}\,c(\boldsymbol{k})e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega_k t)}\right|^2 \tag{7}$$

For a massive particle of mass m, the probability to find it at point $\boldsymbol{r}$ and at time t is given by the squared modulus $|\Psi(\boldsymbol{r},t)|^2$ of its wave function $\Psi(\boldsymbol{r},t)$. This wave function is the Fourier transform of the probability amplitude $g(\boldsymbol{k})$ that a measurement of the particle's momentum gives the value $\hbar\boldsymbol{k}$. For a free particle, this probability amplitude $g(\boldsymbol{k})$ has a time variation in $e^{-i\omega_k t}$. Equality (7) is thus reminiscent of this probability for a massive free particle; however, in view of the $\sqrt{\omega_k}$ factor (proportional to $\sqrt{k}$) in front of $c(\boldsymbol{k})$ in the integral of (7), $W_I(\boldsymbol{r},t)$ is not proportional (at a given instant t) to the modulus squared of the spatial Fourier transform of $c(\boldsymbol{k},t) = c(\boldsymbol{k})e^{-i\omega_k t}$. This means that, limiting ourselves to one-photon states, we can indeed consider the function $c(\boldsymbol{k})$ appearing in (1) as a wave function in momentum space, since $|c(\boldsymbol{k})|^2$ is a probability density for the photon momentum. However, the probability to detect a photon at point $\boldsymbol{r}$ and at time t with a photodetector is not simply the modulus squared of the Fourier transform of that "wave function in momentum space" $c(\boldsymbol{k})e^{-i\omega_k t}$. This confirms that it is not possible, for a photon, to introduce a spatial wave function that is exactly equivalent to that of a massive particle.

1-b. Detection probability amplitude

The right-hand side of (6) is the squared modulus of the function:

$$\mathcal{E}(\boldsymbol{r},t) = \langle 0|E^{(+)}(\boldsymbol{r},t)|\psi\rangle \tag{8}$$

which plays an important role in all the computations to follow; it has the dimensions of an electric field. For the wave packet written in (1), its expression is:

$$\mathcal{E}(\boldsymbol{r},t) = i\int \frac{\mathrm{d}^3k}{(2\pi)^{3/2}}\,\sqrt{\frac{\hbar\omega_k}{2\epsilon_0}}\,c(\boldsymbol{k})e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega_k t)} \tag{9}$$

It should not be confused with the average value in state $|\psi\rangle$ of the operator $E^{(+)}(\boldsymbol{r},t)$ written in (4), since that average value is zero, as we mentioned above. Nor is it, as already pointed out, a wave function for the photon in position space; it is a probability amplitude for the detection (and not the presence) of the photon at point $\boldsymbol{r}$ and time t. When we mention, in this complement, the space time wave packet associated with the photon in state (1), we will always be referring to the amplitude (8).

Note, however, that in the particular case where the function $c(\boldsymbol{k})$ in momentum space is well centered around a value k_m with a dispersion Δk very small compared to k_m, one can neglect in (7) the k variation of $\sqrt{k}$ and replace $\sqrt{k}$ by $\sqrt{k_m}$; the integral in (7) will therefore involve the spatial Fourier transform of $c(\boldsymbol{k})e^{-i\omega_k t}$. This approximation will often be used in what follows.

Comments:

(i) In this section, we have ignored the radiation polarization, assuming that all the plane waves in relation (1) are associated with the same polarization vector $\boldsymbol{\varepsilon}$. When this is not the case, the detection amplitude becomes a three-component vector function. Its component along any given axis yields the detection probability amplitude by a photodetector preceded by a polarization analyser letting through only the light polarized linearly along that axis. This vector detection amplitude is similar to the wave function of a spin 1 particle, which also has three components.

(ii) In this complement, we shall study only wave packets containing a well defined photon number, 1 or 2, which allows directly generalizing the computations of Chapter XX. Wave packets can, however, be built many different ways, without exactly fixing the photon number. It often happens, for example, that one wishes to reproduce a classical field for which each field mode $\mathbf{k}$ has a given amplitude $\alpha(\mathbf{k})$; it is then natural to use a state where each quantum mode is in a coherent state with eigenvalue $\alpha(\mathbf{k})$. In that case, only the average photon number is fixed, not its exact value.

1-c. Temporal variation of the signal

When the photodetector is placed at $\mathbf{r} = 0$, it delivers a signal that is given, according to (7), by:

$$W_I(\mathbf{r} = \mathbf{0}, t) \propto \left| \int \mathrm{d}^3k \, \sqrt{k} \, c(\boldsymbol{k}) e^{-ickt} \right|^2 \tag{10}$$

This signal is proportional to the squared modulus of the Fourier transform of $\sqrt{k}\, c(\boldsymbol{k})$. Let us assume that $c(k)$ is a real positive function of k, and that $|c(k)|^2$ is a function of k centered at $k = k_m$, with a width Δk. If this width Δk of the wave packet is very small compared with the average wave number k_m, one can replace $\sqrt{k}$ by $\sqrt{k_m}$, and the signal becomes proportional to the modulus squared of the Fourier transform of $c(\boldsymbol{k})$. At time $t = 0$, and since we assumed all the $c(k)$ to be positive, all the waves forming the wave packets are in phase and the signal observed on the photodetector is maximum; it is zero for $t = \pm\infty$, and takes on significant values only during a time interval $\Delta t \simeq 1/c\Delta k$ around $t = 0$. This signal describes, in a way, the passage of the wave packet at the detector's position.

To study the detection probability at a point $\mathbf{r} \neq \mathbf{0}$, we just have to replace e^{-ickt} by $e^{i(\mathbf{k}\cdot\mathbf{r}-kct)}$ in the integral of relation (10). As an example, imagine we have a one-dimensional wave packet, all the wave vectors $\boldsymbol{k}$ being parallel to the Oz axis ($k_x = k_y = 0$); the exponential reduces to $e^{-ik(z-ct)}$. The phenomena observed at a point in space of coordinate $z \neq 0$ are thus deduced from those observed at $z = 0$ by a simple time shift equal to z/c: the wave packet moves along the Oz direction with velocity c and without any deformation.

2. One- or two-photon interference signals

We now discuss in terms of wave packets what happens in light interference experiments involving one or two photons.

2-a. How should one compute photon interference?

In non-relativistic quantum mechanics, a particle with a non-zero mass m is described by a wave function $\psi(\boldsymbol{r},t)$ whose squared modulus $|\psi(\boldsymbol{r},t)|^2$ yields the probability density of finding the particle at point $\boldsymbol{r}$ and time t. In a Young's type interference experiment, the wave function, after going through the two slits pierced into a screen, is a linear superposition of two wave functions $\psi_1(\boldsymbol{r},t)$ and $\psi_2(\boldsymbol{r},t)$ originating from the two slits. These two waves overlap in a region of space where the probability density of finding the particle at point $\boldsymbol{r}$ and time t, which is equal to $|\psi_1(\boldsymbol{r},t)+\psi_2(\boldsymbol{r},t)|^2$, contains a term $2\mathrm{Re}\{\psi_1(\boldsymbol{r},t)\psi_2^*(\boldsymbol{r},t)\}$ oscillating in space and time; this results in interference fringes.

However, we recalled in § 1 why we cannot, in general, introduce a spatial wave function for a photon that would be strictly analogous to $\psi(\boldsymbol{r},t)$, and whose squared modulus would yield the probability density for the presence of the photon at a given point. This led us to define an amplitude $\mathcal{E}(\boldsymbol{r},t)$ in (8), whose squared modulus yields the probability density for photodetecting the photon at point $\boldsymbol{r}$ and time t. We are going to show in § 2-b that such amplitudes can actually be used to interpret interference fringes; as an example, we shall study the fringes appearing in the single photodetection signal $w_I(\boldsymbol{r},t)$ observed on a one-photon wave packet after it goes through a screen pierced with two slits. As already underlined, it is important not to confuse $\mathcal{E}(\boldsymbol{r},t)$ with the average value of the electric field in the quantum state under study – which in any case is zero in a one-photon state. In classical electromagnetism, the electric (or magnetic) fields directly interfere; in quantum electromagnetism, one must reason in terms of probability amplitudes.

For field states containing at least two photons, the double photodetection signal $w_{II}(\boldsymbol{r}_B,t_B,\boldsymbol{r}_A,t_A)$ is different from zero. To interpret it in the simplest possible case, we assume (in § 2-c) that the radiation is described by a tensor product of two one-photon wave packets[3]. We will show that interference fringes observable on w_{II} can also be interpreted in terms of products of detection probability amplitudes; these fringes result from interference between transition amplitudes associated with two different paths leading the field from its initial state (where it contains two photons) to the vacuum. Here again, one should reason in terms of interference not directly between average values of electric or magnetic fields, but between paths.

2-b. Interference signal for a one-photon wave packet in two modes

We start with the simplest photon interference experiment, the well-know Young's double slit experiment, but in a case when only one photon at a time passes through the screen pierced with the two slits. The state vector of this single photon is then the sum of two components associated with the passage through one or the other of the two slits. When the photon reaches the interference region, these two components are associated with two different radiation modes.

[3] A simple example of a two-photon entangled state, which is not a product of two one-photon states, will be studied in § 5.

α. One-photon wave packets, after passage through a two-slit screen

We now focus on the radiation state after it passed the two-slit screen. This state is described by a one-photon wave packet, which, in the interaction representation, is of the form:

$$|\psi\rangle = c_1|\psi_1\rangle + c_2|\psi_2\rangle \qquad \text{with:} \quad |c_1|^2 + |c_2|^2 = 1 \tag{11}$$

In this expression, the ket $|\psi_1\rangle$ describes the wave packet emerging from the first slit; as in (1), it can therefore be written with a function $c_1(\mathbf{k})$ that is peaked around the value $\boldsymbol{k}_1$. The ket $|\psi_2\rangle$ describes the wave packet emerging from the second slit, and its function $c_2(\mathbf{k})$ is peaked around the value $\boldsymbol{k}_2$. Since before going through the slits the two wave packets came from the same source, they must be centered around a common frequency $\omega_m = ck_1 = ck_2$; consequently $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ have the same modulus, but their directions can be different. We shall finally assume that the wave packets emerging from the two slits arrive at the same time in the interference region (meaning the optical paths along the two trajectories are equal) and that each wave packet is sufficiently long for the frequency ω_m to be well defined.

As in (8), for each wave packet $|\psi_i\rangle$ we introduce a detection amplitude $\mathcal{E}_i(\boldsymbol{r}, t)$:

$$\langle 0|E^{(+)}(\boldsymbol{r}, t)|\psi_i\rangle = \mathcal{E}_i(\boldsymbol{r}, t) \qquad \text{where:} \quad i = 1, 2 \tag{12}$$

In the interference region, we assume the two modes[4] to be close to plane waves with wave vectors $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$. We then set:

$$\mathcal{E}_i(\boldsymbol{r}, t) = \mathcal{F}_i(\boldsymbol{r}, t)\ e^{i(\boldsymbol{k}_i \cdot \boldsymbol{r} - \omega_m t)} \tag{13}$$

where the function $\mathcal{F}_i(\boldsymbol{r}, t)$ has a much slower space and time variation than the exponential $e^{i(\boldsymbol{k}_i \cdot \boldsymbol{r} - \omega_m t)}$.

β. Calculation of the single photodetection signal

We assume that the field is contained in a box of volume L^3; we use a complete orthonormal set of field modes, with wave vectors $\{\boldsymbol{k}_i\}$, which includes both $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$. Relation (B-3) of Chapter XIX indicates that the positive frequency component of the electric field can be written [5] as:

$$E^{(+)}(\boldsymbol{r}, t) = i \sum_i \sqrt{\frac{\hbar\omega_i}{2\varepsilon_0 L^3}}\ e^{i(\boldsymbol{k}_i \cdot \boldsymbol{r} - \omega_i t)}\ a_i \tag{14}$$

with $\omega_i = c \times k_i$. When this operator acts on the ket (11), all the i terms lead to a zero result, except for the $i = 1$ and $i = 2$ terms. For these two terms, we have $a_i|\psi_{1,2}\rangle = |0\rangle$, so that (11) and (12) lead to:

$$E^{(+)}(\boldsymbol{r}, t)|\psi\rangle = \big[c_1\ \mathcal{E}_1(\boldsymbol{r}, t) + c_2\ \mathcal{E}_2(\boldsymbol{r}, t)\big]|0\rangle \tag{15}$$

[4] Another possibility would be to use Gaussian wave packets with the same focal point, having in the vicinity of that point plane wave structures with wave vectors $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$, and lateral extensions very large compared to the wavelengths $2\pi/k_1$ and $2\pi/k_2$.

[5] For the sake of simplicity, we ignore polarization variables of the field.

The probability for detecting the photon at point $\boldsymbol{r}$ and time t is proportional to the square of norm of this ket, written as:

$$\langle\psi|E^{(-)}(\boldsymbol{r},t)E^{(+)}(\boldsymbol{r},t)|\psi\rangle = \left|c_1\mathcal{E}_1(\boldsymbol{r},t) + c_2\mathcal{E}_2(\boldsymbol{r},t)\right|^2 \tag{16}$$

The equality includes square and cross terms. The square terms can be written, taking (13) into account:

$$|c_i\mathcal{E}_i(\boldsymbol{r},t)|^2 = |c_i\mathcal{F}_i(\boldsymbol{r},t)|^2 \tag{17}$$

and they vary slowly as a function of $\boldsymbol{r}$ and t. The crossed terms are expressed as:

$$c_1c_2^*\ \mathcal{E}_1(\boldsymbol{r},t)\mathcal{E}_2^*(\boldsymbol{r},t) + \text{c.c.} = c_1c_2^*\ \mathcal{F}_1(\boldsymbol{r},t)\mathcal{F}_2^*(\boldsymbol{r},t)\exp\{i[(\boldsymbol{k}_1-\boldsymbol{k}_2)\cdot\boldsymbol{r})]\} + \text{c.c.} \tag{18}$$

and exhibit spatial modulations characteristic of interference phenomena (c.c. stands for complex conjugate).

γ. Discussion

Relation (16) shows that the photodetection signal is the squared modulus of the sum of two amplitudes, $c_1\mathcal{E}_1(\boldsymbol{r},t)$ and $c_2\mathcal{E}_2(\boldsymbol{r},t)$, which interfere. Amplitude $c_1\mathcal{E}_1(\boldsymbol{r},t)$ is the amplitude for detecting at point $\boldsymbol{r}$ and time t the photon in mode $|\psi_1\rangle$; it is equal to the amplitude c_1 of finding the field in state $|\psi_1\rangle$, multiplied by the amplitude $\mathcal{E}_1(\boldsymbol{r},t)$ for detecting the photon at point $\boldsymbol{r}$ and time t when the field is in state $|\psi_1\rangle$. The amplitude $c_2\mathcal{E}_2(\boldsymbol{r},t)$ is interpreted in a similar way. During the detection process, the field goes from state $|\psi\rangle$ written in (11) to the vacuum state $|0\rangle$ following two possible paths: the photon is absorbed either while in mode $|\psi_1\rangle$, or while in mode $|\psi_2\rangle$. As nothing allows deciding which path the system followed, the two corresponding amplitudes interfere. This confirms what we stated above: in the quantum theory of radiation, the interference fringes observed on a photodetector signal are associated with the interference, not between two classical electromagnetic waves, but rather between two transition amplitudes corresponding to different paths (leading the system from the same initial state to the same final state).

2-c. Interference signals for a product of two one-photon wave packets

Let us generalize this type of interpretation, in terms of transition amplitudes, to interference experiments involving two photons and where one measures correlations between signals coming from two photodetectors.

α. State vector for the two photons

We now assume the field contains two photons, and can be described as the product of two wave packets such as the one written in (1):

$$|\psi_{12}\rangle = \int \mathrm{d}^3k\ c_1(\boldsymbol{k}) \int \mathrm{d}^3k'\ c_2(\boldsymbol{k}')\, a_{\boldsymbol{k}}^\dagger a_{\boldsymbol{k}'}^\dagger|0\rangle \tag{19}$$

Is it possible to observe spatial and temporal modulations on the signals w_I and w_{II} coming from one or two detectors placed in that field? We are going to show that the

answer to that question is "no" if one considers only single photodetections, but "yes" if one takes into account the correlations between double photon detections.

We assume the two wave packets in (19) to be well separated: there exists no overlap between the domain D_1 where the function $c_1(\boldsymbol{k})$ is different from zero and the domain D_2 where the function $c_2(\boldsymbol{k})$ is non-zero. Let us introduce the matrix element that generalizes relation (8) to the two-photon case:

$$\langle 0|E^{(+)}(\boldsymbol{r},t)\ E^{(+)}(\boldsymbol{r}',t')|\psi_{12}\rangle \tag{20}$$

We now insert in that expression the plane mode expansion (14) for both electric field operators. To yield a non-zero result, the annihilation operators appearing in these fields must act on a mode that, according to (19), contains one photon. This means that either the mode selected in $E^{(+)}(\boldsymbol{r},t)$ belongs to the D_1 domain and the one selected in $E^{(+)}(\boldsymbol{r}',t')$ to the D_2 domain, or the inverse. In the first case, the scalar product of the vacuum bra and the modes that came into play yields the detection amplitude $\mathcal{E}_2(\boldsymbol{r}',t')$ associated with the second wave packet, multiplied by the detection amplitude $\mathcal{E}_1(\boldsymbol{r},t)$ associated with the first one. In the second case, the wave packets are inverted. The final result is:

$$\langle 0|E^{(+)}(\boldsymbol{r},t)\ E^{(+)}(\boldsymbol{r}',t')|\psi_{12}\rangle = \mathcal{E}_1(\boldsymbol{r},t)\ \mathcal{E}_2(\boldsymbol{r}',t') + \mathcal{E}_2(\boldsymbol{r},t)\ \mathcal{E}_1(\boldsymbol{r}',t') \tag{21}$$

where the functions $\mathcal{E}_1$ and $\mathcal{E}_2$ are the detection amplitudes associated, as defined in (8), with the two individual wave packets included in $|\psi_{12}\rangle$.

β. Single photodetection signal $w_I(\boldsymbol{r},t)$

To get the single photodetection signal, we first compute the result of the action on state (19) of the field positive frequency component $E^{(+)}(\boldsymbol{r}',t')$. As we argued above, to yield a non-zero result, this operator must destroy a photon, either in a mode for which $c_1(\mathbf{k})$ is non-zero, or in a mode for which $c_2(\mathbf{k})$ is non-zero. In the first case, the summation over all the modes involved reconstructs the function $\mathcal{E}_1(\boldsymbol{r}',t')$ multiplied by the vacuum ket associated with these modes; the modes of the other wave packet remain unchanged. In the second case, the two wave packets exchange roles and it is now the function $\mathcal{E}_2(\boldsymbol{r}',t')$ that is reconstructed. This leads to:

$$E^{(+)}(\boldsymbol{r}',t')|\psi_{12}\rangle = \left[\mathcal{E}_1(\boldsymbol{r}',t')\int \mathrm{d}^3k'\ c_2(\boldsymbol{k}')\,a^\dagger_{\boldsymbol{k}'}|0\rangle + \mathcal{E}_2(\boldsymbol{r}',t')\int \mathrm{d}^3k\ c_1(\boldsymbol{k})\,a^\dagger_{\boldsymbol{k}}|0\rangle\right] \tag{22}$$

The probability per unit time to detect a photon at point $\boldsymbol{r}'$ and time t' is the square of this ket's norm. The terms in $|\mathcal{E}_1(\boldsymbol{r}',t')|^2$ and $|\mathcal{E}_2(\boldsymbol{r}',t')|^2$ contain the square of the two wave packets' norms, each equal to one; these terms do not oscillate, neither in space, nor in time. The cross terms are the only ones that could yield spatial and temporal modulations; they contain, however, the scalar product of the two wave packets, which is zero since we assumed the wave packets were orthogonal (there is no overlap between the two D_1 and D_2 domains). This means that, when the field is described by state (19), no interference fringes are observable in the signal of a single photodetector.

The interpretation of this result is similar to the one we gave before. The system can follow two paths: either an absorption of a photon from the first wave packet, or an

absorption of a photon from the second. However, as opposed to what happens when the system started from the initial state (11), the final state of the field is not the same for these two paths: if a photon from the first wave packet has been absorbed, the final state includes a photon from the second wave packet. Consequently, the two final states associated with the two paths are orthogonal, and observing the field's final state one could (in principle) determine which path the system has followed; this is why the two amplitudes cannot interfere.

Comment:

One could consider other states for the two modes, each containing several photons, as for example states $|\alpha_1\rangle \otimes |\alpha_2\rangle$ where each mode is in a coherent state, characterized by a classical normal variable, α_1 for mode $\{\boldsymbol{k}_1\}$, α_2 for mode $\{\boldsymbol{k}_2\}$. We then know that state $|\alpha_1\rangle$ is an eigenstate of operator $E^{(+)}(\boldsymbol{r},t)$ with an eigenvalue value $\mathcal{E}_{\rm cl}^{(+)}(\alpha_1,\boldsymbol{r},t)$ equal to the positive frequency component of the classical field in mode $\{\boldsymbol{k}_1\}$, corresponding to the classical normal variable α_1 (Chapter XVIII, § B-2). A similar result is valid for state $|\alpha_2\rangle$:

$$E^{(+)}(\boldsymbol{r},t)|\alpha_i\rangle = \mathcal{E}_{\rm cl}^{(+)}(\alpha_i,\boldsymbol{r},t)|\alpha_i\rangle \qquad i=1,2 \tag{23}$$

This leads to:

$$E^{(+)}(\boldsymbol{r},t)|\alpha_1,\alpha_2\rangle = \left[\mathcal{E}_{\rm cl}^{(+)}(\alpha_1,\boldsymbol{r},t)+\mathcal{E}_{\rm cl}^{(+)}(\alpha_2,\boldsymbol{r},t)\right]|\alpha_1,\alpha_2\rangle \tag{24}$$

The probability of detecting a photon at point $\boldsymbol{r}$ and time t is equal to the square of the norm of ket (24). It is proportional to:

$$\left|\mathcal{E}_{\rm cl}^{(+)}(\alpha_1,\boldsymbol{r},t)+\mathcal{E}_{\rm cl}^{(+)}(\alpha_2,\boldsymbol{r},t)\right|^2$$

As this is the squared modulus of the sum of two classical fields, it is the usual interference signal of classical fields. As opposed to what we found before for the radiation state (19), when the two modes are in coherent states, the one-photon detection signal exhibit interference. This is an illustration of the quasi-classical character of coherent states.

γ. *Double photodetection signal* $w_{II}(\boldsymbol{r}_B,t_B,\boldsymbol{r}_A,t_A)$

Assuming, as above, the field initial state is given by (19), we now focus on the probability $w_{II}(\boldsymbol{r}_B,t_B,\boldsymbol{r}_A,t_A)$ (per double unit time) that a detector, placed at $\boldsymbol{r}_A$, detects a photon at time t_A and that another detector, placed at $\boldsymbol{r}_B$, detects a photon at time t_B. This probability is proportional to the correlation function (Complement B_{XX}, § 2-d):

$$\langle\psi_{12}|E^{(-)}(\boldsymbol{r}_A,t_A)E^{(-)}(\boldsymbol{r}_B,t_B)E^{(+)}(\boldsymbol{r}_B,t_B)E^{(+)}(\boldsymbol{r}_A,t_A)|\psi_{12}\rangle \tag{25}$$

Since $|\psi_{12}\rangle$ contains only two photons, we can insert in the middle of this expression the projector onto the vacuum state, which leads to the squared modulus of expression (21). We obtain:

$$w_{II}(\boldsymbol{r}_B,t_B;\boldsymbol{r}_A,t_A) \propto \left|\mathcal{E}_2(\boldsymbol{r}_B,t_B)\mathcal{E}_1(\boldsymbol{r}_A,t_A)+\mathcal{E}_1(\boldsymbol{r}_B,t_B)\mathcal{E}_2(\boldsymbol{r}_A,t_A)\right|^2 \tag{26}$$

In addition to the square terms $|\mathcal{E}_2(\boldsymbol{r}_B,t_B)\mathcal{E}_1(\boldsymbol{r}_A,t_A)|^2$ and $|\mathcal{E}_1(\boldsymbol{r}_B,t_B)\mathcal{E}_2(\boldsymbol{r}_A,t_A)|^2$, which have a slow variation with $\boldsymbol{r}_A$, $\boldsymbol{r}_B$, t_A, t_B, we get cross terms:

$$\mathcal{E}_1(\boldsymbol{r}_B,t_B)\mathcal{E}_1^*(\boldsymbol{r}_A,t_A)\mathcal{E}_2(\boldsymbol{r}_A,t_A)\mathcal{E}_2^*(\boldsymbol{r}_B,t_B)+\text{c.c.}$$

which do have spatial and temporal modulations. If, for example, the first wave packet is centered around the values $\boldsymbol{k}_1$ and ω_1, and the second one around the values $\boldsymbol{k}_2$ and ω_2, these modulations are of the form:

$$\exp\Big\{i\big[(\boldsymbol{k}_1-\boldsymbol{k}_2)\cdot(\boldsymbol{r}_A-\boldsymbol{r}_B)-(\omega_1-\omega_2)(t_A-t_B)\big]\Big\}+\text{c.c.} \tag{27}$$

This result is not in contradiction with the fact that the probability of detecting a photon at $\mathbf{r}_A$, t_A, or at $\mathbf{r}_B$, t_B varies slowly with these variables: once a first photon has been detected at $\boldsymbol{r}_A,t_A$, the probability to detect another one at $\boldsymbol{r}_B,t_B$ varies sinusoidally with $\boldsymbol{r}_A-\boldsymbol{r}_B$ and t_A-t_B.

(i) ***Discussion***

The amplitude whose squared modulus appears on the right-hand side of (26) is the sum of two amplitudes associated with two possible paths leading the system from the initial state $|\psi_{12}\rangle$ (containing two photons) to the same final state $|0\rangle$ (where all the modes are empty). Along the first path, with amplitude $\mathcal{E}_2(\boldsymbol{r}_B,t_B)\mathcal{E}_1(\boldsymbol{r}_A,t_A)$, the k_1 mode photon is absorbed at $\boldsymbol{r}_A,t_A$ and the k_2 mode photon is absorbed at $\boldsymbol{r}_B,t_B$. Along the second path, with amplitude $\mathcal{E}_1(\boldsymbol{r}_B,t_B)\mathcal{E}_2(\boldsymbol{r}_A,t_A)$, the opposite happens: the k_2 mode photon is absorbed at $\boldsymbol{r}_A,t_A$ and the k_1 mode photon is absorbed at $\boldsymbol{r}_B,t_B$. As explained before (§ 2-b), interference occurs between the different transition amplitudes associated with two possible paths leading the system from the same initial state to the same final state, as long as there is no way one can determine which path is actually followed.

(ii) ***Another interpretation***

The photodetection signal (25) can also be written in the form:

$$\langle\psi_A|E^{(+)}(\boldsymbol{r}_B,t_B)E^{(+)}(\boldsymbol{r}_B,t_B)|\psi_A\rangle \tag{28}$$

with:

$$\begin{aligned}|\psi_A\rangle &= E^{(+)}(\boldsymbol{r}_A,t_A)|\psi_{12}\rangle\\ &= \Big[\mathcal{E}_1(\boldsymbol{r}_A,t_A)\int \mathrm{d}^3k'\ c_2(\boldsymbol{k}')\,a^\dagger_{\boldsymbol{k}'}|0\rangle+\mathcal{E}_2(\boldsymbol{r}_A,t_A)\int \mathrm{d}^3k\ c_1(\boldsymbol{k})\,a^\dagger_{\boldsymbol{k}}|0\rangle\Big]\end{aligned} \tag{29}$$

where, in the second line, we used relation (22). Signal (28) can be interpreted as the probability of detecting a photon when the field is that state $|\psi_A\rangle$ where the photon has a probability amplitude $\mathcal{E}_1(\boldsymbol{r}_A,t_A)$ to be in the wave packet with amplitude $c_2(\boldsymbol{k}')$, and a probability amplitude $\mathcal{E}_2(\boldsymbol{r}_A,t_A)$ to be in the other wave packet with amplitude $c_1(\boldsymbol{k})$. This situation is quite similar to that encountered in § 2-b-α, where we showed that the photodetection probability of a photon in state (11) exhibits modulations.

In other words, we started from a state $|\psi_{12}\rangle$ with no coherence. It is the detection of a first photon that introduces the state (29) where a second photon is now in a coherent superposition, the coherence arising from the fact that the detected photon can come either from the first wave packet, or from the second. The coefficients of the

superposition (29) depend on the point $\boldsymbol{r}_A$ and the instant t_A where the detection of the first photon occurred. In this description of the phenomena, it is the first detection that introduces quantum correlations between the two modes, and the dependence of these correlations on $\boldsymbol{r}_A$ and t_A explains why the probability of the second detection oscillates as a function of $\boldsymbol{r}_B - \boldsymbol{r}_A$ and $t_B - t_A$.

3. Absorption amplitude of a photon by an atom

We now replace the broadband photodetector by an atom with two discrete levels, a ground level a and an excited level b. This atom is placed at $\boldsymbol{r} = \boldsymbol{0}$, and interacts with the same wave packet $|\psi\rangle$ as that written in (1). We propose to compute the probability amplitude for the atom, initially in state a, to absorb the incident photon and be found in state b at time t.

3-a. Computation of the amplitude

The initial and final states of the process under study are:

$$|\psi_{\text{in}}\rangle = |a;\psi\rangle \qquad |\psi_{\text{fin}}\rangle = |b;0\rangle \tag{30}$$

since the absorption of the photon transfers the radiation from state $|\psi\rangle$ to the vacuum $|0\rangle$. According to relation (B-4) in Chapter XX, the amplitude we are looking for is, to first order in H_I:

$$\langle\psi_{\text{fin}}|\bar{U}(t,-\infty)|\psi_{\text{in}}\rangle = \frac{1}{i\hbar}\int_{-\infty}^{t} \mathrm{d}t' \ \langle\psi_{\text{fin}}|\bar{H}_I(t')|\psi_{\text{in}}\rangle \tag{31}$$

where the bar above the operators indicates they are expressed in the interaction picture, with respect to the non-perturbed Hamiltonian.

The interaction Hamiltonian H_I is given by[6]:

$$H_I = -DE^{(+)}(\boldsymbol{r} = \boldsymbol{0}) \tag{32}$$

The matrix element of $\bar{H}_I(t')$ appearing in (31) equals:

$$\langle\psi_{\text{fin}}|\bar{H}_I(t')|\psi_{\text{in}}\rangle = -\mathcal{D}e^{i\omega_0 t'}\langle 0|E^{(+)}(\boldsymbol{r} = \boldsymbol{0}, t')|\psi\rangle \tag{33}$$

In this equality, $\mathcal{D} = \langle b|D|a\rangle$, $\omega_0 = (E_b - E_a)/\hbar$ is the frequency of the atomic transition, and $E^{(+)}(\boldsymbol{r} = \boldsymbol{0}, t')$ the electric field positive frequency component in the interaction representation. Using notation (8) for the matrix element on the right-hand side of (33), we get:

$$\langle\psi_{\text{fin}}|\bar{H}_I(t')|\psi_{\text{in}}\rangle = -\mathcal{D}\ e^{i\omega_0 t'}\ \mathcal{E}(\boldsymbol{r} = \boldsymbol{0}, t') \tag{34}$$

which allows rewriting the absorption amplitude (31) as:

$$\langle\psi_{\text{fin}}|\bar{U}(t,-\infty)|\psi_{\text{in}}\rangle = -\frac{\mathcal{D}}{i\hbar}\int_{-\infty}^{t} \mathrm{d}t' \ e^{i\omega_0 t'}\ \mathcal{E}(\boldsymbol{r} = \boldsymbol{0}, t') \tag{35}$$

[6] As we have ignored the radiation polarization degrees of freedom, we also ignore here the vector character of the atomic dipole $\boldsymbol{D}$. Operator D appearing in (32) is actually the projection of $\boldsymbol{D}$ onto the radiation polarization vector.

The quantity $-\mathcal{D}\mathcal{E}(\boldsymbol{r}=\boldsymbol{0},t')$ that appears in this expression can be considered to be the interaction Hamiltonian between the atomic dipole $\mathcal{D}$ and a classical field $\mathcal{E}(\boldsymbol{r}=\boldsymbol{0},t')$. The field $\mathcal{E}(\boldsymbol{r}=\boldsymbol{0},t')$ thus appears as the classical field that would yield the same transition amplitude between the two atomic states a and b as a quantum field, when the radiation is in a one-photon wave packet described by $c(\boldsymbol{k})$.

3-b. Properties of that amplitude

Let us return to the wave packet in infinite space (1) and use relation (8) to get the probability amplitude for the detection of the photon at point $\mathbf{r}=\mathbf{0}$. Describing the field operator (4) with an integral over d^3k' instead of d^3k, and using it in (1), the commutation of operators $a(\boldsymbol{k}')$ and $a^\dagger(\boldsymbol{k})$ leads to a $\delta(\boldsymbol{k}-\boldsymbol{k}')$ function. This leads to:

$$\langle 0|E^{(+)}(\boldsymbol{r}=\boldsymbol{0},t')|\psi\rangle = \mathcal{E}(\boldsymbol{r}=\boldsymbol{0},t') = \frac{i}{(2\pi)^{3/2}}\int \mathrm{d}^3k\,\sqrt{\frac{\hbar\omega_k}{2\varepsilon_0}}\;c(k)e^{-i\omega t'} \tag{36}$$

where $\omega_k = ck$. As $c(k)$ is centered around an average wave vector k_m, assumed very large compared to the width Δk of $c(k)$, amplitude (36) can be written as:

$$\mathcal{E}(\boldsymbol{r}=\boldsymbol{0},t') = e^{-i\omega_m t'}\;\mathcal{F}(t') \tag{37}$$

which is the product of a carrier wave of frequency $\omega_m = ck_m$ by an envelope $\mathcal{F}(t')$. This latter function has a t' variation that is slower than the preceding exponential; it can, for example, follow a bell-shaped curve, centered at $t=0$ and with width $\Delta t \simeq 1/c\Delta k$. Inserting (37) into (35), we get:

$$\langle\psi_{\mathrm{fin}}|\bar{U}(t,-\infty)|\psi_{\mathrm{in}}\rangle = -i\int_{-\infty}^{t}\mathrm{d}t'\;e^{i(\omega_0-\omega_m)t'}\;\Omega_1(t') \tag{38}$$

where $\Omega_1(t')$ is the instantaneous Rabi frequency defined as:

$$\hbar\Omega_1(t') = -\mathcal{D}\mathcal{F}(t') \tag{39}$$

Equation (38) allows understanding the behavior of the absorption amplitude of the photon when t increases from $-\infty$ to $+\infty$. As long as $t \ll -\Delta t$, both functions $\mathcal{E}(t')$ and $\Omega_1(t')$ are zero; the incident wave packet has not yet reached the atom's vicinity and no photon absorption can occur. As t increases from $-\Delta t/2$ to $+\Delta t/2$, the wave packet crosses the atom, and the integral in (38) becomes larger. When $t \gg +\Delta t$, the wave packet has left the atom; the absorption amplitude remains constant and equal to:

$$\langle\psi_{\mathrm{fin}}|\bar{U}(+\infty,-\infty)|\psi_{\mathrm{in}}\rangle = -i\int_{-\infty}^{+\infty}\mathrm{d}t'\;e^{i(\omega_0-\omega_m)t'}\;\Omega_1(t') \tag{40}$$

This expression yields the probability amplitude for a photon to have been absorbed once the wave packet has crossed the atom. This confirms the results of Chapter XX, but in the present approach we did not have to artificially introduce any initial or final times for the process.

Let us evaluate an order of magnitude for the amplitude (40). Assume first that $\omega_m = \omega_0$ (resonant wave packet). The integral in (40) is then of the order of $\Omega_1^{\max}\Delta t$,

where $\Omega_1^{\max}$ is the maximum value reached by the Rabi frequency when the atom is at the center of the wave packet, and the envelope $\mathcal{F}(t')$ takes on its largest value. When $\omega_m \neq \omega_0$ (off-resonance wave packet), the absorption amplitude is weaker. According to (40), this amplitude is actually the Fourier transform of $\Omega_1(t')$ at frequency $\omega_0 - \omega_m$. This result simply expresses energy conservation: for the incident photon to be absorbed, its frequency must be equal to the atomic transition frequency. However, as the field envelope varies over time intervals of the order of Δt, the photon average frequency does not have to be strictly equal to the atomic frequency; the two frequencies must be equal to within $\Delta\omega \simeq 1/\Delta t$.

4. Scattering of a wave packet

We now study a process involving two atoms: a wave packet impinges on an atom A placed at $\boldsymbol{r}_A$ on the z axis; after interacting with it, the wave packet is scattered in all directions, and then interacts with a second atom B placed at $\boldsymbol{r}_B$. The incident wave packet, propagating along the Oz direction, is described by the function $c(\boldsymbol{k})e^{-i\omega t}$. As before, we have two main goals. The first one is, while assimilating atom B with a device for measuring the photon scattered by A, to confirm the interpretation of $\mathcal{E}(\mathbf{r}, t)$ as a detection amplitude of a photon at point $\mathbf{r}_B$. The second goal is to study the time dependence of the scattering process itself.

We shall first study the spatial and temporal dependence of the scattered wave packet, in particular when the central frequency ω_m of the incident wave packet is close to the resonant frequency ω_0 of the scattering atom. We shall then compute the probability amplitude for the scattered wave packet to have excited at time t the atom B from its ground state a to its excited state b. As in § 1, we will associate with this amplitude a spatial wave packet describing the passage of the scattered wave packet by point $\boldsymbol{r}_B$.

4-a. Absorption amplitude by atom B of the photon scattered by atom A

We first consider a photon with a wave vector $\boldsymbol{k}_i$ parallel to the z axis, and a frequency $\omega_i = ck_i$. We are looking for the probability amplitude $\langle \boldsymbol{k}_f | S | \boldsymbol{k}_i \rangle$ for this photon to be scattered by atom A located at $\boldsymbol{r}_A$ from state $\boldsymbol{k}_i$ to state $\boldsymbol{k}_f$. This amplitude is given by relation (E-3) of Chapter XX, where we only take into account the resonant α processes (we assume the incident photon frequency to be fairly close to the atomic resonant frequency):

$$\begin{aligned}\langle \boldsymbol{k}_f | S | \boldsymbol{k}_i \rangle &= \langle \psi_{\text{fin}} | \bar{U}(\Delta t, 0) | \psi_{\text{in}} \rangle \\ &= -2\pi i \; A_{fi}^{\alpha}(E_{\text{in}}) \; \delta^{(\Delta t)}(E_{\text{fin}} - E_{\text{in}}) \equiv \mathcal{B}\, \delta^{(\Delta t)}(\omega_f - \omega_i)\end{aligned} \tag{41}$$

In this relation, $A_{fi}^{\alpha}(E_{\text{in}})$ is obtained from relation (E-4) in Chapter XX (here again we assume that the radiation is contained in a box of volume L^3):

$$A_{fi}^{\alpha}(E_{\text{in}}) = -\frac{\hbar\sqrt{\omega_i \omega_f}}{2\varepsilon_0 L^3} \frac{\langle a | \boldsymbol{\varepsilon}_f^* \cdot \boldsymbol{D} | c \rangle \langle c | \boldsymbol{\varepsilon}_i \cdot \boldsymbol{D} | a \rangle}{E_a + \hbar\omega_i - E_c} \; e^{i(\boldsymbol{k}_i - \boldsymbol{k}_f)\cdot \boldsymbol{r}_A} \tag{42}$$

where we have assumed that only one level c contributes, which explains why the sum over c has been suppressed; this is correct if the frequency of the radiation is close to the resonance frequency of one $a \leftrightarrow c$ transition, but far from all the other resonances. Note

that we have added to the right-hand side an exponential factor that comes from the spatial dependence of the electric field: in this complement, as in Chapter XX, we treat the atom's position $\boldsymbol{r}_A$ classically, but we no longer assume the atom to be placed at the coordinate origin. Expression (42) is a product of two matrix elements of the interaction Hamiltonian, one for the absorption of the $\boldsymbol{k}_i$ photon, the other for the emission of the $\boldsymbol{k}_f$ photon, divided by a common energy denominator. The function $\delta^{(\Delta t)}(\omega_f - \omega_i)$ simply expresses energy conservation, within $\hbar/\Delta t$, for the elastic scattering process, as was the case, for example, in §§ B-1-b and E-1-b of Chapter XX. We assume that the interaction time Δt is sufficiently long for this function to be assimilated to a real delta function $\delta(\omega_f - \omega_i)$.

The coefficient $\mathcal{B}$ introduced in the second equality (41) is proportional to expression (42); it contains the product of two matrix elements, which depends on the polar angles θ and φ of the vector $\boldsymbol{k}_f$ with respect to the direction of $\boldsymbol{k}_i$. We characterize this dependence by a function $f(k, \theta, \varphi)$, with:

$$k = |\boldsymbol{k}_i| = |\boldsymbol{k}_f| \tag{43}$$

As we assumed the frequency ω_i of the incident photon to be close to resonance, we can use the results of § C-2 in Chapter XX concerning resonant scattering. As in relation (E-11) of that chapter, we write the energy denominator in the form $\omega_i - \omega_0 + i\Gamma/2$, where Γ is the natural width of the excited state of the scattering atom A. Amplitude (41) then becomes:

$$\langle \boldsymbol{k}_f | S | \boldsymbol{k}_i \rangle = C_k \, \frac{f(k, \theta, \varphi)}{\omega_i - \omega_0 + i\Gamma/2} \, e^{i[(\boldsymbol{k}_i - \boldsymbol{k}_f)\cdot \boldsymbol{r}_A]} \, \delta(\omega_f - \omega_i) \tag{44}$$

where C_k is a coefficient proportional to k.

We now move to the next stage, the interaction of atom B with the $\boldsymbol{k}_f$ photon. As in (33), it is described by a matrix element (here again, we must add an exponential factor to account for the fact that atom B is not at the coordinate origin, but at point $\mathbf{r}_B$):

$$-\mathcal{D} e^{i\omega_0 t'} \langle 0 | E^{(+)}(\boldsymbol{r}_B, t') | a; \boldsymbol{k}_f \rangle \propto e^{i\boldsymbol{k}_f \cdot \boldsymbol{r}_B} e^{i(\omega_0 - \omega_f)t'} \tag{45}$$

We are now looking for the amplitude at time t of the complete process, scattering by A of the $\boldsymbol{k}_i$ photon, with an amplitude given by (44), and absorption by B, with an amplitude given by (45). Consequently, we multiply these two amplitudes and sum the product over all the possible $\boldsymbol{k}_f$ vectors for the scattered photon, and over the linear combination of states $\boldsymbol{k}_i$ forming the incident wave packet.

4-b. Wave packet scattered by atom A

To study the properties of the wave packet scattered by atom A, we successively carry out the two summations.

α. *Summation over all possible directions of the scattered photon*

Let us start with the summation over $\boldsymbol{k}_f$. Taking into account the function $\delta(\omega_f - \omega_i)$ appearing in (44), the summation over the modulus of $\boldsymbol{k}_f$ leads to:

$$k_f = k_i = k \tag{46}$$

Regrouping the $\boldsymbol{k}_f$ dependent terms in (44) and (45), we find that the summation over the directions of $\boldsymbol{k}_f$ introduces the angular integral:

$$\int \mathrm{d}\Omega \; f(k,\theta,\varphi) \; e^{i\boldsymbol{k}_f\cdot(\boldsymbol{r}_B-\boldsymbol{r}_A)} \tag{47}$$

The summation over the polar angles θ, φ of the exponentials describing the phase shift between $\boldsymbol{r}_A$ and $\boldsymbol{r}_B$ of all the plane waves $\boldsymbol{k}_f$ yields a spherical wave centered at $\boldsymbol{r}_A$:

$$\int \mathrm{d}\Omega \; f(k,\theta,\varphi) \; e^{i\boldsymbol{k}_f\cdot(\boldsymbol{r}_B-\boldsymbol{r}_A)} \propto f(k,\theta_B,\varphi_B)\frac{e^{ikr}}{r} \qquad \text{with} \qquad r=|\boldsymbol{r}_B-\boldsymbol{r}_A| \tag{48}$$

where θ_B and φ_B are the polar angles of vector $\boldsymbol{r}_B - \boldsymbol{r}_A$ with respect to the direction $\boldsymbol{k}_i$ of the incident photon. The right-hand side of (48) is reminiscent of a classical result in collision theory – see for example relation (B-12) of Chapter VIII. This means that the sum of all the plane waves scattered by atom A located at $\boldsymbol{r}_A$ has the structure of an outgoing spherical wave with the same wave number k as the $\boldsymbol{k}_f$ waves it is composed of. The amplitude of this spherical wave varies as $1/r$, which ensures that the outgoing energy across a sphere of radius r and surface $4\pi r^2$ does not depend on r.

The fact that the polar angles θ_B and φ_B appearing on the right-hand side of (48) are those of vector $\boldsymbol{r}_B - \boldsymbol{r}_A$ can be understood by stationary phase arguments. The phase factor $e^{i\boldsymbol{k}_f\cdot(\boldsymbol{r}_B-\boldsymbol{r}_A)}$ associated with the scattered wave $\boldsymbol{k}_f$ is equal to $e^{ikr\cos\alpha_f}$, where α_f is the angle between $\boldsymbol{k}_f$ and $\boldsymbol{r}_B - \boldsymbol{r}_A$. Since $kr \gg 1$, this phase factor has a very rapid variation with α_f, except in the vicinity of points where $\cos\alpha_f$ is stationary with respect to α_f, i.e. when $\alpha_f = 0$ for the outgoing wave. The angular integral (47) therefore gets most of its contribution from values of the angles that are close to the polar angles θ_B and φ_B of vector $\boldsymbol{r}_B - \boldsymbol{r}_A$.

Taking (48) into account, we deduce from (44) and(45):

$$\sum_{\boldsymbol{k}_f}\langle b;0|\bar{H}_I(t')|a;\boldsymbol{k}_f\rangle\langle\boldsymbol{k}_f|S|\boldsymbol{k}_i\rangle \propto f(k,\theta_B,\varphi_B)\frac{e^{ikr}}{r}\,\frac{1}{\omega-\omega_0+i\Gamma/2}\,e^{i\boldsymbol{k}_i\cdot\boldsymbol{r}_A}\,e^{i(\omega_0-\omega)t'} \tag{49}$$

where we have replaced ω_i and ω_f by ω.

β. Summation over the energies

The initial state of the scattering process is a superposition of states $|\boldsymbol{k}_i\rangle$ multiplied by $c(\boldsymbol{k}_i)$. We consider a one-dimensional wave packet, propagating along the Oz axis. With a proper choice of the coordinate origin, we can assume atom A is located at $\boldsymbol{r} = 0$, which amounts to replacing $\boldsymbol{r}_A$ by $\boldsymbol{0}$. As before, we assume $c(k)$ is real, so that at $t = 0$, the wave packet is centered at the position $\boldsymbol{r} = 0$ of the scattering atom A.

We now multiply (49) by $c(k)$, integrate over k, and over the time t' from $-\infty$ to t. The amplitude of the absorption by atom B, at time t, of the photon scattered by atom A is therefore proportional to:

$$\int_{-\infty}^{t}\mathrm{d}t' \; e^{i\omega_0 t'}\int \mathrm{d}\omega \; c(\omega)\omega \; f(\omega,\theta_B,\varphi_B)\frac{e^{i\omega r/c}}{r}\,\frac{1}{\omega-\omega_0+i\Gamma/2}\,e^{-i\omega t'} \tag{50}$$

where we have replaced the integral variable k by $\omega = ck$.

Let us compare (50) and (35). The integral over ω appearing in the integral over t' in (50) can be interpreted as the classical field scattered at time t' and at a distance r from point O along an axis with polar angles θ_B, φ_B:

$$\mathcal{E}_{\text{diff}}(r, \theta_B, \varphi_B, t') \propto \int \mathrm{d}\omega \; c(\omega)\omega \; f(\omega, \theta_B, \varphi_B) \frac{e^{i\omega r/c}}{r} \; \frac{1}{\omega - \omega_0 + i\Gamma/2} \, e^{-i\omega t'} \tag{51}$$

It will be useful for what follows to regroup the two exponentials of (50) and to set:

$$\tilde{t} = t' - r/c \tag{52}$$

We then get:

$$\mathcal{E}_{\text{diff}}(r, \theta_B, \varphi_B, t') = \int \mathrm{d}\omega \; c(\omega)\omega \; f(\omega, \theta_B, \varphi_B) \frac{1}{r} \; \frac{1}{\omega - \omega_0 + i\Gamma/2} \, e^{-i\omega \tilde{t}} \tag{53}$$

which means that the wave packet scattered along the direction θ_B, φ_B moves at velocity c, and that its amplitude decreases as $1/r$.

γ. *Spatial and temporal dependence of the scattered wave packet*

We now assume that the frequency width $\Delta\omega$ of the incident wave packet is much smaller than its average frequency ω_m:

$$\Delta\omega \ll \omega_m \tag{54}$$

but we do not make any hypothesis concerning the relative values of $\Delta\omega$ and Γ. The factor $\sqrt{\omega}$ appearing in the previous relations can now be replaced by $\sqrt{\omega_m}$, and comes out of the integral. We can also neglect the variations with ω of $f(\omega, \theta_B, \varphi_B)$ over the Γ interval where the function $(\omega - \omega_0 + i\Gamma/2)^{-1}$ varies significantly. The scattered field $\mathcal{E}_{\text{diff}}(r, \tilde{t})$ can then been seen as the temporal Fourier transform of the product of two functions $c(\omega)$ and $(\omega - \omega_0 + i\Gamma/2)^{-1}$. This field is the convolution of the Fourier transforms of these two functions of ω. Taking into account (36) and (37), the Fourier transform of the first function is:

$$c(\omega) \quad \underset{FT}{\longleftrightarrow} \quad \mathcal{E}(\tilde{t}) = e^{-i\omega_m \tilde{t}} \; \mathcal{F}(\tilde{t}) \tag{55}$$

For the second function, we get:

$$\frac{1}{\omega - \omega_0 + i\Gamma/2} \quad \underset{FT}{\longleftrightarrow} \quad e^{-i\omega_0 \tilde{t}} \; \theta(\tilde{t}) \; e^{-\Gamma \tilde{t}/2} \tag{56}$$

where $\theta(\tilde{t})$ is the Heaviside function, equal to 1 for $\tilde{t} > 0$, and to 0 for $\tilde{t} < 0$. This leads to:

$$\mathcal{E}_{\text{diff}}(r, \tilde{t}) \propto \frac{1}{r} \int_{-\infty}^{+\infty} \mathrm{d}\tau \; e^{-i\omega_m \tau} \; \mathcal{F}(\tau) \; e^{-i\omega_0 (\tilde{t} - \tau)} \; \theta(\tilde{t} - \tau) \; e^{-\Gamma(\tilde{t} - \tau)/2} \tag{57}$$

δ. Study of two limiting cases

Two interesting cases occur when the width $\Delta\omega$ of the incident wave packet is either very large or very small compared to the natural width Γ of the excited state of atom A.

$\Delta\omega \gg \Gamma$ ***limit***

The incident wave packet passes through a given point in a time $1/\Delta\omega$ that is very short compared to the radiative lifetime $1/\Gamma$ of the excited state. The envelope $\mathcal{F}(\tau)$ of the incident wave packet is different from zero only during a time interval much shorter than the characteristic times of the Fourier transform of $(\omega - \omega_0 + i\Gamma/2)^{-1}$. We can thus set $\tau = 0$ in the last two terms of (57), which yields:

$$\mathcal{E}_{\text{diff}}(r,\tilde{t}) \propto \frac{1}{r}\left[\int_{-\infty}^{+\infty} \mathrm{d}\tau\ e^{i(\omega_0-\omega_m)\tau}\ \mathcal{F}(\tau)\right]\ \left[e^{-i\omega_0\tilde{t}}\ \theta(\tilde{t})\ e^{-\Gamma\tilde{t}/2)}\right] \tag{58}$$

The term in the first bracket is proportional to the excitation amplitude of the scattering atom by the incident wave packet. The second bracket describes a free oscillation at the atomic frequency ω_0, starting at time $\tilde{t} = 0$ and damped over a time $2/\Gamma$.

The physical meaning of this result is as follows. The incident wave packet spends a very short time near atom A, and hence excites it in a percussive manner before moving away with velocity c. Once the incident wave packet is gone, the atomic dipole thus excited oscillates freely at frequency ω_0, until it is damped by spontaneous emission. This situation is the analog of the percussive excitation of an oscillator in classical mechanics.

$\Delta\omega \ll \Gamma$ ***limit***

We can now replace in (57) the function $\mathcal{F}(\tau)$ by $\mathcal{F}(\tilde{t})$ as $\tilde{t} - \tau$ cannot be larger, in modulus, than $1/\Gamma$. This is because of the presence of the last exponential term in (57) and the fact that, when $\Delta\omega \ll \Gamma$, the envelope of $\tilde{\mathcal{E}}(\tilde{t})$ varies very slowly over that time interval. One can then rewrite (57) in the form:

$$\mathcal{E}_{\text{diff}}(r,\tilde{t}) \propto \mathcal{F}(\tilde{t})e^{-i\omega_m\tilde{t}}\int_{-\infty}^{+\infty} \mathrm{d}\tau\ e^{i\omega_m(\tilde{t}-\tau)}e^{-i\omega_0(\tilde{t}-\tau)}\ \theta(\tilde{t}-\tau)\ e^{-\Gamma(\tilde{t}-\tau)/2)} \tag{59}$$

Let us make the change of variable $\tau' = \tilde{t} - \tau$ in the integral over τ. Taking (56) into account, we see that this integral is actually the Fourier transform of $e^{-i\omega_0\tau'}\theta(\tau')e^{-\Gamma\tau'/2}$ calculated at ω_m, which is $(\omega_m - \omega_0 + i\Gamma/2)^{-1}$. This leads to:

$$\mathcal{E}_{\text{diff}}(r,\tilde{t})\ \propto\ \mathcal{F}(\tilde{t})\, e^{-i\omega_m\tilde{t}}\ \frac{1}{\omega_m - \omega_0 + i\Gamma/2} \tag{60}$$

The physical meaning of this result is as follows. When $\Delta\omega \ll \Gamma$, the wave packet takes a long time passing atom A, whose dipole undergoes forced oscillation at the frequency ω_m. It thus emits radiation at the same frequency, with an amplitude that follows adiabatically the slow variation of the envelope $\mathcal{F}(\tilde{t})$ of the incident wave packet; this explains the first term in (60). The second term describes the linear response of the dipole with eigenfrequency ω_0 and damping time $2/\Gamma$ to an excitation of frequency ω_m. In this case, the oscillator's amplitude follows adiabatically the excitation's amplitude.

5. Example of wave packets with two entangled photons

In § 2-c, we considered two-photon states that were tensor products of two one-photon wave packets; the two photons described by these states were not entangled. There obviously exist a number of two-photon states that cannot be described in the form of a product of two one-photon states, and which thus describe entangled photons. In this last section, we shall focus on such an example where the entangled photons appear in an optical nonlinear process, called *parametric down-conversion*. This process has the advantage of producing pairs of photons bunched in time. Detecting one photon of the pair at a given time t allows predicting the second photon will be detected to within a very short time interval.

5-a. Parametric down-conversion

Computations involved in parametric down-conversion are similar to computations we already discussed. We shall simply outline the general ideas allowing a physical understanding of the process, without going into details that would unnecessarily lengthen the present complement.

α. Description of the process

In § E-1 of Chapter XX, we studied the elastic scattering of a photon by an atom. Figures 2 and 3 of Chapter XX show two possible diagrams representing such a process: an incident photon with angular frequency ω_i is absorbed and an ω_f photon emitted while the atom goes back to its initial level. Energy conservation of the total system atom + photon requires that $\omega_i = \omega_f$. In the present complement, we study a nonlinear scattering process during which, as before, an incident photon of angular frequency ω_0 is absorbed by an atom in state a, but where there are now two photons with angular frequencies ω_1 and ω_2 that are emitted. At the end of the scattering process the atom returns to state a; energy conservation now requires that $\omega_0 = \omega_1 + \omega_2$. Figure 1 gives two possible representations of such a process, analogous to those of Figure 2.α and 3.α in Chapter XX.

Several temporal orders are possible for the absorption and emission processes. For example, Figures 3.α and 3.β of Chapter XX do not have the same temporal order for the absorption and emission occurring in the scattering of one photon. For the three-photon process considered here, including one absorption and two emissions, $3! = 6$ possible temporal orders should, a priori, be considered; Figure 1 represents only one of these six possible orders.

β. Scattering amplitude

The principle for calculating the scattering amplitude of parametric down-conversion is similar to the one that led us to formulas (E-3) to (E-5) of Chapter XX, but involves now three, instead of two, interactions with the field; two relay states (instead of one) come into play. As an example, for the process represented in Figure 1, we must consider the following states:

- initial state: $|a;\omega_0\rangle$, with energy $E_{\rm in} = E_a + \hbar\omega_0$
- first relay state: $|b;0\rangle$, with energy $E_{\rm rel\ 1} = E_b$

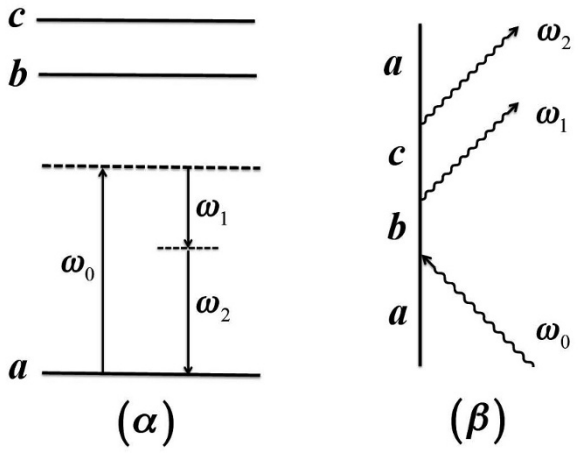


Figure 1: An incident photon, with angular frequency ω_0 is scattered by an atomic system in the initial state a. At the end of the scattering process, the atomic system has returned to state a, while two new photons have appeared with angular frequencies ω_1 and ω_2. Energy conservation requires that $\omega_0 = \omega_1 + \omega_2$. In the left hand side of the figure, the absorption (emission) processes are represented with upwards (downwards) arrows; in the right-hand side, these processes are shown with incoming (outgoing) wiggly arrows that also symbolize the photon propagation.

- second relay state: $|c;\omega_1\rangle$, with energy $E_{\text{rel }2} = E_c + \hbar\omega_1$
- final state: $|a;\omega_1,\omega_2\rangle$, with energy $E_{\text{fin}} = E_a + \hbar\omega_1 + \hbar\omega_2$

The probability amplitude associated with this process is obtained by generalizing relation (E-4) of Chapter XX. Within a constant and non-significant factor, it is the product of a function $\delta(\omega_1+\omega_2-\omega_0)$, imposing energy conservation[7], by the following expression:

$$-\left(\frac{\hbar}{2\varepsilon_0 L^3}\right)^{3/2}\sqrt{\omega_0\omega_1\omega_2}\sum_{b,c}\frac{\langle a|\boldsymbol{\varepsilon}_2^*\cdot\boldsymbol{D}|c\rangle\langle c|\boldsymbol{\varepsilon}_1^*\cdot\boldsymbol{D}|b\rangle\langle b|\boldsymbol{\varepsilon}_0\cdot\boldsymbol{D}|a\rangle}{(E_a+\hbar\omega_2-E_c)(E_a+\hbar\omega_0-E_b)} \tag{61}$$

where $\boldsymbol{\varepsilon}_0$, $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$ are, respectively, the polarizations of the photons of frequencies ω_0, ω_1 and ω_2. Compared to relation (E-4) of Chapter XX, expression (61) now contains three (instead of two) matrix elements in the numerator, and two (instead of one) energy denominators containing the differences in energy between the initial state and either the relay 1 or the relay 2 state.

Six similar amplitudes can be written, generalizing equations (E-3) to (E-5) of Chapter XX and corresponding to different temporal orders for the absorption and emission processes. Once they have been added, one must also sum these amplitudes over all

[7] As in § 4-a, we assume the total interaction time Δt to be sufficiently long for the function $\delta^{(\Delta t)}$ to be assimilated to a real delta function.

the atomic relay states b and c. All the contributions to the total amplitude contain the same function $\delta(\omega_1 + \omega_2 - \omega_0)$.

The final state of the system "atom + radiation" at the end of the scattering process is the sum over ω_1 and ω_2 of the components thus obtained, with the condition $\omega_1 + \omega_2 = \omega_0$. It can be written:

$$|\Psi\rangle = |a\rangle \otimes |\psi_R\rangle \tag{62}$$

where:

$$|\psi_R\rangle = \sum_{\mathbf{k}_1,\mathbf{k}_2} \delta(\omega_1 + \omega_2 - \omega_0)\ g(\mathbf{k}_1, \mathbf{k}_2)|\mathbf{k}_1, \mathbf{k}_2\rangle \tag{63}$$

with $\omega_{1,2} = c\,|\mathbf{k}_{1,2}|$. This state cannot be written as the product of two field states; it is therefore entangled (see Chapter XXI).

The function $g(\mathbf{k}_1, \mathbf{k}_2)$ characterizing the field state is the result of the ω_1 and ω_2 dependence of the scattering amplitudes, as well as of the density of final states appearing in the summations over the continuums[8] ω_1 and ω_2 (summation over the moduli of the two vectors $\mathbf{k}_1$ and $\mathbf{k}_2$). We assume here that the energies of all the relay states are far away from any resonance, so that the $g(\mathbf{k}_1, \mathbf{k}_2)$ dependence on $|\mathbf{k}_1|$ and $|\mathbf{k}_2|$ does not present any kind of narrow structure. In other words, all the energy differences $\Delta E = E_{\text{in}} - E_{\text{rel}}$ in the denominators of the scattering amplitudes are of the order of a fraction of an optical frequency. We mentioned in comment (i) at the end of § 1 in Complement A_{XX} that the time spent in a relay state during the scattering process is, according to the time-energy uncertainty relation, of the order of $\hbar/|\Delta E|$. This means that the times separating the emission of the two photons ω_1 and ω_2 cannot differ by more than a few optical periods, i.e. a few tens of femtoseconds. This qualitative argument shows that the two photons ω_1 and ω_2 are emitted quasi-simultaneously.

Comment:

If the interaction Hamiltonian appearing in the three matrix elements of the scattering amplitude is the electric dipole Hamiltonian, and if the atomic states have a well defined parity, this atomic parity changes with each interaction. After the three interactions, the parity is therefore changed, which forbids the final atomic state to be the same as the initial state. Consequently, the parametric down-conversion process we just studied can occur only when the atomic states do not have a well defined parity. Such a situation is encountered when the atomic Hamiltonian is not invariant upon reflection. This happens for example when the atom is inserted in a crystal where the local crystalline field, which has the symmetry of an external electric field, is not invariant upon spatial reflection.

5-b. Temporal correlations between the two photons generated in parametric down-conversion

We now compute the double detection signal from the two photons generated in parametric down-conversion. The experiment we analyze is schematized in Figure 2. An incident pump beam, with frequency ω_0, propagates along a direction with unit vector

[8] Two continuums of final states come into play in this problem, but the condition $\omega_1 + \omega_2 = \omega_0$ reduces it to one.

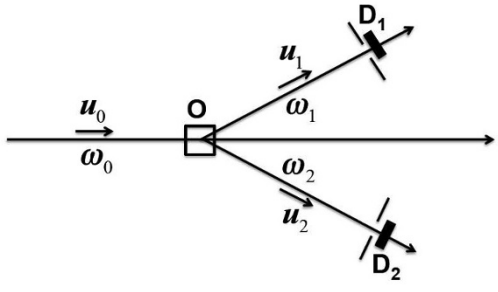


Figure 2: A pump beam with angular frequency ω_0, propagating along a direction with unit vector $\boldsymbol{u}_0$, impinges on a nonlinear crystal placed in O. The parametric down-conversion process generates two beams of frequencies ω_1 and ω_2, with $\omega_1 + \omega_2 = \omega_0$. Diaphragms allow fixing the directions $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ of these two beams. The two detectors D_1 and D_2 register the arrivals of the photons and permit studying their temporal correlations.

$\boldsymbol{u}_0$, and impinges onto a nonlinear crystal O containing atoms performing the conversion. Two diaphragms placed in front of the two detectors D_1 and D_2 allow selecting two directions, with unit vectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, for the two beams generated by parametric down-conversion.

We focus on the temporal, rather than spatial, aspect of the phenomenon. For the sake of simplicity, we assume that the three field states appearing in Figure 2 are plane waves, infinite in the two transverse directions. The only variable characterizing the modes involved is therefore the longitudinal component of the vector $\mathbf{k}$ or, equivalently, the frequency ω. The incident photon is described by a wave packet, characterized in frequency space by a real function $c(\omega_0)$, centered at ω_m^0 and of width $\Delta\omega_0 \ll \omega_m^0$. The center of the incident wave packet arrives at crystal O at time $t = 0$, and passes through it in a time of the order of:

$$\Delta t \simeq 1/c\Delta k = 1/\Delta\omega_0 \tag{64}$$

The two-photon wave packet generated by parametric down-conversion is described by an expression similar to (63), in which we now use the variables $\omega_{1,2} = c\,|\mathbf{k}_{1,2}|$; this wave packet depends, to a certain extent, on ω_1 and ω_2 via the function $g(\omega_1, \omega_2)$.

α. *Double photodetection signal $w_{II}(\boldsymbol{r}_1, t; \boldsymbol{r}_2, t+\tau)$*

We first compute the absorption amplitude of the two photons[9], one at time t by detector D_1 located at $\boldsymbol{r}_1$, the other at time $t+\tau$ by detector D_2 located at $\boldsymbol{r}_2$:

$$\langle 0|E^{(+)}(\boldsymbol{r}_2, t+\tau)E^{(+)}(\boldsymbol{r}_1, t)|\psi_R\rangle = \frac{\hbar}{2\varepsilon_0 L^3}\sum_{\omega_0}\sum_{\omega_1,\omega_2} c(\omega_0)\ \sqrt{\omega_1\omega_2}\ g(\omega_1,\omega_2)$$
$$\times\, e^{i[\boldsymbol{k}_2\cdot\boldsymbol{r}_2-\omega_2(t+\tau)]}\ e^{i[\boldsymbol{k}_1\cdot\boldsymbol{r}_1-\omega_1 t]}\ \delta(\omega_1+\omega_2-\omega_0) \tag{65}$$

[9]The signal w_{II} is the squared modulus of this amplitude.

In this equation, $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ are the wave vectors of the two photons propagating freely along $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. We note L_1 and L_2 the distances between O and D_1, O and D_2, and define $T_1 = L_1/c$ and $T_2 = L_2/c$ as the times taken by the photons to travel these two distances. We have:

$$\begin{aligned} \boldsymbol{k}_1 \cdot \boldsymbol{r}_1 &= \frac{\omega_1}{c} L_1 = \omega_1 T_1 \\ \boldsymbol{k}_2 \cdot \boldsymbol{r}_2 &= \frac{\omega_2}{c} L_2 = \omega_2 T_2 \end{aligned} \tag{66}$$

so that (65) can be rewritten in the form:

$$\begin{aligned} \langle 0|E^{(+)}(\boldsymbol{r}_2, t+\tau)E^{(+)}(\boldsymbol{r}_1, t)|\psi_R\rangle = \frac{\hbar}{2\varepsilon_0 L^3} \sum_{\omega_0} \sum_{\omega_1,\omega_2} c(\omega_0)\ \sqrt{\omega_1\omega_2}\ g(\omega_1,\omega_2) \\ \times\, e^{i\omega_2[T_2-t-\tau]} e^{i\omega_1[T_1-t]}\ \delta(\omega_1+\omega_2-\omega_0) \end{aligned} \tag{67}$$

We now replace the two variables ω_1 and ω_2 by a single variable $\delta\omega$ by setting:

$$\begin{aligned} \omega_1 &= \frac{\omega_0}{2} + \delta\omega \\ \omega_2 &= \frac{\omega_0}{2} - \delta\omega \end{aligned} \tag{68}$$

Condition $\omega_1 + \omega_2 = \omega_0$ is then automatically satisfied, so that the delta function appearing on the second line of (67) is no longer necessary. The summation over ω_1 and ω_2 becomes a summation over $\delta\omega$, and the function $g(\omega_1, \omega_2)$ is replaced by a function $\widetilde{g}(\delta\omega)$. If we assume, for the sake of simplicity, that $T_1 = T_2 = T$, we finally obtain:

$$\begin{aligned} \langle 0|E^{(+)}(\boldsymbol{r}_2, t+\tau)E^{(+)}(\boldsymbol{r}_1, t)|\psi\rangle \propto \frac{\hbar}{2\varepsilon_0 L^3} \sum_{\omega_0} c(\omega_0)\ e^{i\omega_0[T-t-\tau]/2} e^{i\omega_0[T-t]/2} \\ \times \sum_{\delta\omega} \sqrt{(\frac{\omega_0}{2}+\delta\omega)(\frac{\omega_0}{2}-\delta\omega)}\ \widetilde{g}(\delta\omega)\ e^{i\delta\omega\ \tau} \end{aligned} \tag{69}$$

β. Discussion

The τ dependence of the double photodetection signal is given by the summation over $\delta\omega$ on the second line of (69). Going to the continuous limit, it is therefore the Fourier transform of $\sqrt{(\frac{\omega_0}{2}+\delta\omega)(\frac{\omega_0}{2}-\delta\omega)} \times \widetilde{g}(\delta\omega)$. Nevertheless, the product $\sqrt{(\frac{\omega_0}{2}+\delta\omega)(\frac{\omega_0}{2}-\delta\omega)}$ varies very slowly with $\delta\omega$ and can be taken as constant, as can be the state densities introduced when replacing the discrete summation by an integral. We also saw above (§ 5-a-β) that the variation of g as a function of ω_1 and ω_2, hence the variation of $\widetilde{g}(\delta\omega)$, is very slow as long as no resonant (or quasi-resonant) relay states are involved in the scattering process. We must thus take the Fourier transform of a function of $\delta\omega$ that has a very large width, of the order of a fraction of the optical frequency. This means that the double photodetection signal is different from zero only if the two photodetections are separated by a time interval of the order of a few optical periods. In other words, the two detections are always quasi-simultaneous.

Consider now the summation over ω_0 in the first line of (69). We are going to see that the t dependence of the signal involves time scales much longer than those

characterizing the variation with τ of the signal w_{II}. To show this, we replace τ by 0 on the right-hand side; after going to the continuous limit, we get:

$$\int \mathrm{d}\omega_0 \; c(\omega_0) \; e^{i\omega_0(T-t)} \tag{70}$$

which is the Fourier transform of the incident wave packet. This packet arrives at point O at $t = 0$, and for the entire packet to pass that point, it takes a certain time Δt; this time interval Δt is much longer, in general, than the time characterizing the τ dependence of signal w_{II}. Relation (70) thus indicates that both detectors yield (almost simultaneously) a signal at any time t within a time interval Δt centered around $t = T$; this time corresponds to the arrivals at D_1 and D_2 of the photons generated at O by the incident wave packet. But once a photon is detected by one of the detectors, the other photon is detected practically at the same instant by the other detector. Such a temporal correlation could not be predicted by a semiclassical treatment.

These results remain valid when the parametric down-conversion process is produced, not by a single incident photon described by a wave packet, but rather by a continuous laser excitation. The two beams generated by the parametric down-conversion process then contain a series of pairs of photons, that are detected at the same instant; they are referred to as *twin photons*.

Such *twin beams* can excite two-photon transitions in a much more efficient way than ordinary beams. This is because, in the absence of resonant relay states in the two-photon absorption process, an argument similar to that presented above shows that the two absorptions must be separated by a very short time interval (the two photons must interact quasi-simultaneously with the absorbing atom). The two incident photons must impinge on the atom at exactly the same time, which can be the case for twin beams (with ordinary beams, one can only observe two-photon absorptions due to accidental coincidences)

In practice, radiation parametric down-conversion is often performed, not on an isolated atom, but rather on atoms or molecules in a solid. It is then imperative to take into account the interference between beams generated in different parts of the solid, and identify the conditions for getting a constructive interference. The refractive optical index of the medium in which the beams propagate then plays an important role, which leads to the so called *phase matching condition*. This discussion, outside the scope of the present complement, is treated in detail in quantum optics books [65] [66].

Chapter XXI

Quantum entanglement, measurements, Bell's inequalities

A Introducing entanglement, goals of this chapter 2188

B Entangled states of two spin-1/2 systems 2190

B-1 Singlet state, reduced density matrices 2191

B-2 Correlations 2191

C Entanglement between more general systems 2193

C-1 Pure entangled states, notation 2193

C-2 Presence (or absence) of entanglement: Schmidt decomposition 2193

C-3 Characterization of entanglement: Schmidt rank 2196

D Ideal measurement and entangled states 2196

D-1 Ideal measurement scheme (von Neumann) 2196

D-2 Coupling with the environment, decoherence; "pointer states" 2199

D-3 Uniqueness of the measurement result 2201

E "Which path" experiment: can one determine the path followed by the photon in Young's double slit experiment? 2202

E-1 Entanglement between the photon states and the plate states 2203

E-2 Prediction of measurements performed on the photon 2204

F Entanglement, non-locality, Bell's theorem 2204

F-1 The EPR argument 2205

F-2 Bohr's reply, non-separability 2207

F-3 Bell's inequality 2208

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.

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

We discuss in this last chapter an essential concept of quantum mechanics, entanglement. This will highlight a number of aspects of quantum mechanics that have no equivalent in classical physics.

A. Introducing entanglement, goals of this chapter

Consider two physical systems A and B, each having a state space $\mathcal{E}_A$ and $\mathcal{E}_B$; they can be grouped into a total system $A+B$, whose state space is the tensor product $\mathcal{E}_A \otimes \mathcal{E}_B$. If we assume system A is described by a normalized quantum state $|\varphi_A\rangle$ belonging to $\mathcal{E}_A$, system B by a normalized quantum state $|\chi_B\rangle$ belonging to its space state $\mathcal{E}_B$, the ket $|\Phi\rangle$ describing the total system is the tensor product:

$$|\Phi\rangle = |\varphi_A\rangle \otimes |\chi_B\rangle \tag{A-1}$$

In this case, each of the three physical systems A, B, and $A + B$ is described by a state vector, which is the most precise possible description in quantum mechanics.

The situation is different when the state vector of the global system is no longer a simple product. Let us note $|\varphi_A\rangle$, $|\zeta_A\rangle$, .., orthonormalized quantum states belonging to the state space $\mathcal{E}_A$ of the first system, and $|\chi_B\rangle$, $|\xi_B\rangle$,..., orthonormalized quantum states belonging to the state space $\mathcal{E}_B$ of the second. We can then build products different from (A-1) for the state of the global system, for example:

$$|\Phi'\rangle = |\zeta_A\rangle \otimes |\xi_B\rangle \tag{A-2}$$

Now, we can also, in view of the superposition principle, form any linear combination $|\Psi\rangle$ of $|\Phi\rangle$ and $|\Phi'\rangle$, which will no longer be a simple product:

$$|\Psi\rangle = \alpha\,|\varphi_A\rangle \otimes |\chi_B\rangle + \beta\,|\zeta_A\rangle \otimes |\xi_B\rangle \tag{A-3}$$

In this relation, the complex coefficients α and β can take on any value, as long as they obey the normalization condition:

$$|\alpha|^2 + |\beta|^2 = 1 \tag{A-4}$$

We shall assume, however, that neither of these coefficients is zero so that (A-3) is not reduced to a simple product:

$$\alpha\beta \neq 0 \tag{A-5}$$

A state such as (A-3), which contains a coherent superposition of two (or more) components, each component being a product, is called an "entangled state". The general property associated with these states is called "quantum entanglement". It expresses the fact that the quantum state of each subsystem is, in a way, conditioned by the state of the other.

In Complement $\mathrm{E_{III}}$, we introduced the concept of a density operator, which provides a more general description of a physical system than a state vector. The density operator of the total physical system $A + B$, whose state vector is known, is simply the projector onto $|\Psi\rangle$:

$$\rho_{A+B} = |\Psi\rangle\,\langle\Psi| \tag{A-6}$$

whose trace is equal to one:

$$\mathrm{Tr}\left\{\rho_{A+B}\right\} = \langle\Psi\,|\Psi\rangle = 1 \tag{A-7}$$

When a physical system can be described by a state vector, it is said to be in a "pure state". Its density operator obeys the relation:

$$\left[\rho_{A+B}\right]^2 = \rho_{A+B} \tag{A-8}$$

and thus:

$$\mathrm{Tr}\left\{\left[\rho_{A+B}\right]^2\right\} = 1 \tag{A-9}$$

Under such conditions, we can choose to describe the total physical system either by its state vector $|\Psi\rangle$, or by the density operator ρ_{A+B}. We are going to show that this is no longer the case for the two subsystems A and B, for which only the density operator can be used.

Imagine, for example, that we are only interested in measurements performed on subsystem A. We saw in § 5-b of Complement $\mathrm{E_{III}}$ that, when the total system is entangled as in (A-3), there generally does not exist any state vector belonging to $\mathcal{E}_A$ that allows computing the probabilities of measurements performed solely on A. Instead, one must necessarily use a density operator ρ_A obtained by taking a partial trace (taken over the state space $\mathcal{E}_B$ of the non-observed system – we recall in § B-1 below how to compute the matrix elements of a partial trace):

$$\rho_A = \mathrm{Tr}_B\left\{\rho_{A+B}\right\} \tag{A-10}$$

Like operator (A-6), this operator is Hermitian, non-negative, and its trace is equal to one; it is however not the projector onto a single state vector. When the system is described by the entangled state (A-3), this density operator is given by:

$$\rho_A = |\alpha|^2\,|\varphi_A\rangle\,\langle\varphi_A| + |\beta|^2\,|\zeta_A\rangle\,\langle\zeta_A| \tag{A-11}$$

which is actually the sum of two projectors. Consequently, subsystem A is in state $|\varphi_A\rangle$ with a probability $|\alpha|^2$, and in state $|\zeta_A\rangle$ with a probability $|\beta|^2$: as opposed to the state of $A+B$, the quantum state of A is not known with certainty, but only with a certain probability. The density operator ρ_A can be called a "statistical mixture", underlying the fact that the results of measurements performed on A are predicted by computing averages on the (non-observed) properties of B. We then get the inequality:

$$\mathrm{Tr}\left\{\left[\rho_A\right]^2\right\} \leq 1 \tag{A-12}$$

where the equality occurs only if one of the two coefficients α or β is zero; the equivalent of relation (A-9) is, in general, not satisfied by ρ_A. Inequality (A-12) expresses the fact that, as the quantum state of A is known only in a statistical way, the quantum description of A is less precise than the description of the total system $A+B$. This discussion can be easily generalized to the case where $|\Psi\rangle$ is the superposition of, not only two components as in (A-3), but of three or more.

We find ourselves in a situation that might look rather surprising, as it does not have any equivalent in classical physics. We do know that a perfect classical description

of the total system $A + B$ automatically implies that each of its two subsystems is also perfectly described. This is because a complete description of the state of the total system is simply the collection of all the complete descriptions of its subsystems. As an example, a perfect classical description of the solar system is simply the knowledge of all the positions and velocities of the planets, satellites, and all their constituent particles. In quantum mechanics, things are drastically different: the most precise description of the total system by a state vector (pure state) does not imply, in general, that its subsystems can be described with the same precision. This difference radically changes the usual relation between the parts and the whole of a physical system. Schrödinger, who first introduced in 1935 the words "quantum entanglement" commented on this new concept [67]: "As far as I am concerned, I would not call this property *one* but rather *the* characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought. By the interaction, the two representatives [the quantum states] have become entangled... Another way of expressing the peculiar situation is: the best possible knowledge of a whole does not necessarily include the best possible knowledge of all its parts, even though they may be entirely separate and therefore virtually capable of being 'best possibly known' i.e., of possessing, each of them, a representative (state vector) of its own."

In a general way, when the state vector of a global system is not a simple product, and quantum entanglement occurs, the quantum predictions for observations on part of the system can become rather unexpected. This chapter will discuss a number of special physical effects related to quantum entanglement. As a general introduction, § B studies the simple case of two spin-1/2 systems entangled in a singlet state. This example is generalized, in § C, to any physical system, and we present some of the properties of entangled quantum states. The relations between entanglement and quantum measurements is discussed in § D, using in particular the ideal measurement scheme proposed by von Neumann. In § E, we describe an experiment where one tries to observe interference fringes of a particle going through a two-slit plate while determining at the same time which slit the particle went through; if this were possible, one would face a contradiction. However, a partial trace operation on an entangled state of the particle and the plate allows proving the coherence of the quantum formalism and illustrates an aspect of complementarity. Finally, § F discusses the relations between entanglement and quantum non-locality, in the framework of the general Einstein, Podolsky and Rosen argument, and of Bell's theorem.

B. Entangled states of two spin-1/2 systems

We first discuss a very simple case, which will prove useful for the rest of the chapter: each of the two systems A and B is a spin 1/2; each state space $\mathcal{E}_{A,B}$ is then spanned by the two eigenstates $|\pm\rangle$ of the spin component on the Oz axis. We assume these two states are entangled in a singlet state, such as the one written in relation (B-22) of Chapter X:

$$\begin{aligned} |\Psi\rangle &= \frac{1}{\sqrt{2}}\Big[\,|A:+\rangle \otimes |B:-\rangle - |A:-\rangle \otimes |B:+\rangle\Big] \\ &= \frac{1}{\sqrt{2}}\Big[\,|+,-\rangle - |-,+\rangle\Big] \end{aligned} \tag{B-1}$$

(on the second line, we have simplified the notation, assuming that the first index in the ket refers to spin A, and the second to spin B).

B-1. Singlet state, reduced density matrices

In the basis of the 4 kets $|+,+\rangle$, $|+,-\rangle$, $|-,+\rangle$, $|-,-\rangle$ taken in that order, the matrix representing the density operator ρ_{A+B} is written:

$$(\rho_{A+B}) = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{B-2}$$

It is easy to check, performing a matrix product, that $(\rho_{A+B})^2 = (\rho_{A+B})$, hence that relation (A-9) is verified: this means the total system is in a pure state.

As indicated in § 5-b of Complement E_{III}, the matrix representing the density operator ρ_A is obtained by taking a partial trace, i.e. by adding the matrix elements of (ρ_{A+B}) that are diagonal with respect to the quantum numbers of the second spin (this amounts to summing over the states of the non-observed spin):

$$\langle m_A|\,\rho_A\,|m'_A\rangle = \sum_{m_B} \langle m_A, m_B|\,\rho_{A+B}\,|m'_A, m_B\rangle \tag{B-3}$$

This leads to:

$$(\rho_A) = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{B-4}$$

We then get:

$$(\rho_A)^2 = \frac{1}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{B-5}$$

and thus $\mathrm{Tr}\left\{(\rho_A)^2\right\} = 1/2$; this means that spin A is not in a pure state. By symmetry, the same result would obviously be found for (ρ_B). Note that after taking the partial trace, all the non-diagonal elements (coherences) of (B-2) have completely disappeared. When they are considered as an isolated system, each of the two spins is in a "completely depolarized" state, and measuring its Oz spin component (or any of its spin component, for that matter) will yield the results + or − with the same probability 1/2. At the level of each individual spin, the minus sign that characterizes the entanglement of the state vector (B-1) becomes irrelevant; on the other hand, we are going to show that this entanglement yields very strong correlations between the results of measurements pertaining to both spins.

B-2. Correlations

Imagine now that we perform simultaneously measurements on both spins, the first one along a direction in the plane xOz, making an angle θ_A with the Oz axis, the second along a direction in that same plane, making an angle θ_B with Oz. The results

we are about to obtain will be important for the discussion of Bell's theorem in § F-3-a. Relations (A-22) of Chapter IV (taking the angle φ equal to zero and $\theta = \theta_{A,B}$) yield the expressions for the eigenvectors of the measurements in the state spaces $\mathcal{E}_A$ and $\mathcal{E}_B$:

$$\begin{aligned} |+\rangle_\theta &= \cos\frac{\theta}{2}\,|+\rangle + \sin\frac{\theta}{2}\,|-\rangle \\ |-\rangle_\theta &= -\sin\frac{\theta}{2}\,|+\rangle + \cos\frac{\theta}{2}\,|-\rangle \end{aligned} \tag{B-6}$$

In the space of the two-spin states, the ket corresponding to a double + result of the measurement is written:

$$\begin{aligned} |+\rangle_{\theta_A} \otimes |+\rangle_{\theta_B} = \cos\frac{\theta_A}{2}\cos\frac{\theta_B}{2}\,|+,+\rangle + \cos\frac{\theta_A}{2}\sin\frac{\theta_B}{2}\,|+,-\rangle \\ + \sin\frac{\theta_A}{2}\cos\frac{\theta_B}{2}\,|-,+\rangle + \sin\frac{\theta_A}{2}\sin\frac{\theta_B}{2}\,|-,-\rangle \end{aligned} \tag{B-7}$$

This means that, when the system is in the singlet state (B-1), the probability amplitude for obtaining this double result is:

$$\begin{aligned} \left[\langle+|_{\theta_A} \otimes \langle+|_{\theta_B}\right]|\Psi\rangle &= \frac{1}{\sqrt{2}}\left[\cos\frac{\theta_A}{2}\sin\frac{\theta_B}{2} - \sin\frac{\theta_A}{2}\cos\frac{\theta_B}{2}\right] \\ &= -\frac{1}{\sqrt{2}}\sin\frac{\theta_A-\theta_B}{2} \end{aligned} \tag{B-8}$$

The probability of the result $(+,+)$ when measuring the components of both spins along the θ_A and θ_B directions is therefore:

$$\mathcal{P}_{++}(\theta_A,\theta_B) = \frac{1}{2}\sin^2\frac{\theta_A-\theta_B}{2} \tag{B-9}$$

One can redo the same calculations for the three other possible pairs of results, $(+,-)$, $(-,+)$ and $(-,-)$. This does not present any difficulty but it is easier to note that changing θ_A into $\theta_A+\pi$ exchanges the two eigenkets of (B-6), and hence the results + and − for the first spin A; the same operation can be done with the second spin B. We then make these changes in (B-9) and get the probabilities for the 4 possible results in the form:

$$\begin{aligned} \mathcal{P}_{++}(\theta_A,\theta_B) &= \mathcal{P}_{--}(\theta_A,\theta_B) = \frac{1}{2}\sin^2\frac{\theta_A-\theta_B}{2} \\ \mathcal{P}_{+-}(\theta_A,\theta_B) &= \mathcal{P}_{-+}(\theta_A,\theta_B) = \frac{1}{2}\cos^2\frac{\theta_A-\theta_B}{2} \end{aligned} \tag{B-10}$$

When both spins are measured, strong correlations between the results appear[1]. These correlations are the direct consequence of the entanglement present in the singlet state vector (B-1).

[1]The probabilities cannot, in general, be factored. As an example, relation (B-10) shows that $\mathcal{P}_{++}/\mathcal{P}_{-+}$ is different from $\mathcal{P}_{+-}/\mathcal{P}_{--}$. This means that the ratio of the probabilities of obtaining for the first spin the result + or the result − depends on the state of the second spin, clearly showing correlations.

C. Entanglement between more general systems

The concept of entanglement is obviously not limited to the singlet state of two spin-1/2 particles. We now study how to characterize the presence of entanglement when the total system is in a pure state.

C-1. Pure entangled states, notation

We consider two quantum systems A and B belonging, respectively, to state spaces $\mathcal{E}_A$ (with dimension P) and $\mathcal{E}_B$ (with dimension Q). The normalized state vector $|\Psi\rangle$ describing the total system $A+B$ belongs to the tensor product space $\mathcal{E}_A \otimes \mathcal{E}_B$, with dimension PQ. Some of the states $|\Psi\rangle$ can be written as a simple product:

$$|\Psi\rangle = |\varphi_A\rangle \otimes |\varphi_B\rangle \tag{C-1}$$

where $|\varphi_A\rangle$ and $|\varphi_B\rangle$ are any normalized kets of $\mathcal{E}_A$ and $\mathcal{E}_B$, respectively. In such a case, the two physical subsystems A and B are not entangled; each of them, as well as the total state, can be described by a state vector (pure state). On the other hand, the majority of the states $|\Psi\rangle$ cannot be factored this way, and must necessarily be written as a sum of products (the singlet state studied above is such an example); the two subsystems A and B are then entangled.

It is not always obvious to guess from the expression of any given state vector $|\Psi'\rangle$ if it can actually be written as a simple tensor product. This ket has, in general, $P \times Q$ components, and is expressed as:

$$|\Psi'\rangle = \sum_{p=1}^{P} \sum_{q=1}^{Q} c_{p,q} \ |\chi_A^p\rangle \otimes |\xi_B^q\rangle \tag{C-2}$$

where the $|\chi_A^p\rangle$ as well as the $|\xi_B^q\rangle$ are orthonormalized kets. Now if we expand the kets $|\varphi_A\rangle$ and $|\varphi_B\rangle$, appearing in the tensor product (C-1), onto the kets $|\chi_A^p\rangle$ and $|\xi_B^q\rangle$ as

$$|\varphi_A\rangle = \sum_{p=1}^{P} x_p \ |\chi_A^p\rangle \qquad \text{and} \qquad |\varphi_B\rangle = \sum_{q=1}^{Q} y_q \ |\xi_B^q\rangle \tag{C-3}$$

we obtain for a ket $|\Psi\rangle$ of the type (C-1):

$$|\Psi\rangle = \sum_{p=1}^{P} \sum_{q=1}^{Q} x_p \ y_q \ |\chi_A^p\rangle \otimes |\xi_B^q\rangle \tag{C-4}$$

It is not obvious at all, just from the knowledge of the coefficients $c_{p,q}$ of $|\Psi'\rangle$, to know if they can be factored into an expression of this type, leading to a product as in (C-1). We present in the next section a systematic method for asserting if this factorization is possible and actually performing it.

C-2. Presence (or absence) of entanglement: Schmidt decomposition

It can be shown (see the demonstration below) that any pure state $|\Psi\rangle$ describing the ensemble of the two physical systems A and B can be written in the form:

$$|\Psi\rangle = \sum_i \sqrt{q_i} \, |u_i, w_i\rangle \tag{C-5}$$

where the $|u_i\rangle$ are a set of orthonormal vectors in the state space of the first system, and the $|w_i\rangle$ another set of orthonormal vectors in the second state space. This expression is the Schmidt decomposition of a pure state, also called the "biorthonormal decomposition".

Whether the state of the total system $|\Psi\rangle$ is entangled or not, it always has a corresponding density operator, written in (A-6). By taking partial traces, each of the subsystems can be described by the density operators:

$$\rho_A = \mathrm{Tr}_B \left\{\rho_{A+B}\right\} \quad ; \quad \rho_B = \mathrm{Tr}_A \left\{\rho_{A+B}\right\} \tag{C-6}$$

Performing these partial traces using (C-5), we get two symmetric expressions:

$$\rho_A = \sum_i q_i \, |u_i\rangle \, \langle u_i| \tag{C-7}$$

and:

$$\rho_B = \sum_i q_i \;\; |w_i\rangle \, \langle w_i| \tag{C-8}$$

This means that, when the total system is in a pure state, the two partial density operators always have the same eigenvalues[2]. In the particular case where they are all zero except one, each of the two subsystems is in a pure state and state $|\Psi\rangle$ can be factored: no entanglement exists in the total system. Most of the time, however, several eigenvalues are non-zero, in which case $(\rho_A)^2$ is obviously not equal to ρ_A, and the same is true for ρ_B; entanglement is then present in the pure state $|\Psi\rangle$.

Demonstration of relation (C-5):

As the two operators ρ_A and ρ_B are Hermitian, non-negative and have a trace equal to unity, their corresponding matrices can be diagonalized to yield real eigenvalues included between 0 and 1. We call $|u_i\rangle$ the normalized eigenvectors of ρ_A (the index i takes on P different values, where P is the dimension of subsystem A) and q_i the corresponding eigenvalues, all positive or zero (but not necessarily different. Similarly, the eigenvectors of ρ_B are noted $|v_l\rangle$ (where l takes on Q different values, Q being the dimension of the second subsystem B), and r_l the corresponding eigenvalues. The two partial density operators can then be expanded as:

$$\rho_A = \sum_{i=1}^{P} q_i \;\; |u_i\rangle \, \langle u_i| \quad \text{and} \quad \rho_B = \sum_{l=1}^{Q} r_l \;\; |v_l\rangle \, \langle v_l| \tag{C-9}$$

with $0 \leq q_i,\, r_l \leq 1$.

State $|\Psi\rangle$ can then be expanded on the basis of the tensor products $\{|A : u_i\rangle \otimes |B : v_l\rangle\}$, that we shall simply note $\{|u_i, v_l\rangle\}$ assuming the first ket represents a state of A and the second a state of B; we call $c_{i,l}$ the components of $|\Psi\rangle$ in this basis and get:

$$|\Psi\rangle = \sum_{i=1}^{P} \sum_{l=1}^{Q} c_{i,l} \, |u_i, v_l\rangle \tag{C-10}$$

[2]Note that this is not necessarily the case if the total system is described by a statistical mixture rather than a pure state. As an example, we can assume that ρ_{A+B} equals a tensor product $\rho_A \otimes \rho_B$, where ρ_A and ρ_B can be chosen arbitrarily, and hence have different eigenvalues.

We now introduce the ket $|\overline{w}_i\rangle$, belonging to the state space $\mathcal{E}_B$ (this ket is not necessarily normalized), as:

$$|\overline{w}_i\rangle = \sum_{l=1}^{Q} c_{i,l} \, |v_l\rangle \tag{C-11}$$

Expansion (C-10) for $|\Psi\rangle$ now simply becomes:

$$|\Psi\rangle = \sum_{i=1}^{P} |u_i, \overline{w}_i\rangle \tag{C-12}$$

We also know that the matrix elements of the partial trace ρ_A are:

$$\langle u_i| \, \rho_A \, |u_j\rangle = \sum_m \langle u_i, v_m \, |\Psi\rangle \, \langle\Psi| \, u_j, v_m\rangle \tag{C-13}$$

Now expression (C-12) for $|\Psi\rangle$ leads to:

$$|\Psi\rangle \, \langle\Psi| = \sum_{i',j'} |u_{i'}, \overline{w}_{i'}\rangle \, \langle u_{j'}, \overline{w}_{j'}| \tag{C-14}$$

Inserting this result into (C-13), we are only left with terms for which $i' = i$ and $j' = j$, which yields:

$$\langle u_i| \, \rho_A \, |u_j\rangle = \sum_m \langle v_m| \, \overline{w}_i\rangle \, \langle \overline{w}_j| \, v_m\rangle = \langle \overline{w}_j| \, \overline{w}_i\rangle \tag{C-15}$$

This means that:

$$\rho_A = \sum_{i,j} |u_i\rangle \, \langle u_i| \, \rho_A \, |u_j\rangle \, \langle u_j| = \sum_{i,j} \, |u_i\rangle \, \langle u_j| \times \langle \overline{w}_j \, |\overline{w}_i\rangle \tag{C-16}$$

Now in the basis $\{|u_i\rangle\}$ we have used, we know that ρ_A is diagonal and given by expression (C-9); the comparison with (C-16) shows that we must necessarily have:

$$\langle \overline{w}_j \, |\overline{w}_i\rangle = \delta_{i,j} \ \times q_i \tag{C-17}$$

For non-zero eigenvalues q_i, this relation shows that one can define a set of orthonormal vectors $|w_i\rangle$ belonging to the state space of system B as:

$$|w_i\rangle = \frac{1}{\sqrt{q_i}} \, |\overline{w}_i\rangle \tag{C-18}$$

For all the values of the index i associated with eigenvalues q_i equal to zero, that same relation shows that the kets $|\overline{w}_i\rangle$ are zero.

Replacing in (C-12) the $|\overline{w}_i\rangle$ by $\sqrt{q_i} \, |w_i\rangle$, we complete the demonstration of equality (C-5), and of relations (C-7) and (C-8) which follow directly.

C-3. Characterization of entanglement: Schmidt rank

The number of non-zero q_i eigenvalues, i.e. the number of non-zero terms in (C-5), is called the "Schmidt rank" of $|\Psi\rangle$ and noted R. When $R = 1$, the state of the total system is not entangled, the two subsystems each being in a pure state. When $R = 2$, we are in the case studied in § B where the two subsystems are entangled, when $R = 3$ we get a more complex entanglement, etc. This gives us a criterion for deciding whether a general ket (C-2) is entangled or not. We just have to compute a partial density operator for one of the subsystems and the number of its non-zero eigenvalues (for the sake of simplicity, we shall obviously choose the subsystem whose state space's dimension is the smallest). If that number equals 1, the eigenvector associated with that non-zero eigenvalue becomes one of the factors of the decomposition, which makes it easy to find the other. If that number is greater than 1, however, the decomposition into a single tensor product is no longer possible.

In a way, entanglement is symmetrically shared between A and B. For example, it is not possible for one of the two subsystems to be in a pure state and the other in a statistical mixture. The rank R must be lower than the smallest of the dimensions P and Q of the state spaces of A and B; to allow a high rank entanglement, the two subsystems must thus have state spaces with high enough dimensions.

Comment:

When all the q_i eigenvalues of ρ_A (and of ρ_B) are distinct, the Schmidt decomposition is unique. This is because decompositions (C-9) and (C-8) of ρ_B on the projectors onto its eigenvectors must then necessarily coincide; the set of eigenvectors $|w_i\rangle$ is identical to that of the $|v_l\rangle$. In this case, the eigenvectors of the partial density operators directly yield the unique Schmidt decomposition. On the other hand, when certain q_i eigenvalues are degenerate, this is no longer true. As an example, we saw that for the singlet state (B-1), the two partial density matrices have two eigenvalues, both equal to 1/2; this singlet state, decomposed in (B-1) into products of eigenvectors of the Oz spin components, can equally well be decomposed into products of eigenvectors of the spin components on any spatial direction. There are an infinity of possible Schmidt decompositions for that state.

D. Ideal measurement and entangled states

Entanglement also plays an essential role in any quantum measurement process, as it generally appears while the measured system S and the measuring apparatus M interact. Furthermore, we shall see that it even propagates further and brings the environment of the measuring apparatus into play.

D-1. Ideal measurement scheme (von Neumann)

Von Neumann's quantum measurement scheme proposes a general framework that allows characterizing the quantum measurement process in terms of entanglement appearing (or disappearing) in the state vector describing the total system $S + M$. The two systems S and M are initially described by a factored state $|\Psi_0\rangle$; however, as they

interact during a certain time, they reach an entangled state $|\Psi'\rangle$. After the measurement, we assume they no longer interact, imagining for example they have moved far away from each other.

In the state space of S, with dimension N_S, the physical quantity measured on S is described by an operator O_A whose normalized eigenvectors are the kets $|a_k\rangle$ with eigenvalues a_k (that we shall assume non-degenerate, to simplify the notation):

$$O_A\,|a_k\rangle = a_k\,|a_k\rangle \tag{D-1}$$

Initially, state $|\varphi_0\rangle$ of S is any linear combination of the $|a_k\rangle$:

$$|\varphi_0\rangle = \sum_{k=1}^{N_S} c_k\,|a_k\rangle \tag{D-2}$$

with complex coefficients c_k, having only the constraint that the sum of their squared moduli be equal to 1 (normalization condition). As for the measuring apparatus M, we assume it is, initially, always in the same normalized quantum state $|\Phi_0\rangle$. The initial state of the total system is then:

$$|\Psi_0\rangle = |\varphi_0\rangle \otimes |\Phi_0\rangle \tag{D-3}$$

D-1-a. Basic process

We start with the particular case where the measurement result is certain and where the system S is initially in one of the eigenstates associated with the measurement: $|\varphi_0\rangle = |a_k\rangle$. In that case:

$$|\Psi_0\rangle = |a_k\rangle \otimes |\Phi_0\rangle \tag{D-4}$$

Once the measurement is done, S stays in the same state $|a_k\rangle$, but the measuring apparatus reaches a state $|\Phi_k\rangle$ different from $|\Phi_0\rangle$ and which depends on k: this is a necessary condition for the result to be experimentally accessible. This is because the position of the "pointer" used for the reading of the result (a needle in a macroscopic apparatus, the recording of the result in a memory, etc.) must necessarily depend on k to allow for the acquisition of the data. It is also natural to assume that the different states $|\Phi_k\rangle$ are orthogonal to each other, since the pointer necessarily involves a large number of atoms whose different states will allow a macroscopic observer to read the result. The measurement process for the total system can be summed up, in this simple case, as follows:

$$|\Psi_0\rangle = |a_k\rangle \otimes |\Phi_0\rangle \implies |\Psi'\rangle = |a_k\rangle \otimes |\Phi'_k\rangle \tag{D-5}$$

where $|\Phi'_k\rangle$ is a normalized state of M. At this stage, no correlation or entanglement has appeared between the measuring apparatus and the measured system. This is what happens in the simple case where the measurement result is certain.

In the general case, the initial state of system S is a superposition (D-2) of eigenstates $|a_k\rangle$. In this case, state (D-4) must be replaced by the linear combination, with the same coefficients:

$$|\Psi_0\rangle = \sum_k c_k\ |a_k\rangle \otimes |\Phi_0\rangle \tag{D-6}$$

As Schrödinger's equation is linear, we get:

$$|\Psi_0\rangle \Longrightarrow |\Psi'\rangle = \sum_k c_k \ |a_k\rangle \otimes |\Phi'_k\rangle \tag{D-7}$$

which now becomes a state in which the measuring apparatus is entangled with the measured system S. The states of S and M are therefore strongly correlated: when the "pointer's position" is associated with one of the state vectors $|\Phi'_k\rangle$, the state of the system S must be described by the ket $|a_k\rangle$ associated with a definite eigenvalue of O_A.

After the measurement, one can no longer attribute a state vector (pure state) to the system S, which can only be described by a partial density operator. As the states $|\Phi'_k\rangle$ are normalized and orthogonal to each other, this density operator is given by:

$$\rho'_S = \mathrm{Tr}_M \left\{ |\Psi'\rangle \langle\Psi'| \right\} = \sum_k |c_k|^2 \ |a_k\rangle \langle a_k| \tag{D-8}$$

This relation was to be expected: it simply states that the system has a probability $|c_k|^2$ of being in the state $|a_k\rangle$ associated with the measurement result a_k, which is in agreement with the Born rule for the usual probabilities. This useful formula sums up in a simple way a number of characteristics of the quantum measurement postulate. Note, however, that at this stage, all the possible results are still present in the partial trace, as they are considered possible, even after the measurement. Nothing at this point tells us that only one result is actually measured when the experiment is performed, nor that the squares of the coefficients $|c_k|$ can be interpreted as classical probabilities associated with mutually exclusive observations. The evolution predicted by Schrödinger's equation cannot, by itself alone, explain the uniqueness of the results observed at the macroscopic level. This is why von Neumann introduced the postulate of the state vector's reduction (also called the "wave packet reduction" or "wave packet collapse", *cf.* Chapter III, § B-3-c); more detail on this point will be given in § D-3.

D-1-b. Dynamics of the entanglement process

A simple interaction Hamiltonian between systems S and M can explain the appearance of entanglement between these systems, and lead to relations (D-5) or (D-7). As an example, imagine this interaction Hamiltonian H_{int} can be written as:

$$H_{int} = g \ O_A \otimes P_M \tag{D-9}$$

where O_A is the operator (acting only on S) already introduced above, P_M an operator acting only on M, and g a coupling constant. We shall also assume that, in the state space of M, there exists a Hermitian conjugate operator X_M of the operator P_M:

$$[X_M, P_M] = i\hbar \tag{D-10}$$

This commutation relation means that P_M generates the translation operators with respect to X_M. In other words, the action of $e^{-i\Delta x P_M/\hbar}$ on any eigenvector $|x_M\rangle$ of X_M:

$$X_M \, |x_M\rangle = x_M \, |x_M\rangle \tag{D-11}$$

leads to a translation by Δx of the eigenvalue x_M:

$$e^{-i\Delta x P_M/\hbar} \, |x_M\rangle = |x_M + \Delta x\rangle \tag{D-12}$$

where Δx is any real number – see relation (13) of Complement $\mathrm{E_{II}}$.

We assume that $|\Phi_0\rangle$ (state of the measuring apparatus before the measurement) is a normalized eigenstate of X_M with eigenvalue x_0, and we ignore[3] any other source for the evolution of the total system other than the interaction between S and M. The evolution operator between the time $t = 0$ before the measurement, and the time $t = \tau$ when the interaction is over, is:

$$U(\tau, 0) = e^{-ig\tau O_A P_M/\hbar} \tag{D-13}$$

Its action on the ket (D-4) yields:

$$U(\tau, 0)\,|a_k\rangle \otimes |\Phi_0(x_0)\rangle = |a_k\rangle \otimes |\Phi_0(x_0 + g\tau a_k)\rangle \tag{D-14}$$

where the variables in parentheses in the states of the measuring apparatus[4] refer to the eigenvalues of X_M. This means that the states $|\Phi'_k\rangle$ introduced in (D-5) are the kets:

$$|\Phi'_k\rangle = |\Phi(x_0 + g\tau a_k)\rangle \tag{D-15}$$

These relations show that, as far as the measuring apparatus M is concerned, the eigenvalue of X has been shifted by a quantity $g\tau a_k$ that depends on the eigenvalue a_k of S. The observable X_M therefore plays the role of a "pointer's position" in the measuring apparatus (measuring needle), which yields the measurement result once the two systems have interacted. As for the observable O_A, it pertains to the system S being measured by the pointer's position.

If now the initial state is in a coherent superposition as in (D-6), the state after the interaction (D-7) is written:

$$|\Psi'\rangle = \sum_k c_k \;\; |a_k\rangle \otimes |\Phi(x_0 + g\tau a_k)\rangle \tag{D-16}$$

which is a biorthonormal decomposition such as the one obtained in § C-2. If, initially, the system S is not in an eigenstate of O_A, the interaction with the measuring apparatus changes its state into a statistical mixture (D-8). On the other hand, if the system is initially in an eigenstate of O_A, it will stay in the same eigenstate after the measurement: the measurement process does not change its state. The measurement is then said to be a "quantum non-demolition" measurement, or QND measurement.

D-2. Coupling with the environment, decoherence; "pointer states"

We now examine under which conditions the interaction and entanglement process we have considered constitutes a good measurement. A first obvious condition to be satisfied is for the states $|\Phi'_n\rangle$ of the measuring apparatus to store the information about

[3] To avoid this hypothesis, the computation could be performed within the interaction point of view (exercise 15 in Complement $\mathrm{L_{III}}$) with respect to free evolutions of both S and M; this would somewhat complicate the results. However, as we focus here on the dynamics induced by their mutual interaction, we shall keep the computations simple and assume that these free evolutions have a negligible effect during the duration τ of the interaction.

[4] Needless to say, a measuring apparatus is macroscopic and has many other degrees of freedom apart from the pointer's position. For the sake of simplicity, these other degrees of freedom have not been introduced in the notation.

the measurement result in a robust way, and prevent it from being destroyed as M continues to evolve on its own. This condition will be fulfilled if X_M is a constant of the motion of M or, in other words, if X_M commutes with the Hamiltonian H_M of M. In addition, the measuring apparatus cannot remain isolated from its environment, even at a microscopic level. In view of its function, the apparatus must be able to interact and be correlated with a measurement recording device, and even with the observer collecting the measurement result; it is, by definition, an "open" apparatus, which can interact with the outside world. In any case, for this apparatus to be completely isolated, would require that none of its atoms, electrons, etc., be in interaction and correlated with any of the environment particles, which is obviously impossible to achieve with a macroscopic apparatus.

This means that, as far as the coupling between M and the environment E is concerned, an entanglement phenomenon occurs, reminiscent of the one discussed for S and M. We must determine which basis, in the state space of M, will lead for the entangled state $M + E$ to a biorthonormal decomposition similar to (D-16). The computation of the partial density operator of M is similar to the computation that yielded (D-8) for the entanglement between S and M: it is in the basis of this biorthonormal decomposition that the partial density operator of M (which plays the role of S) remains diagonal; with another basis, the density matrix will have in general non-diagonal elements. As, furthermore, the entanglement continues to propagate further and further into the environment, it is necessary that the relevant basis of M remains constant in time. It is thus important to find this privileged basis.

Depending on the circumstances, the coupling between a measuring apparatus and its environment can take on various forms, in general complex due to the large number of degrees of freedom involved; several time constants come into play. Different models have been proposed to account for this coupling and the dynamics it produces. Without going into any details we shall make a few general remarks. The measurement process involves a whole chain of amplification between S and the macroscopic pointer, which can be composed of mesoscopic or macroscopic objects sensitive to the environment. Entanglement propagates along that chain via local interactions: the interaction potentials are diagonal in the position representation, and have a microscopic range. Consequently, they cannot couple quantum states corresponding to macroscopically different positions of the objects concerned; the branches of the state vector corresponding to different spatial positions propagate independently. This means that the coupling with the environment tends to favor the basis of states where the positions of the different elements of the measuring apparatus, including in particular its pointer, occupy well defined positions in space. The corresponding preferred basis in the state space of the measuring apparatus M, in which its density matrix remains diagonal over time, is called the basis of the "pointer states". In this basis, and only in this one, defined by pointer localization criteria, the entanglement with E is prone to destroy the coherences (non-diagonal elements of the density matrix), without changing the diagonal elements (meaning the positions of the pointer's particles).

To sum up, several conditions are necessary for a device to be considered as an acceptable measuring apparatus M for a physical quantity of S. In the first place, the coupling between S and M must be capable of transferring the right information from one to the other. The transferred information must then be conserved over time, while M continues its own evolution, and is coupled with the environment E. Obviously, these are

necessary conditions. In practice, an effective measuring apparatus must be conceived taking into account many other imperatives, such as high sensitivity, or strong protection against unavoidable external perturbation.

D-3. Uniqueness of the measurement result

As mentioned above, nothing in the dynamics associated with Schrödinger's equation can explain the uniqueness of the results observed at the macroscopic level. This is not surprising as (D-8) is a direct consequence of Schrödinger's equation, which is incapable of stopping on its own the endless propagation of the "von Neumann chain", as we shall now discuss.

D-3-a. The infinite von Neumann chain

Let us go back to the ideal measurement scheme of § D-1. After the measurement, the state of $S+M$ is the entangled state (D-7), a superposition of components associated with all the possible measurement results. One may wonder if, using a second measuring apparatus M_2 to observe M, one might be able to resolve this superposition and obtain a unique result. In fact, the same entanglement process that occurred between S and M will occur again, leading to a final state:

$$|\Psi''\rangle = \sum_k c_k \ |a_k\rangle \otimes |\Phi'_k\rangle \otimes |\Xi'_k\rangle \tag{D-17}$$

where the kets $|\Xi'_k\rangle$ represent the states of the second measuring apparatus M_2, orthogonal to each other for different values of k. Adding a third measuring apparatus M_3 will, obviously, only continue further the entanglement's progression, each additional apparatus playing the role of an environment for the previous one. This chain of measuring apparatus may continue all the way to infinity without permitting at any stage the resolution of the superposition, and the demonstration of the uniqueness of the measurement result. This is called the von Neumann chain (and the logical problem it poses is called the "von Neumann's infinite regress").

The well-known "Schrödinger's cat paradox" involves a similar situation. The system S is supposed to be a radioactive nucleus in a superposition of two states, $|a_1\rangle$ where the nucleus is still in the excited state, and $|a_2\rangle$ where it has disintegrated, emitting a particle. The kets $|\Phi_k\rangle$, $|\Xi_k\rangle$, etc. represent the states of the measuring apparatus that can detect this particle, and then trigger a mechanical system killing the cat in the case of positive detection. The last of these kets $|Z\rangle$ characterizes the cat, which can therefore be in state $|Z_1\rangle$ where it is still alive, or in state $|Z_2\rangle$ where it is dead. Schrödinger points out the absurdity of a physical description involving a cat that can be at the same time both in an alive and a dead state.

As we just discussed, the uniqueness of the measurement results cannot be proven with Schrödinger's equation; this equation merely predicts that the pointer of a measuring apparatus, and any other macroscopic object, can become superpositions of states located at points very far away in space. Because of the linearity of Schrödinger's equation, nothing prevents the different components of the state vector from propagating further and further away, without this infinite chain of entanglements ever reducing into a single one of its components. It is precisely to solve this problem that von Neumann introduced a specific postulate: the postulate of the reduction of the state vector (Chapter III (§ B-3-c) which "forces" the uniqueness of the measurement result.

D-3-b. Postulate of reduction of the state vector

The postulate of reduction (or collapse) of the state vector is also called the "projection postulate", or the "postulate of collapse of the wave packet". As we saw in Chapter III (§ B-3-c, this postulate states that, once the measurement has been performed, one must suppress the summations appearing in (D-7), (D-8), and (D-17): one only keeps, among all the terms, the component $k = m$ corresponding to the measurement result actually observed. After the measurement, the state vector becomes again a simple product from which entanglement has disappeared; S is once again in a pure state. This means that the entanglement, initially created by the measurement, disappears once the result has been recorded.

This postulate, as efficient as it may be, is somewhat difficult to interpret. Using this postulate amounts to considering that the state vector can evolve under the influence of two different processes: a "normal" continuous evolution, obeying Schrödinger's differential equation, and a sudden discontinuous evolution upon measurement, governed by the von Neumann reduction postulate. Obviously, this duality immediately introduces the question of the limit between these two evolutions: from which time on, exactly, should we consider that the measurement has been performed? In other words, how far does the coherent superposition (D-17) propagate? Which physical processes constitute a measurement, as opposed to those leading to a continuous Schrödinger evolution? These difficult questions were the motivation for introducing other interpretations of quantum mechanics. As an example, there are "non-standard" interpretations where Schrödinger's equation is modified by the adjunction of a small stochastic term. This term is chosen so as to be totally negligible at the microscopic level, while coming into play at a certain macroscopic level; its role is to suppress all the macroscopically different components of the state vector, except one. Both Schrödinger and von Neumann dynamics are then unified into a single equation for the evolution of the state vector. Many other interpretations have been proposed: additional variables, modal interpretation, Everett, all suggesting different solutions for the problem. The interested reader may consult reference [68].

E. "Which path" experiment: can one determine the path followed by the photon in Young's double slit experiment?

Let us now return to a question already discussed in Complement D_I. In Young's double slit experiment where the photon may follow two different paths to reach the detection screen, is it possible to observe interference fringes between these paths and simultaneously obtain information as to which path the photon followed? Figure 1 of Complement D_I, reproduced here in the above Figure 1, shows an interference experimental set-up using a plate pierced with two slits F_1 and F_2; this plate is mobile in a direction perpendicular to the incident photon. As it receives momentum transfers Δp_1 and Δp_2 that will be different, depending on whether the photon goes through F_1 or F_2, one could naively imagine observing interference while knowing through which slit the particle went through. However, using the momentum-position uncertainty relations applied to this mobile plate, we showed that the interference fringes were blurred as soon as the momentum transfers Δp_1 and Δp_2 were sufficiently different to provide this information. The reason is that if we want to be able to distinguish these two momentum transfers, the momentum uncertainty of the mobile plate must be less than the modulus of $\Delta p_1 - \Delta p_2$. A simple calculation then shows that when this condition is met, the uncertainty in the

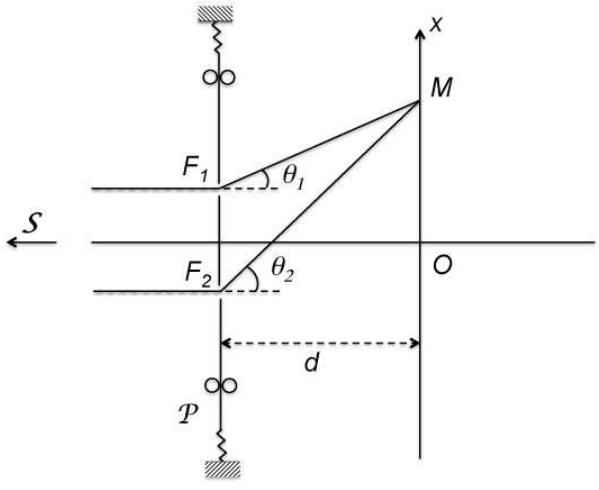


Figure 1: Young's double slit experiment using a plate $\mathcal{P}$, mobile along the x axis, and pierced with two slits F_1 and F_2. A photon, emitted by a source $\mathcal{S}$ assumed to be far away at infinity, reaches the detection screen at point M. The x component of the momentum transferred by the photon to the plate $\mathcal{P}$ depends on whether it goes through slit F_1 or slit F_2.

position of the plate must necessarily be larger than the fringe spacing, which blurs out the fringes. It is impossible to know which of the slits the photon went through without destroying at the same time the interference pattern.

We shall take the analysis a step further and consider the entanglement between the plate and the paths followed by the photon. This should allow us to envisage intermediary situations where partial information on the particle's path can be obtained.

E-1. Entanglement between the photon states and the plate states

Consider the path F_1M followed by the photon if it goes through F_1 and arrives at M on the detection screen (Figure 1). We call $|\psi_1\rangle$ the photon state when it follows that path and transfers a momentum Δp_1 to the mobile plate as it goes through F_1. In that case, after the photon's transit, the state of the plate is:

$$|E_1\rangle = \exp(i\Delta p_1 X/\hbar)|E_0\rangle \tag{E-1}$$

where $|E_0\rangle$ is the initial state of the plate and $\exp(i\Delta p_1 X)$ the momentum space translation operator, by a quantity Δp_1. The state of the global system photon + plate, along the path F_1M, is therefore $|\psi_1\rangle \otimes |E_1\rangle$. A similar reasoning would yield the result $|\psi_2\rangle \otimes |E_2\rangle$ along the path F_2M. As Schrödinger's equation is linear, the state of the global system after the photon has crossed the plate is:

$$|\Psi\rangle = |\psi_1\rangle \otimes |E_1\rangle + |\psi_2\rangle \otimes |E_2\rangle \tag{E-2}$$

which clearly shows the entanglement between the photon[5] and the plate.

E-2. Prediction of measurements performed on the photon

Measurements performed only on the photon after its crossing the plate can be predicted from the reduced density operator ρ, which is the partial trace over the plate variables of the global system density operator $|\Psi\rangle\langle\Psi|$. The matrix elements of ρ are obtained via the standard calculation of a partial trace (Complement $\mathrm{E_{III}}$, § 5-b) and lead to the operator expression:

$$\rho = |\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2| + |\psi_1\rangle\langle\psi_2|\langle E_2|E_1\rangle + |\psi_2\rangle\langle\psi_1|\langle E_1|E_2\rangle \tag{E-3}$$

(in this equation we did not write the factors of $|\psi_1\rangle\langle\psi_1|$ and $|\psi_2\rangle\langle\psi_2|$, which are the scalar products $\langle E_1|E_1\rangle$ and $\langle E_2|E_2\rangle$, both equal to 1 if the state $|E_0\rangle$ is normalized). The interference between the two paths is described by the terms in $|\psi_1\rangle\langle\psi_2|$ and $|\psi_2\rangle\langle\psi_1|$, which are multiplied by the scalar products $\langle E_2|E_1\rangle$ and $\langle E_1|E_2\rangle$.

Two extreme cases then appear. If the two states $|E_1\rangle$ and $|E_2\rangle$ are very close to each other, the two scalar products are practically equal to 1 and the interference terms in (E-3) are barely modified by the presence of the factors $\langle E_2|E_1\rangle$ and $\langle E_1|E_2\rangle$: the interference is then quite visible on the detection screen. In that case, however, the states $|E_1\rangle$ and $|E_2\rangle$ are too close to give any information as to whether the photon went through F_1 or F_2. In the other extreme case where $|E_1\rangle$ and $|E_2\rangle$ are very different from each other, their scalar product is practically zero: the interference terms disappear from (E-3), but one can, in principle, determine which of the two slits the photon went through. The present calculation allows studying intermediate situations where the scalar products $\langle E_2|E_1\rangle$ and $\langle E_1|E_2\rangle$ take on values included between 0 and 1. They describe how the contrast of the fringes diminishes when $\langle E_2|E_1\rangle$ and $\langle E_1|E_2\rangle$ continuously decrease from 1 to 0.

Actually, these scalar products can easily be computed from (E-1) and the equivalent relation for $|E_2\rangle$. This leads to:

$$\langle E_2|E_1\rangle = \langle E_0|\exp\left[i(\Delta p_1 - \Delta p_2)x\right]|E_0\rangle \tag{E-4}$$

Using the expression of $\Delta p_1 - \Delta p_2$ (noted $p_1 - p_2$ in Complement $\mathrm{D_I}$) and equations (6) and (7) of that complement, we can show that $\langle E_2|E_1\rangle$ and $\langle E_1|E_2\rangle$ are equal to the overlap integrals between the plate initial wave function, and that same wave function translated in momentum space by the amount h/a where a is the fringe spacing.

F. Entanglement, non-locality, Bell's theorem

We now present two important theorems, the EPR (for Einstein, Podolsky and Rosen) theorem, and Bell's theorem, which are related, the second actually being a logical continuation of the first. The EPR theorem was presented in an article published in 1935 by these three authors [69], and is one of the episodes of the famous discussion between Einstein and Bohr concerning the foundations of quantum mechanics (in particular during the Solvay conferences). Einstein's position was that the entire physical world had

[5] All the conclusions of this section remain valid for Young's interference type experiments performed with a massive particle instead of a photon.

to be expressed in the general framework of relativity, where the concept of space-time events is fundamental. Bohr had a different point of view, and considered that quantum theory demanded abandoning a description of microscopic events in space-time terms, while of course conforming to the actual predictions of relativity.

F-1. The EPR argument

The EPR theorem can be stated as follows: "*If all the predictions of quantum mechanics are correct (even for systems made of several remote particles) and if physical reality can be described in a local (or separable) way, then quantum mechanics is necessarily incomplete: some 'elements of reality' exist in Nature that are ignored by this theory* ".

To demonstrate their theorem, Einstein, Podolsky and Rosen imagined an experiment where two physical systems, originating for example from a common source S and described by an entangled quantum state, are then measured in remote regions of space. Historically, EPR developed their argument for correlated particles whose position and momentum are measured. It is however simpler to present an equivalent version of the argument concerning spins and discrete results, a version initially proposed by Bohm (and often called for that reason EPRB).

F-1-a. Exposing the argument

Imagine that two spin 1/2 particles are emitted by a source S in a singlet state (B-1), which is an entangled state where the spins are strongly correlated. The particles then move towards two remote regions of space, without their spins interacting with the outside world; the initial spin entanglement remains unchanged. In these remote regions of space, the particle spin components are measured along a direction defined by angle a for the region on the left, and by angle b for the region on the right (Fig. 2). One often calls Alice and Bob the two observers who perform the measurements in the two different laboratories, which can be very far away from each other. Alice chooses the direction a freely, which defines her "measurement type". With a spin 1/2, she can only obtain two results, that we will note $+1$ or -1, whatever measurement type was chosen. In a similar way, Bob chooses the direction b arbitrarily and obtains one of the two results $+1$ or -1. In the thought EPRB experiment, one assumes for simplicity that the two spins, once they have been emitted by the source, will only interact with the measuring apparatus (without having any free evolution, as was the case above). Standard quantum mechanics then predicts (§ B) that the distances and instants at which the measurements are performed do not play any role in the probability of obtaining the different possible double results.

To keep things simple, let us assume Alice and Bob limit their choices to a finite number of directions a and b for their respective measurements. It may then happen, by chance, that their chosen directions are parallel. Now if the angles a and b are chosen to be equal (parallel measurement directions), relations (B-10) indicate that the results will be always opposite for the two measurements: each time Alice observes ± 1, Bob observes the opposite value ∓ 1. This remains valid even if the measurements are performed at points greatly separated in space, whatever the choice $a = b$, and even if the two observers operate in totally independent ways, in their own regions of space; for example, they could make their choice at the last moment, even after the emission and

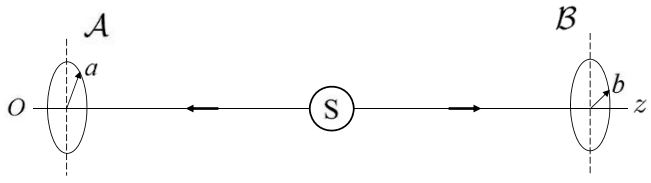


Figure 2: In an EPRB experiment, a source S emits pairs of particles in a singlet spin state (entangled quantum state). These particles then propagate along the Oz axis, towards two remote regions of space $\mathcal{A}$ and $\mathcal{B}$, where Stern-Gerlach apparatus are used by observers Alice and Bob to measure the components of their spin along directions perpendicular to Oz. For the first particle, the measurement direction is defined by angle a, and for the second, by angle b. Each measurement yields a result $+1$ or -1, and one looks for correlations between these results when the experiment is repeated a great number of times.

propagation of the pair of particles.

Let us assume Alice performs her measurement along a direction a before Bob starts doing his own measurement. When Alice finishes her measurement, it becomes certain that, should Bob decide to chose a direction b parallel to a, he will observe the opposite result; the result is certain in that particular case. Such a certainty can only come from the fact that the particle measured by Bob possesses a physical property that determines this certain result; this property (called "element of reality" by EPR) will influence the way that particle interacts with the measuring apparatus in B and will determine the result. On the other hand, the particle propagating towards Bob cannot be influenced by events occurring in Alice's laboratory. This means that this physical property we are discussing existed before the measurement performed by Alice.

The reasoning is obviously symmetrical and establishes that, before any measurement, the particles already possessed physical properties that determined the outcome of the future measurements. As the direction a chosen by Alice was random, it means that the particles must possess enough properties to determine the results for any analysis directions chosen by the observers. Now quantum mechanics does not predict the existence of such properties, as it only gives a description of the particles via a singlet state vector, which always predicts a totally random result for the first measurement. Furthermore, there exists no quantum state for which all the spin components on arbitrary directions can be simultaneously determined (the corresponding operators do not commute with each other). This means that quantum mechanics accounts only partially for the physical properties of the system; it is therefore incomplete.

F-1-b. Assumptions and conclusions

Let us discuss in more detail the logical structure of the EPR argument.

(i) It starts by assuming that the predictions of quantum mechanics for the proba-

bilities of measurement results are correct. The argument thus assumes that the perfect correlations predicted by this theory are always observed, whatever the distance between the two measuring apparatus.

(ii) Another essential ingredient of the EPR argument is the concept of "elements of reality" defined with the following criterion [69]: "if, without in any way disturbing a system, we can predict with certainty the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity". In other words, a certainty cannot be built on nothing: an experimental result known beforehand can only be the consequence of a preexisting physical quantity.

(iii) And last, but not least, the EPR argument brings in the notion of space-time and locality: the elements of reality they discuss are attached to regions of space where the experiments take place, and they cannot suddenly vary (and certainly not appear) under the influence of events occurring in a very distant region of space. Einstein wrote in 1948 [70]: "Physical objects are thought of as arranged in a space-time continuum. An essential aspect of this arrangement of things in physics is that they lay claim, at a certain time, to an existence independent of one another, provided these objects are situated in different parts of space". To sum up, one can say that the basic conviction of EPR is that regions of space contain their own elements of reality (attributing distinct elements of reality to separated regions of space is sometimes called "separability"), and that their time evolution is local – one often refers to "local realism" in the literature to qualify the ensemble of the EPR hypotheses.

Basing their argument on these hypotheses, EPR show that, for any chosen values of a and b, the measurement results are functions:

(i) of the individual properties of the spins the particles carry with them (the EPR elements of reality);

(ii) and of the orientations a, b of the Stern and Gerlach analyzers (which is obvious).

It follows that the results are given by well defined functions of these variables, meaning that no non-deterministic process occurs: a particle with spin brings along all the necessary information to yield the result of a future measurement, whatever the choice of the orientation a (for the first particle) or b (for the second). This implies that all the components of each spin must have simultaneously well determined values.

F-2. Bohr's reply, non-separability

Bohr rapidly replied [71] to the EPR article presenting their argument. In Bohr's view, the only physical system to be considered is the entire experimental set-up, including the measured quantum system and all the measuring apparatus, which are treated classically. It is thus meaningless to try and select among this ensemble subsystems having individual physical properties. The physical system Bohr considers is a whole that one should not attempt to separate into parts. This is often called the "non-separability" rule. In other words, Bohr considers that spatial separation does not lead to separability.

It is not the EPR reasoning that Bohr criticizes, but he considers that their starting assumptions are not relevant in the framework of quantum physics. From Bohr's point of view, the EPR criterion for elements of reality "contains an essential ambiguity when applied to quantum phenomena". Along the same line, more than ten years later (in 1948), Bohr made his point of view explicit [72]: "Recapitulating, the impossibility of subdividing the individual quantum effects and of separating a behavior of the objects

from their interactions with the measuring instrument serving to define the conditions under which the phenomena appear implies an ambiguity in assigning conventional attributes to atomic objects, which calls for a reconsideration of our attitude towards the problem of physical explanation ". It is thus the very need for a physical explanation involving such a subdivision that is questioned by Bohr.

Bohr refutes Einstein's basic idea, namely that one can attribute distinct physical properties to two objects located in very remote space-time regions. He believes that "quantum non-separability" applies, even in such a situation. It is understandable that Einstein was unwilling to abandon concepts that are the pillars of special and general relativity (gravitation).

F-3. Bell's inequality

In 1964, more than thirty years after the publishing of the EPR argument, an article by Bell shed an entirely new light on the question [73]. This article, in a way, took up the EPR argument from the point at which its authors had left it. Taking at their face value the existence of the EPR elements of reality, and using the same local realism considerations, Bell showed that there is actually no way to complete quantum mechanics without changing its predictions, at least in some cases. This means that one must either accept that certain predictions of quantum mechanics are sometimes incorrect, or abandon certain EPR hypotheses, however natural they may seem.

F-3-a. Bell's theorem

Following Bell's idea, let us assume that λ represents the "elements of reality" associated with the spins; λ is, actually, just a concise notation that could represent a multiple component vector, so that the number of elements of reality contained in λ is totally arbitrary. One can even include in λ components that play no particular role in the problem; the only important hypothesis is that λ must contain enough information to yield the results of all the possible spin measurements. For each pair of spins emitted in the course of the experiment, λ is fixed.

Another commonly used notation for the two measurement results is A and B, not to be confused with the small letters a and b used for the parameters of the two measuring apparatus. As expected, A and B are functions not only of λ, but also of the measurement parameters a and b. However, locality requires that b has no influence on result A (since the distance between the two measurements' locations is arbitrarily large); conversely, a has no influence on result B. We shall note $A(a,\lambda)$ and $B(b,\lambda)$ the corresponding functions, which can take on two values, $+1$ or -1. Figure 3 schematizes the experiment we are discussing.

To establish Bell's theorem, it is sufficient to take into account only two directions for each individual measurement; we shall then use the simpler notation:

$$A \equiv A(a,\lambda) \qquad A' \equiv A(a',\lambda) \tag{F-1}$$

and:

$$B \equiv B(b,\lambda) \qquad B' \equiv B(b',\lambda) \tag{F-2}$$

For each emitted pair of particles, as λ is fixed, the four above numbers have well defined values, which can only be ± 1.

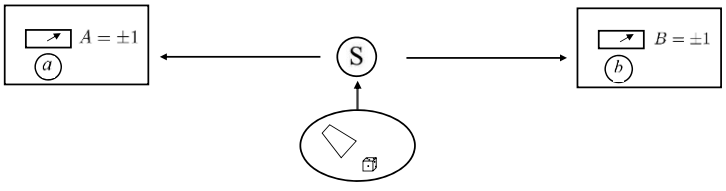


Figure 3: Source S emits particles toward two measuring apparatus located far away from each other, each being set up with its own measurement parameter, respectively a and b; each apparatus yields a ±1 result. The oval under the source symbolizes a fluctuating random process, which controls the particles' emission process, and hence their properties. Correlations between the measured results are observed; they are due to the common random properties the particles have acquired upon their emission by the fluctuating process.

Consider then the sum of products:

$$M(\lambda) = AB - AB' + A'B + A'B' \tag{F-3}$$

that can also be written as:

$$M(\lambda) = A(B - B') + A'(B + B') \tag{F-4}$$

If $B = B'$, the above expression reduces to $2A'B = \pm 2$; if $B = -B'$, it reduces to $2AB = \pm 2$. In both cases, we see that $M = \pm 2$.

If we now take the average value $\overline{M}$ of $M(\lambda)$ over a large number of emitted pairs (average over λ), we get:

$$\overline{M} = \overline{AB}_\lambda - \overline{AB'}_\lambda + \overline{A'B}_\lambda + \overline{A'B'}_\lambda \tag{F-5}$$

where $\overline{AB}_\lambda$ denotes the average value over λ of the product $AB \equiv A(a,\lambda)B(b,\lambda)$, and a similar notation has been used for the 3 other terms. As each $M(\lambda)$ value can only be ± 2, we necessarily have:

$$-2 \leq \overline{M} \leq +2 \tag{F-6}$$

This result is the so called BCHSH (Bell, Clauser, Horne, Shimony et Holt) form of Bell's theorem. This inequality must be satisfied by all sorts of measurements yielding random results, whatever mechanism creates the correlations, as long as the locality condition is obeyed: A does not depend on the measurement parameter b, and B does not depend on a.

This means that any theory that fits in the framework of "local realism" must lead to predictions satisfying relation (F-6). Realism is necessary since we used in the demonstration the concept of EPR elements of reality to introduce the functions A and

B; locality is also essential as it forbids A to depend on b and, conversely, B to depend on a.

The simplicity of this demonstration is such that the inequality may be expected to remain valid in many situations. This actually happens any time the observed correlations can be explained by fluctuations having a past common cause; they are then referred to as "classical correlations". In such cases, each time the experiment is performed, this common cause is fixed, and the 4 numbers A, B, A' and B' take on well-defined values (even though they are, a priori, unknown) all equal to ± 1; the number M is therefore also well defined, equal to -2 or $+2$. Whatever the M values found in a random series of N measurements, it is mathematically impossible for the sum of these M values to be greater than $2N$ or smaller than $-2N$. Consequently, the average value obtained by dividing this sum by N necessarily obeys (F-6): the mere existence of these 4 numbers is sufficient to obtain the inequality.

F-3-b. Contradictions

It would seem natural for any reasonable physical theory to automatically lead to predictions satisfying this inequality. Now, surprisingly enough, this is not the case for quantum mechanics and, furthermore, this contradiction has been experimentally confirmed.

α. *Contradictions with quantum mechanics predictions*

Relations (B-10) allow computing the average value of the product of the ± 1 results obtained in the measurements of the two spins along directions making an angle θ_{ab} with each other. This average value is given by (we write $\hat{A}$ and $\hat{B}$ to emphasize that these letters now denote operators, not numbers):

$$\left\langle \hat{A}(a)\hat{B}(b) \right\rangle = \mathcal{P}_{+,+} \; + \; \mathcal{P}_{-,-} \; - \; \mathcal{P}_{+,-} \; - \; \mathcal{P}_{-,+} = -\cos\theta_{ab} \tag{F-7}$$

This expression is the quantum equivalent of the average value over the λ variable of the product of results $A(a,\lambda)B(b,\lambda)$ in a theory with local realism. To get the quantum equivalent $\langle Q \rangle$ of the combination of the four products of results as they appear in (F-3), we must compute the same combination of average values of these products of results, which yields:

$$\begin{aligned} \langle Q \rangle &= \left\langle \hat{A}(a)\hat{B}(b) \right\rangle - \left\langle \hat{A}(a)\hat{B}(b') \right\rangle + \left\langle \hat{A}(a')\hat{B}(b) \right\rangle + \left\langle \hat{A}(a')\hat{B}(b') \right\rangle \\ &= -\cos\theta_{ab} + \cos\theta_{ab'} - \cos\theta_{a'b} - \cos\theta_{a'b'} \end{aligned} \tag{F-8}$$

Imagine now that the four directions are in the same plane, and that the vectors, arranged in the order $\mathbf{a}$, $\mathbf{b}$, $\mathbf{a}'$ and $\mathbf{b}'$, all make a 45° angle with the preceding vector (Fig. 4); all the cosines are then equal to $1/\sqrt{2}$, except for $\cos\theta_{ab'}$ that is equal to $-1/\sqrt{2}$. We then get $\langle Q \rangle = -2\sqrt{2}$; exchanging the directions of $\mathbf{b}$ and $\mathbf{b}'$, we get $\langle Q \rangle = 2\sqrt{2}$. In both cases, BCHSH inequality (F-6) is violated by a factor $\sqrt{2}$, i.e. by more than 40 %. In spite of the seemingly simple cosine variation of expression (F-7), we just showed that no theory with local reality is able to account for it, as this would violate inequality (F-6). This means that the EPR-Bell argument leads to an important quantitative contradiction with quantum mechanics, proving it to be a theory that does not comply with local realism in the sense of EPR.

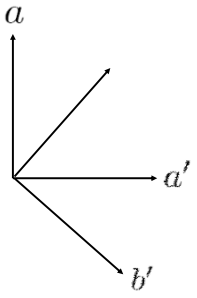


Figure 4: Position of the four vecteurs a, b, a' and b' leading to a maximum violation of BCHSH inequalities for two spin-1/2 particles in a singlet state. These vectors define the spin components to be measured, along a or a' for the spin on the left, and along b or b' for the spin on the right. This means that the entire experiment needs four different set-ups. The only pair of vectors leading to a negative correlation between the two measurement results is (a, b'), as the angle between the corresponding directions is larger than 90°.

How is it possible to encounter such a contradiction and how come such an apparently flawless argument does not apply to quantum mechanics? Several answers can be given:

(i) Had Bohr been aware of Bell's theorem, he would very likely have rejected the existence of the 4 preexisting numbers A, A', B, B'. If these numbers do not exist, the argument of § F-3-a is no longer possible and the BCHSH inequality disappears. Bohr would have considered Bell's theorem as mathematically correct in probability theory, but totally irrelevant in quantum mechanics, as being improper for the quantum description of the experiment under study.

Even if he had accepted to reason about these numbers, as unknown variables to be determined later as is often the case in algebra, would the inequality have survived? The answer is no, still reasoning with Bohr's logic. As already mentioned in § F-2, Bohr's point of view is that the entire experiment must be considered as a whole. One cannot distinguish two separate measurements that would be performed, each on one of the particles: the only true measurement process concerns the ensemble of both particles together. A fundamentally indeterminist and delocalized process occurs in the whole region of space containing the entire experiment.

The functions A and B then both depend on both the measurement parameters, and must be written as $A(a, b)$ and $B(a, b)$; this immediately forces an abandon of locality. Instead of the 2 numbers A and A', we now have 4 numbers, $A = A(a, b)$, $A' = A(a', b)$, as well as $A'' = A(a, b')$ and $A''' = A(a', b')$; the same is true for B and B', which must be replaced by 4 numbers. We now must deal with a total of 8 numbers instead of 4. The demonstration of the BCHSH inequality is then no longer possible and the contradiction

disappears.

(ii) One may prefer a more local point of view for the measurement process, which allows keeping the concept of a measurement on a single particle. To avoid the contradiction with quantum mechanical predictions, one must then consider that it is meaningless to attribute four well defined values A, $A^{'}$, B, $B^{'}$ to each pair. Since only a maximum of two of them can be measured in a given experiment, we should not be able to talk about these four numbers or argue about them even as unknown quantities. A well know phrase of Péres [75] very clearly sums up this point of view: "unperformed experiments have no results". Wheeler [76] expresses the same idea as he writes: "no phenomenon is a real phenomenon until it is an observed phenomenon.".

β. Contradictions with experimental results

The question was: does Bell's theorem allows pointing out very particular situations where quantum mechanics is no longer valid? Or, on the contrary, are the predictions of quantum mechanics always valid, which immediately entails that certain hypotheses leading to the inequalities must be abandoned? A great number of experiments have been performed from 1972 on; they all confirmed the predictions of quantum mechanics, measuring, sometimes with great precision [77], the violation of Bell's inequalities.

After a moment of doubt, it now seems well established that quantum mechanics yields perfectly correct predictions, even in situations where it implies a violation of Bell's inequalities. However plausible they might look, one must abandon at least one of the hypotheses that led to these inequalities.

Conclusion

The concept of quantum entanglement is quite essential; it leads to situations where certain types of correlations, totally impossible in classical physics, can be produced and observed. These situations can occur even when the observations are performed in regions of space arbitrarily remote from one another. A fundamental idea of quantum mechanics, without any classical counterpart, is that the most precise description of a whole does not necessarily entail an equivalently precise description of its parts. This means that there exists no theory both local and realistic, for describing a system containing two remote and entangled particles (it would contradict quantum mechanics).

Entanglement also plays an essential role in the measurement processes, and comes into play at different levels: entanglement between the measured system S and the measuring apparatus M, between M and the environment E, between two environments E and E', and so forth. We also discussed how entanglement determines the contrast of the fringes observed in an interference experiment where a particle has to cross a plate pierced with two holes.

In addition to these important aspects, entanglement also plays an essential role in quantum computing: one seeks to take advantage of the parallel evolution of the various entangled branches of the state vector to perform computations. This domain of research has undergone intense development in recent years, but is too extensive to be treated in the present volume. The reader may want to consult specialized books on the subject, as for example that of D.Mermin [78]. Entanglement also plays a central role in quantum cryptography, whose aim is to fabricate devices for secure quantum key distribution that

cannot be spied on, as any eavesdropping is detectable; a review on this subject can be found in the article by N.Gisin, G.Ribordy, W.Tittel and H.Zbinden [79].

There still remains the fact that, in the presence of entanglement, and in particular during a measurement process, the standard interpretation of quantum mechanics may present some difficulties. Schrödinger's evolution equation does not predict the uniqueness of the measurement result observed in the macroscopic world. To obtain this uniqueness in the framework of the theory, one can introduce an ad hoc postulate, such as the von Neumann postulate of reduction of the state vector. It then raises the question of where to set the border: when exactly should one stop using the continuous evolution of Schrödinger's equation and impose the reduction of the state vector? How can one reconcile the intrinsic irreversibility of this ad hoc postulate with the reversibility of Schrödinger's equation?

Another open question concerns the status of the state vector. We have used it throughout this book as a mathematical tool, good for computing probabilities, but what does it really represent? Does it directly describe physical reality? Or does it simply give information about physical reality? A number of quantum mechanics interpretations have been proposed (see reference [68]) that discuss this fundamental difficulty.

A_{XXI} : DENSITY OPERATOR AND CORRELATIONS; SEPARABILITY	A_{XXI} : This complement introduces von Neumann statistical entropy associated with a density operator, discussing its properties and establishing some important inequalities it must satisfy. Also discussed are the differences between classical and quantum correlations (arising from quantum entanglement effects). The concept of "quantum non-separability" is introduced.
B_{XXI} : GHZ STATES; ENTANGLEMENT SWAPPING	GHZ states provide an example of conflict between quantum mechanics and the usual concept of local realism. The contradiction is even stronger than for Bell's inequalities, as it is expressed as an opposition in signs. Entanglement swapping allows entangling particles without them ever having to interact with each other.
C_{XXI} : MEASUREMENT INDUCED RELATIVE PHASE BETWEEN TWO CONDENSATES	When two Bose-Einstein condensates overlap, their relative phase is a priori totally undetermined. However, such a phase may appear, induced by a measurement process sensitive to that phase. As measurements proceed, this relative phase will progressively acquire a more precise value.
D_{XXI} : EMERGENCE OF A RELATIVE PHASE WITH SPIN CONDENSATES; MACROSCOPIC NONLOCALITY AND THE EPR ARGUMENT	This complement is an extension of the previous one, studying the case where the two condensates are formed of particles with spins. The same phenomenon occurs: the emergence of a relative phase, but in a context where the EPR argument is harder to refute because of the macroscopic character of the measured quantities. Furthermore, situations may arise where Bell's inequalities are violated, which proves that the measurement induced phase between the two condensates is of a non-classical nature.

Complement A_{XXI}

Density operator and correlations; separability

1 Von Neumann statistical entropy 2217
1-a General definition 2217
1-b Physical system composed of two subsystems 2219
2 Differences between classical and quantum correlations . . 2221
2-a Two levels of correlations 2221
2-b Quantum monogamy 2221
3 Separability 2223
3-a Separable density operator 2223
3-b Two spins in a singlet state 2225

In Chapter XXI, we mainly considered global systems $A + B$ described by a state vector (pure state). This complement will examine what happens when these global systems are described by a density operator (statistical mixture); we shall study the correlations quantum mechanics predicts in that case between the two subsystems A and B. We start by introducing, in § 1, the concept of statistical entropy, which yields a useful measure of their degree of correlation. We then analyze, in § 2, the differences between classical correlations (introduced at the probability level) and the quantum correlations (which can arise from the coherent superposition of state vectors). Finally, in § 3, we will come back to the important concept of separability already introduced in § F-2 of Chapter XXI.

1. Von Neumann statistical entropy

The statistical entropy introduced by von Neumann permits, in a straightforward way, to distinguish between a pure state and a statistical mixture; in the latter case, it also yields a measurement of the statistical character of the information known about the physical system. It is also a useful tool for studying in a quantitative way the amount of correlation between two physical systems.

1-a. General definition

With any density operator ρ, we associate a statistical entropy S by the relation:

$$S = -k_B \mathrm{Tr}\left\{\rho \ln \rho\right\} \tag{1}$$

where k_B is the Boltzmann constant. As ρ is Hermitian, this operator can be diagonalized. Noting n_k its eigenvalues, we get:

$$S = -k_B \sum_k n_k \ln n_k \tag{2}$$

Since all the n_k are included between 0 and 1, we necessarily have:

$$S \geq 0 \tag{3}$$

where the equality occurs only if ρ has one eigenvalue equal to 1, all the others being equal to zero. The entropy associated with ρ is therefore equal to zero only if this operator is a projector, and hence corresponds to a pure state. On the other hand, whenever ρ describes a statistical mixture, S is different from zero. It takes on its maximum value when the density operator ρ has equal populations in all the system's accessible states, i.e. if it is proportional to the identity operator in the state space. To prove this, let us vary each n_k by an amount $\mathrm{d}n_k$, and impose a zero variation for the sum over k of (2), while maintaining constant the sum of all n_k using a Lagrangian multiplier λ. We then get:

$$\mathrm{d}S - \lambda k_B \sum_k \mathrm{d}n_k = -k_B \sum_k \left[1 + \mathrm{Log}\ n_k + \lambda\right] \mathrm{d}n_k = 0 \tag{4}$$

For this expression to be zero for any $\mathrm{d}n_k$ means that all the $\ln\ n_k$, and hence all the n_k themselves, must be equal.

One can associate a concept of information, or rather a lack of information, with the entropy S. When the physical system is in a pure state, that state provides the maximum information on the system, compatible with quantum mechanics. In this situation, there is no lack of information and $S = 0$. On the other hand, when the system is spread over several pure states with comparable probabilities, a large value of S means that a lot of information about the system is lacking.

Comment:

The statistical entropy S characterizes the populations of the density matrix n_k (§ 4-c of Complement $\mathrm{E_{III}}$), but not the corresponding eigenvectors. Moreover, the same density operator can in general be obtained from several different statistical mixtures of pure states (*cf.* comment at the end of §4-a of Complement $\mathrm{E_{III}}$); the value of the entropy does not distinguish between these different mixtures.

A statistical mixture of several density operators can only increase the entropy of the system. Imagine, for example, that the density operator ρ is actually the combination of several density operators ρ_n with probabilities p_n (all positive, and whose sum over n is equal to 1), written as:

$$\rho = \sum_{\mathbf{n}} p_n\ \rho_n \tag{5}$$

Noting S_n the entropies associated with ρ_n:

$$S_n = -k_B \mathrm{Tr}\left\{\rho_n\ \ln \rho_n\right\} \tag{6}$$

we can write[1]:

$$S \geq \sum_n p_n\ S_n \tag{7}$$

[1]This properties is often called "entropy concavity".

Demonstration:

In § 1-b of Complement G_{XV}, we showed that, when ρ and ρ' are two density operators with traces equal to 1, one always has the following inequality:

$$\mathrm{Tr}\,\{\rho\,\ln\rho\} \geq \mathrm{Tr}\left\{\rho\,\ln\rho'\right\} \tag{8}$$

(the equality occurring only if $\rho' = \rho$). We can then write:

$$\begin{aligned} -\frac{S}{k_B} = \mathrm{Tr}\,\{\rho\,\ln\rho\} &= \sum_n p_n\ \mathrm{Tr}\,\{\rho_n\,\ln\rho\} \\ &\leq \sum_n p_n\ \mathrm{Tr}\,\{\rho_n\,\ln\rho_n\} = -\frac{1}{k_B}\sum_n p_n\ S_n \end{aligned} \tag{9}$$

which establishes relation (7).

1-b. Physical system composed of two subsystems

We now compare the entropies S_{A+B}, S_A and S_B associated, respectively, with the total density operator ρ_{A+B} and with the partial density operators ρ_A and ρ_B. We are going to show that $S_A + S_B = S_{A+B}$ when the systems A and B are not entangled, and that $S_A + S_B \geq S_{A+B}$ otherwise.

α. *Pure state*

Imagine first that the total system is in a pure entangled state. We have seen that, in that case, the two subsystems A and B are not described by pure but by statistical mixtures of states, so that:

$$\begin{aligned} S_A &= -k_B \mathrm{Tr}\,\{\rho_A\,\ln\rho_A\} \geq 0 \\ S_B &= -k_B \mathrm{Tr}\,\{\rho_B\,\ln\rho_B\} \geq 0 \end{aligned} \tag{10}$$

As the entropy S_{A+B} associated with a pure state is zero, it follows that:

$$S_A + S_B \geq S_{A+B} \tag{11}$$

(the equality corresponds to the special case where the pure state $|\Psi\rangle$ is a product, without entanglement, and where the Schmidt rank is equal to 1; see Chapter refch21, § C-3).

We can also use the Schmidt decomposition for $|\Psi\rangle$, which yields relations (C-7) and (C-8) of Chapter XXI, to get:

$$S_A = -k_B \sum_i q_i\,\ln\,q_i = S_B \tag{12}$$

Both entropies of the two subsystems are thus always equal whenever the total system is in a pure state.

β. Statistical mixture

When the total system is described by a density operator ρ_{A+B} not necessarily corresponding to a pure state, its entropy S_{A+B} may not be equal to zero. We are going to show, however, that this entropy S_{A+B} always remains lower or equal to the sum of the entropies of each subsystem, meaning that relation (11) remains valid in this more general case; this property is referred to as the "entropy subadditivity". The equality in (11) is obtained solely in the case where ρ_{A+B} is a product:

$$\rho_{A+B} = \rho_A \otimes \rho_B \tag{13}$$

which corresponds to the case of two subsystems, separately described by statistical mixtures, while remaining uncorrelated. The difference $S_A + S_B - S_{A+B}$ yields an estimate of the loss of precision between the quantum description of the total system, and the separate quantum descriptions of the two subsystems.

Demonstration:

According to inequality (8), we can write:

$$\mathrm{Tr}\left\{\rho \ln\rho\right\} \geq \mathrm{Tr}\left\{\rho \ln\left(\rho_A \otimes \rho_B\right)\right\} \tag{14}$$

We note $|u_i\rangle$ the eigenvectors of ρ_A with eigenvalues q_i, and $|v_l\rangle$ the eigenvectors of ρ_B with eigenvalues s_l. Let us now compute the trace of the right-hand side of (14) in the basis $\{|u_i, v_l\rangle\}$ of the tensor products of the eigenvectors of the two operators, with respective eigenvalues q_i and s_l; we get:

$$\begin{aligned}
\mathrm{Tr}\Big\{\rho\ \mathrm{Log}\left(\rho_A \otimes \rho_B\right)\Big\} &= \sum_{i,l} \langle u_i, v_l|\, \rho\ \mathrm{Log}\left(\rho_A \otimes \rho_B\right) |u_i, v_l\rangle \\
&= \sum_{i,l} \langle u_i, v_l|\, \rho\ \mathrm{Log}\left(q_i \times s_l\right) |u_i, v_l\rangle \\
&= \sum_{i,l} \langle u_i, v_l|\, \rho\, |u_i, v_l\rangle \mathrm{Log}\left(q_i\right) + \sum_{i,l} \langle u_i, v_l|\, \rho\, |u_i, v_l\rangle \mathrm{Log}\left(s_l\right)
\end{aligned} \tag{15}$$

Let us now choose $\rho = \rho_{A+B}$. The first term on the right-hand side can be written as:

$$\begin{aligned}
\sum_{i,l} \langle u_i, v_l|\, \rho_{A+B}\ \ |u_i, v_l\rangle \ln\left(q_i\right) &= \sum_i \langle u_i|\, \rho_A\ \ |u_i\rangle \ln\left(q_i\right) \\
&= \sum_i \langle u_i|\, \rho_A\ \ln\rho_A\, |u_i\rangle \\
&= \mathrm{Tr}\left\{\rho_A\ \ln\rho_A\right\}
\end{aligned} \tag{16}$$

The second term on the right-hand side of (15) yields a similar expression, where B replaces A. Finally, inequality (14) can be written as:

$$\mathrm{Tr}\left\{\rho_{A+B}\ \ln\rho_{A+B}\right\} \geq \mathrm{Tr}\left\{\rho_A\ \ln\rho_A\right\} + \mathrm{Tr}\left\{\rho_B\ \ln\rho_B\right\} \tag{17}$$

and leads to (11). The equality occurs if and only if (14) becomes an equality, i.e. if ρ_{A+B} is equal to the product (13).

2. Differences between classical and quantum correlations

Quantum mechanics offers more possibilities than classical physics for describing correlations between physical systems. We now briefly discuss such examples.

2-a. Two levels of correlations

The concept of correlation is not, intrinsically, a quantum notion, and it is well known in classical physics. It is then based on probabilistic calculations, which results in the linear weighting of a certain number of possibilities. In this classical context, one introduces a distribution yielding the probability for the first system to occupy a certain state, and the second system, another state; the two systems are correlated when this distribution is not a simple product. If, on the other hand, the distribution turns out to be a product, the two systems are not correlated; a measurement on one of the systems does not change the information about the other. In particular, this is what happens if the states of the two systems, and consequently that of the total system, are perfectly well defined, a case where the notion of correlation between the two systems becomes irrelevant. This means that the notion of correlation between two classical systems is closely linked to an imperfect definition of the state of the total system.

In quantum mechanics, things are totally different. To begin with, even if a physical system is perfectly well defined by a state vector, many of its physical properties are not so precisely defined: during several realizations of the experiment, their measurement can provide fluctuating results. These results can nevertheless be correlated: as an example, we saw in § B of Chapter XXI that the components of each of the two spin-1/2 particles are completely indeterminate but perfectly correlated. Such correlations appear directly at the level of the state vector itself, which can be written as a linear superposition of states where the spins have various orientations. The correlations are therefore related to the the quantum mechanical superposition principle; this is totally different from the combinations of probabilities, which are quadratic functions of that state vector. Letting correlations appear directly at the probability amplitude level, one has access to a level that is, in a way, "a step ahead" of the linear weighting of classical probabilities, and maintains the possibility of quantum interference effects. Note, however, that the existence of this level of combinations does not exclude classical probabilities from coming into play. One can also assume, in quantum mechanics, that the state of the total system is only known in a probabilistic way, so that the two probability levels may coexist. To sum up, it is clear that the concept of quantum correlations covers many more possibilities than correlations in classical physics[2].

2-b. Quantum monogamy

Another purely quantum property is that, if a physical system A is strongly entangled with a physical system B, it cannot be strongly entangled with another system C. Such a property does not have any equivalent in classical physics, where, obviously, nothing prevents correlating a third system C with two others A and B, all the while keeping

[2]We shall introduce, in § 3, a criterion (negativity of the coefficients of the total density operator expansion into a sum of products) for confirming the quantum nature of the correlations between two subsystems.

their initial correlation. This quantum property is often referred to as "entanglement monogamy".

Let us assume, for example, that two spins are in a state of the same type as the singlet state (B-1) of Chapter XXI:

$$|\Psi_{AB}\rangle = \frac{1}{\sqrt{2}}\left[|A:+;B:-\rangle + e^{i\xi}\,|A:-;B:+\rangle\right] \tag{18}$$

(the singlet state is obtained for $\xi = \pi$). How can we add an additional spin without destroying the correlation between the first two? One could imagine the three spin state to be written as:

$$|\Psi_{ABC}\rangle = |\Psi_{AB}\rangle \otimes |C:\theta\rangle = \frac{1}{\sqrt{2}}\left[|A:+;B:-;C:\theta\rangle + e^{i\xi}\,|A:-;B:+;C:\theta\rangle\right] \tag{19}$$

where $|\theta\rangle$ is any normalized state for the third spin. This ket obviously conserves the same entanglement between spins A and B as in state (18), but the third spin is then totally uncorrelated with the first two.

Another possibility is to choose as a state vector:

$$|\Psi_{ABC}\rangle = \frac{1}{\sqrt{2}}\left[|A:+;B:-;C:\theta_1\rangle + e^{i\xi}\,|A:-;B:+;C:\theta_2\rangle\right] \tag{20}$$

The density operator ρ_{AB} describing spins A and B is obtained by taking the partial trace (Complement E_{III}, § 5-b):

$$\rho_{AB} = \mathrm{Tr}_C\left\{|\Psi_{ABC}\rangle\langle\Psi_{ABC}|\right\} \tag{21}$$

Computing the matrix elements of this partial trace shows that:

$$\begin{aligned}\rho_{AB} = \frac{1}{2}\Big[& \langle\theta_1|\theta_1\rangle \;\; |A:+;B:-\rangle\langle A:+;B:-| \\ & + \;\; \langle\theta_2|\theta_2\rangle \;\; |A:-;B:+\rangle\langle A:-;B:+| \\ & + e^{i\xi}\langle\theta_1|\theta_2\rangle \;\; |A:-;B:+\rangle\langle A:+;B:-| \\ & + \;\; e^{-i\xi}\langle\theta_2|\theta_1\rangle \;\; |A:+;B:-\rangle\langle A:-;B:+|\Big]\end{aligned} \tag{22}$$

One can then distinguish several cases:

– If $|\theta_1\rangle = |\theta_2\rangle$, we find again (19), and the third spin is not entangled with the first two. The density operator ρ_{AB} is then written:

$$\begin{aligned}\rho_{AB} = {} & |A:+;B:-\rangle\langle A:+;B:-| \;\; + \;\; |A:-;B:+\rangle\langle A:-;B:+| \\ & + e^{i\xi}\,|A:-;B:+\rangle\langle A:+;B:-| \;\; + \;\; e^{-i\xi}\,|A:+;B:-\rangle\langle A:-;B:+|\end{aligned} \tag{23}$$

which is simply the projector onto state (18); it conserves all the entanglement of spins A and B.

– The opposite case is when $|\theta_1\rangle$ and $|\theta_2\rangle$ are orthogonal, so that $|\Psi_{ABC}\rangle$ becomes a so called GHZ state (Greenberger, Horne and Zeilinger; *cf.* Complement B_{XXI}:

$$|\Psi_{GHZ}\rangle = \frac{1}{\sqrt{2}}\left[|A:+;B:-;C:+\rangle + e^{i\xi}\,|A:-;B:+;C:-\rangle\right] \tag{24}$$

where, when one goes from the first component to the second, the three spins switch from one state to the orthogonal one. The second line in (22) then cancels out and the partial trace ρ_{AB} becomes:

$$\rho_{AB} = \frac{1}{2}\left[\,|A:+;B:-\rangle\,\langle A:+;B:-| + |A:-;B:+\rangle\,\langle A:-;B:+|\,\right] \tag{25}$$

which is a statistical mixture of two possibilities, with probabilities 1/2: the two spins are either in the state $|A:+;B:-\rangle$, or in the state $|A:-;B:+\rangle$. The quantum coherences between these two states (terms dependent on the phase ξ) have totally disappeared. The correlation between the two spins A and B is then of a classical nature[3], and no entanglement comes into play.

– In the intermediate situation where $|\theta_1\rangle$ and $|\theta_2\rangle$ are neither parallel nor orthogonal, we see from (22) that a certain coherence remains (non-diagonal elements). The more parallel $|\theta_1\rangle$ and $|\theta_2\rangle$ are, the more the partial density operator resembles that of the two initial spins which remain entangled, whereas the third one becomes less and less entangled with the first two; conversely, the more orthogonal they are, the more the initial spins lose their correlation, which becomes entirely transmitted to the three spin level.

This is actually a general property: when two physical systems are maximally entangled, a principle of mutual exclusion makes it impossible to entangle them with a third system. Mathematically, this property is expressed by the Coffman-Kundu-Wooters inequality [80].

3. Separability

From Bohr's point of view (§ F-2 of Chapter XXI), one must give up the notion of separability. Even when two physical subsystems are well separated in space, it does not infer that they have, each of them separately, their own physical properties (as EPR assumed); only the total system, including the measuring apparatus, can have such properties. On the other hand, we saw in § 2 that in quantum mechanics, there are two ways for introducing correlations between two systems: a classical way (by assuming they have given probabilities to be found in such or such individual correlated states), and a quantum way (by assuming entanglement directly at the level of a common state vector). We also know that, even though there are situations where quantum mechanics predicts violations of Bell' inequalities and hence of local realism, there are many others where its predictions obey these inequalities; a violation is, in a way, the signature of an ultra-quantum situation. It is thus interesting to look for a criterion allowing a distinction between these two types of correlations.

3-a. Separable density operator

Consider a total system described by a density operator ρ_{A+B} and composed of two subsystems A and B. Let us assume that ρ_{A+B} can be expanded as a series of density operators ρ_A^n and ρ_B^n pertaining to each of the two subsystems, with real and

[3]The density operator is separable in a sense that will be defined in § 3, and therefore cannot lead to violations of Bell's inequalities.

positive π_n coefficients whose sum equals to one, and can be assimilated to probabilities:

$$\rho_{A+B} = \sum_n \pi_n \ \rho_A^n \otimes \rho_B^n \tag{26}$$

with:

$$0 \leq \pi_n \leq 1 \qquad \text{and} \qquad \sum_n \pi_n = 1 \tag{27}$$

Intuitively, one can guess that the correlations contained in ρ_{A+B} must then be of a classical nature. The total system is, with a probability π_n, described by a density operator that is a product, without correlations, of density operators each describing one of the subsystems. The correlations between these subsystems are therefore introduced in a classical way, even if nothing prevents each subsystem from exhibiting strongly quantum individual properties.

Any density operator that can be decomposed as in (26) with positive π_n coefficients is, by definition, said to be "separable" [81, 82]. On the other hand, if any decomposition of ρ_{A+B} such as (26) necessarily includes π_n coefficients that are not real and positive, the density operator ρ_{A+B} contains quantum entanglement and is said to be non-separable. When the total system $A + B$ is separable, correlation measurements between physical properties of the two subsystems A and B can never lead to violations of Bell's inequalities. These violations are thus a sure sign of the non-separability of the density operator.

Demonstration

To show this, let us assume we perform two simultaneous measurements on the systems A and B, the first one depending on the measurement parameter a, and the second, on the measurement parameter b. We note $P_a(R_A)$ the projector acting in the state space of A and corresponding to the measurement result R_A (this projector is the sum of the projectors onto the eigenvectors associated with that measurement). In a similar way, we note $P_b(R_B)$ the projector in the state space of B corresponding to the measurement result R_B. When the total system is described by the density operator (26), the joint probability of obtaining both results R_A and R_B is:

$$\mathcal{P}(R_A, R_B) = \sum_n \pi_n \ \mathcal{P}_a^n(R_A) \times \mathcal{P}_b^n(R_B) \tag{28}$$

with:

$$\begin{aligned} \mathcal{P}_a^n(R_A) &= \mathrm{Tr}_A\{\rho_A^n P_a(R_A)\} \\ \mathcal{P}_b^n(R_B) &= \mathrm{Tr}_B\{\rho_B^n P_b(R_B)\} \end{aligned} \tag{29}$$

As all the numbers appearing on the right-hand side of (28) are positive, this equality has a natural interpretation in classical physics, which is the framework of our present argument. We are dealing with two levels of probabilities. At one level, the total system is prepared, with probability π_n, in a state where the two subsystems are uncorrelated. At a second level, for each value of n, the individual states of the subsystems are only known in a statistical way via the probability $\mathcal{P}_a^n(R_A)$ of a result R_A, and the probability $\mathcal{P}_b^n(R_B)$ of a result R_B.

We now show that if a relation of the type (28) is verified for all the measurement parameters a and b, with any positive probabilities $\mathcal{P}_a^n(R_A)$ and $\mathcal{P}_b^n(R_B)$, and with positive values for all the π_n, Bell's inequalities are always satisfied. Let's assume both results R_A and R_B can take on the values ± 1.

To start with, we assume that the physical properties of each pair of systems A and B depend on a classical random variable μ; this variable takes on a series of values μ_n, corresponding to each term in the summation over n appearing in (28), each with the probability π_n. This means that this summation over n can be interpreted as an average value over the random variable μ.

We then assume that the properties of the classical system A depend on a different random variable ν_A, which determines the result of the measurement performed on A. As an example, one can imagine that ν_A is regularly distributed on the segment $[0,1]$; result R_A is a function of ν_A, and takes the value $+1$ on a fraction of the segment of length $\mathcal{P}_a^n(R_A = +1)$, and the value -1 on the rest of the segment. This function models the probability written on the first line of (29); the measurement result is thus a function of ν_A, of the measurement parameter a, and of μ (which replaces n). Finally, we introduce the random variable ν_B, which determines the result of the measurement performed on B, with a distribution modeling in a similar way the probability written on the second line of (29), for any value of n and any choice of the measurement parameter b.

If we now regroup the three variables μ, ν_A and ν_B, as being the three components of a single variable λ, we reproduce the exact same hypotheses stated at the beginning of § F-3-a in Chapter XXI: the measurement results are functions, the first one of λ and a and the other one of λ and b. The same reasoning then leads to Bell's inequalities. Note that, at no point in this classical reasoning, did we have to consider the ensemble $A+B$ as a whole; it was thus to be expected that Bell's inequality would be established in this case.

3-b. Two spins in a singlet state

Let's go back to the example of two spin-1/2 particles in a singlet state:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\Big[\,|+,-\rangle - |-,+\rangle\Big] \tag{30}$$

In the basis of the 4 kets $|+,+\rangle$, $|+,-\rangle$, $|-,+\rangle$, $|-,-\rangle$ taken in that order, the matrix representing the density operator ρ_{A+B} is written:

$$(\rho_{A+B}) = |\Psi\rangle\langle\Psi| = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{31}$$

This matrix density (ρ_{A+B}) has non-diagonal elements between states $|+,-\rangle$ and $|-,+\rangle$, as, for example:

$$-1 = \langle +,-|\,\rho_{A+B}\,|-,+\rangle \tag{32}$$

To obtain such a non-diagonal term by a sum of products such as (26), will require:

$$-1 = \sum_n \pi_n\ \langle +|\,\rho_A^n\,|-\rangle\ \langle -|\,\rho_B^n\,|+\rangle \tag{33}$$

This demands introducing at least one term n that contains partial density operators ρ_A^n and ρ_B^n, both having non-diagonal elements. Now each of these two operators is a positive-definite operator. This means, for ρ_A^n for example, that it must have populations (diagonal matrix elements) $\langle +| \rho_A^n |+\rangle$ and $\langle -| \rho_A^n |-\rangle$ in the two individual spin states, and the same is true for ρ_B^n. The corresponding n term will necessarily introduce in (ρ_{A+B}) populations in the 4 states $|+,+\rangle$, $|+,-\rangle$, $|-,+\rangle$, $|-,-\rangle$; it will then be impossible to cancel those populations by adding other products of density operators (whose populations are positive) with positive π_n coefficients. Consequently, this density operator (ρ_{A+B}) is non-separable, and this is why it can lead to violations of Bell's inequalities.

Complement B_{XXI}

GHZ states, entanglement swapping

1 Sign contradiction in a GHZ state 2227
1-a Quantum calculation 2227
1-b Reasoning in the local realism framework 2230
1-c Discussion; contextuality 2231
2 Entanglement swapping 2232
2-a General scheme 2232
2-b Discussion 2235

Greenberger, Horne and Zeilinger (GHZ) showed in 1989 [83, 84] that violations of local realism even more spectacular than violations of Bell's inequalities could be observed on systems containing more than two correlated particles. These violations involve a contradiction in sign (and hence a violation of 100 %) for perfect correlations between measurement results, as opposed to inequalities violated by 40% for imperfect correlations. Observation of these violations requires creating an initial entanglement between three particles or more, as will be discussed in § 1. Another example, involving the entanglement of more than two particles, is the "entanglement swapping" method, explained in § 2. This method highlights a surprising property of entanglement: the possibility of entangling together two quantum systems, without them ever having to interact with each other.

1. Sign contradiction in a GHZ state

We consider a system composed of three spin-1/2 particles, as it is the simplest case for explaining how the GHZ contradiction can appear.

1-a. Quantum calculation

The three-spin system is described by the normalized quantum state:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\Big[\,|+,+,+\rangle + \eta\,|-,-,-\rangle\Big] \tag{1}$$

In this equality, the states $|\pm\rangle$ symbolize the eigenstates of the spin components along the Oz axis in an $Oxyz$ reference frame; to simplify the notation of the three particle ket, the spins are not numbered: the first sign corresponds to the state of the first spin, the second to that of the second, and similarly for the third spin. The number η stands for either $+1$, or -1. We now look for the quantum probabilities of measurement results of the components of each of the spins $\sigma_{1,2,3}$ of the three particles along two possible directions: either along the Ox direction, or along the perpendicular Oy direction (Fig. 1).

We start with the measurement of the product $\sigma_{1y}\,\sigma_{2y}\,\sigma_{3x}$. As we now show, $|\Psi\rangle$ is an eigenvector of this operator product, with eigenvalue $-\eta$, which means that the

measurement result is certain. The action of the first operator[1] is written:

$$\begin{aligned}\sigma_{3x}|\Psi\rangle &= \frac{1}{2}\left[\sigma_+(3)+\sigma_-(3)\right]|\Psi\rangle \\ &= \frac{1}{2\sqrt{2}}\left[\eta\sigma_+(3)|-,-,-\rangle+\sigma_-(3)|+,+,+\rangle\right] \\ &= \frac{1}{\sqrt{2}}\left[\eta|-,-,+\rangle+|+,+,-\rangle\right] \end{aligned} \tag{2}$$

The second operator then yields:

$$\begin{aligned}\sigma_{2y}\sigma_{3x}|\Psi\rangle &= \frac{1}{2i}\left[\sigma_+(2)-\sigma_-(2)\right]\sigma_{3x}|\Psi\rangle \\ &= \frac{1}{i\sqrt{2}}\left[\eta|-,+,+\rangle-|+,-,-\rangle\right] \end{aligned} \tag{3}$$

Finally, the product of the three operators yields:

$$\begin{aligned}\sigma_{1y}\sigma_{2y}\sigma_{3x}|\Psi\rangle &= \frac{1}{2i}\left[\sigma_+(1)-\sigma_-(1)\right]\sigma_{2y}\sigma_{3x}|\Psi\rangle \\ &= -\frac{1}{\sqrt{2}}\left[\eta|+,+,+\rangle+|-,-,-\rangle\right] \\ &= -\eta|\Psi\rangle \end{aligned} \tag{4}$$

The probability of observing the result $-\eta$ is:

$$\mathcal{P}(\sigma_{1y}\,\sigma_{2y}\,\sigma_{3x} \Longrightarrow -\eta) = 1 \tag{5}$$

whereas the probability $\mathcal{P}(\sigma_{1y}\,\sigma_{2y}\,\sigma_{3x} \Longrightarrow +\eta)$ of observing the other result is zero.

As the three spins play the same role, it is clear that $|\Psi\rangle$ is also an eigenvector of the two operator products $\sigma_{1x}\,\sigma_{2y}\,\sigma_{3y}$ and $\sigma_{1y}\,\sigma_{2x}\,\sigma_{3y}$, with eigenvalues $-\eta$. The corresponding probabilities are therefore:

$$\begin{aligned}&\mathcal{P}(\sigma_{1x}\,\sigma_{2y}\,\sigma_{3y} \Longrightarrow -\eta) = 1 \\ &\mathcal{P}(\sigma_{1y}\,\sigma_{2x}\,\sigma_{3y} \Longrightarrow -\eta) = 1\end{aligned} \tag{6}$$

It is thus certain that the three products take on the value $-\eta$.

We now consider the measurement result of the product of the three spin components along the same Ox axis. We use again (2), but (3) is now replaced by:

$$\begin{aligned}\sigma_{2x}\sigma_{3x}|\Psi\rangle &= \frac{1}{2}\left[\sigma_+(2)+\sigma_-(2)\right]\sigma_{3x}|\Psi\rangle \\ &= \frac{1}{\sqrt{2}}\left[\eta|-,+,+\rangle+|+,-,-\rangle\right] \end{aligned} \tag{7}$$

and (4) by:

$$\begin{aligned}\sigma_{1x}\sigma_{2x}\sigma_{3x}|\Psi\rangle &= \frac{1}{2}\left[\sigma_+(1)+\sigma_-(1)\right]\sigma_{2x}\sigma_{3x}|\Psi\rangle \\ &= \frac{1}{\sqrt{2}}\left[\eta|+,+,+\rangle+|-,-,-\rangle\right] \end{aligned} \tag{8}$$

[1] The Pauli matrices are defined in Complement A_{IV}; the operators $\sigma_\pm = \sigma_x \pm i\sigma_y$ obey the relation $\sigma_\pm|\mp\rangle = 2|\pm\rangle$.

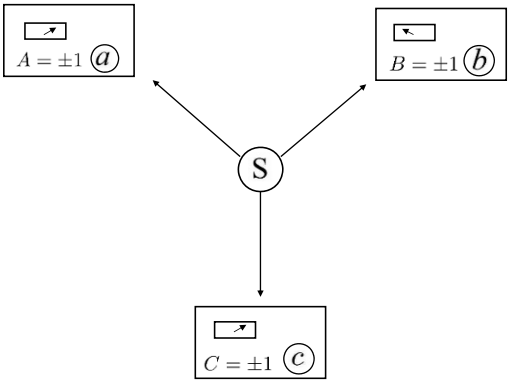


Figure 1: Schematic set-up for observing the GHZ contradictions. Three spins, initially in state $|\Psi\rangle$ written in (1), are measured in three different regions of space. In each of these regions a measuring apparatus is placed, with a setting enabling the local observer to choose between two possible spin component measurements, either along Ox, or along Oy. Whatever choices the three observers make, the results given by the three apparatus are $A = \pm 1$, $B = \pm 1$ and $C = \pm 1$.

This shows that $|\Psi\rangle$ is also an eigenvector of the operator product $\sigma_{1x}\,\sigma_{2x}\,\sigma_{3x}$, but this time with the eigenvalue $+\eta$. It follows that:

$$\mathcal{P}(\sigma_{1x}\,\sigma_{2x}\,\sigma_{3x} \Longrightarrow +\eta) = 1 \tag{9}$$

One can conclude, with certainty, that the measurement result of this product will be equal to $+\eta$.

Quantum mechanical measurement of a product of commuting operators:

Three operators such as σ_{1y}, σ_{2y} and σ_{3x}, acting on different spins, commute with each other; they form a CSCO (Complete Set of Commuting Observables) in the state space of the three spins. One can thus build a basis of eigenvectors $|\xi_1, \xi_2, \xi_3\rangle$ common to the three operators, labeled by the eigenvalue $\xi_1 = \pm 1$ of σ_{1y}, the eigenvalue $\xi_2 = \pm 1$ of σ_{2y} and the eigenvalue $\xi_3 = \pm 1$ of σ_{3x}. Any vector $|\Psi\rangle$ can be decomposed onto this basis as:

$$|\Psi\rangle = \sum_{\xi_1,\xi_2,\xi_3} c\,(\xi_1, \xi_2, \xi_3)\ \ |\xi_1, \xi_2, \xi_3\rangle \tag{10}$$

The action of the operator product $\sigma_{1y}\sigma_{2y}\sigma_{3x}$ on any ket is therefore to simply multiply each of its component $c\,(\xi_1, \xi_2, \xi_3)$ by the product $\xi_1\,\xi_2\,\xi_3$. Now the vector $|\Psi\rangle$ written in (1) is an eigenvector of that operator product, with the eigenvalue $-\eta$. The uniqueness of decomposition (10) then means that the only non-zero $c\,(\xi_1, \xi_2, \xi_3)$ coefficients are those

for which:

$$\xi_1\,\xi_2\,\xi_3 = -\eta \tag{11}$$

Suppose we measure, in a first experiment, the component σ_{1y} of the first spin. The result $\xi_1 = \pm 1$ is random. After the measurement, the projection postulate leads to a state that depends on this result, obtained by keeping in (1) only half of the components – those that correspond to the observed ξ_1 value. The components of the projected state vector still obey relation (11), where ξ_1 is now fixed. Similarly, if we continue the experiment and measure σ_{2y} for the second spin, the result $\xi_2 = \pm 1$ is also random, but the components of the new projected state vector still obey that same relation. As now ξ_1 and ξ_2 are both known, the same is true for ξ_3, whose value is determined by the first two measurements.

To sum up, the results observed for each spin component measurement fluctuate from one experiment to another, but these fluctuations are correlated and the product of the three results remain constant. One can obviously do the same analysis for the other sets of operators considered above, σ_{1x}, σ_{2y} and σ_{3y} for example.

1-b. Reasoning in the local realism framework

Let us leave, for a moment, standard quantum mechanics and examine what a local realistic theory (in the EPR sense of these words) would predict in such a situation. As we are in a particularly simple case where the initial quantum state is an eigenvector of all the observables coming into play (all the results are certain), one could expect nothing particular to happen. On the contrary, we now show that a complete contradiction appears between local realism and the predictions of quantum mechanics.

The local realism argument we present is a direct generalization of that used to obtain Bell's inequalities in § F-3-a of Chapter XXI. We first notice that the perfect correlations imply that the measurement result of a spin component along Ox (or Oy) of any particle can be deduced from the results of measurements performed on other particles, at arbitrarily large distances. The EPR argument then requires the existence of elements of reality corresponding to these two component directions, that we shall note $A_{x,y} = \pm 1$ for the first spin, $B_{x,y} = \pm 1$ for the second, and finally $C_{x,y} = \pm 1$ for the third. According to the EPR argument, for each experiment (i.e. for each emission of a group of three particles), these six numbers have well determined values, even though they are a priori unknown. These numbers are simply the results that shall be obtained, should measurements be performed later on. As an example, a measurement on the first spin will necessarily yield A_x if the chosen analysis direction is along Ox, or A_y if it is along Oy, independently of the type of measurements performed on the other two spins.

To have an agreement with the three equalities (5) and (6) imposes that:

$$\begin{aligned} A_y B_y C_x &= -\eta \\ A_x B_y C_y &= -\eta \\ A_y B_x C_y &= -\eta \end{aligned} \tag{12}$$

Now, in the logic of local realism, the same values of A, B and C can also be used for an experiment where the three spin components are measured along the same Ox direction: the result should simply be the product $A_x B_x C_x$. As the squares of the numbers A_y^2,

etc., are always equal to +1, we can obtain that product by multiplying the lines of (12), which yields:

$$A_x B_x C_x = -\eta \tag{13}$$

That is where the contradiction shows up: equality (9) predicts that the measurement of $\sigma_{1x}\,\sigma_{2x}\,\sigma_{3x}$ must always yield the result $+\eta$, which has the opposite sign! There could not be a greater contradiction between local realism and quantum mechanical predictions.

1-c. Discussion; contextuality

Compared with the violations of Bell's inequality, the GHZ contradiction seems far more spectacular, since a 100 % contradiction is obtained with 100 % certainty. From an experimental point of view, however, the necessity to bring into play three remote and entangled particles is a complex challenge.

To easily identify the three spins (deciding which measurement pertains to spin noted A, to spin noted B, and to spin noted C), and to be sure the three measurements are performed far from each other, let us assume the spins each occupy a different region of space. When the spatial variables are taken into account, the ket (1) can be rewritten more explicitly in the form:

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\,|1:u_a\rangle\,|2:u_b\rangle\,|3:u_c\rangle \otimes \Big[\,|1:+;2:+;3:+\rangle + \eta\;\;|1:-;2:-;3:-\rangle\Big] \tag{14}$$

where $|u_{a,b,c}\rangle$ are three orbital states whose wave functions do not overlap. They can for example be entirely localized in separate boxes where the measurements are performed. One then assumes that none of the particles will be left unmeasured and that each of them is separately observed. The experimental procedure is to first choose, for each box, a component Ox or Oy, then perform the three corresponding measurements in each box, obtain the three results $A_{x,y}$, $B_{x,y}$ and $C_{x,y}$, and finally compute their product.

A first necessary verification is to perform a large number of experiments and measure successively the three products$A_y B_y C_x$, $A_x B_y C_y$ and $A_y B_x C_y$, to be sure that the perfect correlations predicted by quantum mechanics are indeed observed (it is an essential step for the EPR argument, which infers from it the existence of 6 separate elements of reality). One then measures the product $A_x B_x C_x$ and, if quantum mechanics is right, one will observe a sign opposite to the EPR prediction. This means that the value obtained in a measurement of σ_{1x} (for example) depends on the Ox or Oy components measured on the other spins; this remains true even if the corresponding operators commute with σ_{1x}. This leads us to the general concept of "quantum contextuality": in an experiment where several commuting observables are measured, one must take into account, according to Bohr's prescription, the ensemble of the experimental set-up (the whole context of the system to be measured); it would not be correct to reason as if these measurements were independent processes.

Experimental tests of GHZ equalities have been performed [85, 86]. These experiments require three particles to be placed in the quantum state (14), which is not an easy task. Nevertheless, using elaborate quantum optics techniques, the correctness of quantum mechanical predictions has been verified in such a case, with experiments involving 3 or 4 entangled photons, as well as with NMR (Nuclear Magnetic Resonance) techniques.

2. Entanglement swapping

We now describe the "entanglement swapping" method, which enables entangling particles coming from independent sources (i.e. having no common past) through a quantum measurement process.

2-a. General scheme

Consider two sources S_{12} and S_{34} each creating a pair of entangled photons (Fig. 2). The first one creates a photon with momentum $\mathbf{k}_1$ and another one with momentum $\mathbf{k}_2$, whose polarizations are entangled in states H (horizontal polarization, in the plane of the figure) and V (vertical polarization, perpendicular to the plane of the figure). In a similar way, the second source creates a photon with momentum $\mathbf{k}_3$ and another one with momentum $\mathbf{k}_4$, whose polarizations are entangled in the same way. The initial state describing the two pairs is the tensor product of two states, each describing two particles:

$$|\Psi\rangle = \frac{1}{2}\left[|\mathbf{k}_1, H; \mathbf{k}_2, V\rangle + |\mathbf{k}_1, V; \mathbf{k}_2, H\rangle\right] \otimes \left[|\mathbf{k}_3, H; \mathbf{k}_4, V\rangle + |\mathbf{k}_3, V; \mathbf{k}_4, H\rangle\right] \tag{15}$$

While the two photons emitted by a given source are strongly entangled, no entanglement exists between the two pairs of photons, emitted by each of the two sources. It is useful to introduce the four different states pertaining to the wave vectors $\mathbf{k}_i$, $\mathbf{k}_j$:

$$\begin{aligned}
\left|\Phi^B_{i,j}\right\rangle_{(\eta)} &= \frac{1}{\sqrt{2}}\left[|\mathbf{k}_i, H; \mathbf{k}_j, H\rangle + \eta\,|\mathbf{k}_i, V; \mathbf{k}_j, V\rangle\right] \\
\left|\Theta^B_{i,j}\right\rangle_{(\eta)} &= \frac{1}{\sqrt{2}}\left[|\mathbf{k}_i, H; \mathbf{k}_j, V\rangle + \eta\,|\mathbf{k}_i, V; \mathbf{k}_j, H\rangle\right]
\end{aligned} \tag{16}$$

with, here again, $\eta = \pm 1$. These states (often called "Bell states" in the literature, hence the superscript B) form an orthonormal basis of the state space associated with particles i and j. One can show that (the computation is straightforward but a bit tedious and will not be detailed here):

$$\left|\Phi^B_{1,4}\right\rangle_{(+1)} \otimes \left|\Phi^B_{2,3}\right\rangle_{(+1)} - \left|\Phi^B_{1,4}\right\rangle_{(-1)} \otimes \left|\Phi^B_{2,3}\right\rangle_{(-1)} = |H; V; V; H\rangle + |V; H; H; V\rangle \tag{17}$$

and that:

$$\left|\Theta^B_{1,4}\right\rangle_{(+1)} \otimes \left|\Theta^B_{2,3}\right\rangle_{(+1)} - \left|\Theta^B_{1,4}\right\rangle_{(-1)} \otimes \left|\Theta^B_{2,3}\right\rangle_{(-1)} = |H; V; H; V\rangle + |V; H; V; H\rangle \tag{18}$$

(to simplify the notation, it is implicitly assumed, on the right-hand side of both equations, that the order of the particle's momenta is always $\mathbf{k}_1$, $\mathbf{k}_2$, $\mathbf{k}_3$ and $\mathbf{k}_4$). We can then write state (15) in the form:

$$\begin{aligned}
|\Psi\rangle = \frac{1}{2}\Big[&\left|\Phi^B_{1,4}\right\rangle_{(+1)} \otimes \left|\Phi^B_{2,3}\right\rangle_{(+1)} - \left|\Phi^B_{1,4}\right\rangle_{(-1)} \otimes \left|\Phi^B_{2,3}\right\rangle_{(-1)} + \\
&+ \left|\Theta^B_{1,4}\right\rangle_{(+1)} \otimes \left|\Theta^B_{2,3}\right\rangle_{(+1)} - \left|\Theta^B_{1,4}\right\rangle_{(-1)} \otimes \left|\Theta^B_{2,3}\right\rangle_{(-1)}\Big]
\end{aligned} \tag{19}$$

Figure 2 schematizes the experiment to be performed. After they are emitted, the particles with momenta $\mathbf{k}_2$ and $\mathbf{k}_3$ undergo a measurement in which they interfere. This is achieved by sending these two particles to a beam splitter BS, followed by two

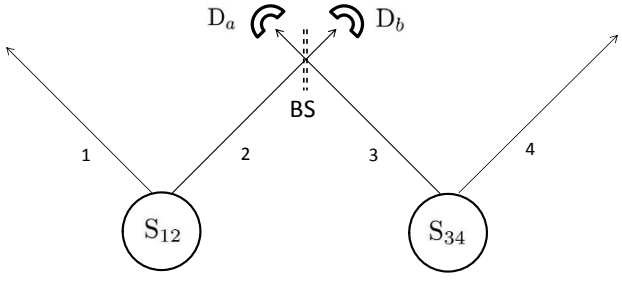


Figure 2: Schematic diagram of the "entanglement swapping" method. Two sources S_{12} and S_{34} each emit a pair of entangled particles, with wave vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ for the first one, $\mathbf{k}_3$ and $\mathbf{k}_4$ for the second. These sources are independent. A beam splitter BS is inserted in the path of particles $\mathbf{k}_2$ and $\mathbf{k}_3$; it is followed by two detectors D_a and D_b that measure the particle number in each of the exit channels a and b. This measurement has the effect of projecting the state vector, hence bringing the two particles $\mathbf{k}_1$ and $\mathbf{k}_4$ into a totally entangled state, even though these particles have never interacted.

detectors D$_a$ and D$_b$ measuring which exit channel were followed by the particles. If the two particles exit through two different channels, the corresponding eigenvector for this measurement result is the state $\left|\Theta^B_{2,3}\right\rangle_{(-1)}$; this is because, as we show below, the three other states $\left|\Theta^B_{2,3}\right\rangle_{(+1)}$, $\left|\Phi^B_{2,3}\right\rangle_{(+1)}$ and $\left|\Phi^B_{2,3}\right\rangle_{(-1)}$ correspond to situations where the two particles exit through the same channel. The measurement thus projects state (19) onto the last of its four components. The net result is that if the two particles with momenta $\mathbf{k}_2$ and $\mathbf{k}_3$ exit through different channels (which happens one out of four times), the two particles with momenta $\mathbf{k}_1$ and $\mathbf{k}_4$ reach the state $\left|\Theta^B_{1,4}\right\rangle_{(-1)}$. This means that the two non-observed particles reach a totally entangled state though they can be arbitrarily far from each other. It is worth noting that the initial entanglement concerns the two particles $\mathbf{k}_1$ and $\mathbf{k}_2$, and, separately, the two particles $\mathbf{k}_3$ and $\mathbf{k}_4$. Performing a suitable measurement on a particle of each pair, one projects the two remaining particles into a strongly entangled state, even though they never interacted at any stage of the process.

Demonstration:

Let us show that $\left|\Theta^B_{2,3}\right\rangle_{(-1)}$ is an initial state of two interfering particles that will lead to their exiting through different channels. We introduce for that purpose the two creation operators $a^\dagger_{\mathbf{k}_2,H}$ and $a^\dagger_{\mathbf{k}_2,V}$ in the individual state with wave vector $\mathbf{k}_2$ and polarization H or V, as well as the two operators $a^\dagger_{\mathbf{k}_3,H}$ and $a^\dagger_{\mathbf{k}_3,V}$ in the individual state with wave vector $\mathbf{k}_3$ and polarization H or V. The state $\left|\Theta^B_{2,3}\right\rangle_{(-1)}$ can be written:

$$\left|\Theta^B_{2,3}\right\rangle_{(-1)} = \frac{1}{\sqrt{2}}\left[a^\dagger_{\mathbf{k}_2,H}a^\dagger_{\mathbf{k}_3,V} - a^\dagger_{\mathbf{k}_2,V}a^\dagger_{\mathbf{k}_3,H}\right]|0\rangle \tag{20}$$

As the particles go through the beam splitter, their polarizations are not modified, but

their wave vectors are. In terms of creation operators, this leads to the unitary transformations:

$$\begin{aligned} a^\dagger_{\mathbf{k}_2,H} &\Rightarrow \frac{1}{\sqrt{2}}\left[a^\dagger_{\mathbf{k}_2,H} + ia^\dagger_{\mathbf{k}_3,H}\right] \\ a^\dagger_{\mathbf{k}_3,H} &\Rightarrow \frac{1}{\sqrt{2}}\left[ia^\dagger_{\mathbf{k}_2,H} + a^\dagger_{\mathbf{k}_3,H}\right] \end{aligned} \tag{21}$$

where the i factors come from the phase change in a light beam as it undergoes internal reflection. Similar equalities are obtained for the V polarization, so that:

$$\begin{aligned} &a^\dagger_{\mathbf{k}_2,H}\ a^\dagger_{\mathbf{k}_3,V} - a^\dagger_{\mathbf{k}_2,V}\ a^\dagger_{\mathbf{k}_3,H} \\ &\Rightarrow \frac{1}{2}\Big[\left(a^\dagger_{\mathbf{k}_2,H} + ia^\dagger_{\mathbf{k}_3,H}\right)\left(ia^\dagger_{\mathbf{k}_2,V} + a^\dagger_{\mathbf{k}_3,V}\right) \\ &\qquad\qquad - \left(a^\dagger_{\mathbf{k}_2,V} + ia^\dagger_{\mathbf{k}_3,V}\right)\left(ia^\dagger_{\mathbf{k}_2,H} + a^\dagger_{\mathbf{k}_3,H}\right)\Big] \end{aligned} \tag{22}$$

As creation operators in different modes commute with each other, this operator is equal to:

$$\left[a^\dagger_{\mathbf{k}_2,H}\ a^\dagger_{\mathbf{k}_3,V} - a^\dagger_{\mathbf{k}_2,V}\ a^\dagger_{\mathbf{k}_3,H}\right] \tag{23}$$

so that state $\left|\Theta^B_{2,3}\right\rangle_{(-1)}$ is transformed, after the beam splitter, into:

$$\left|\Theta^B_{2,3}\right\rangle_{(-1)} \Rightarrow \frac{1}{\sqrt{2}}\left[a^\dagger_{\mathbf{k}_2,H}\ a^\dagger_{\mathbf{k}_3,V} - a^\dagger_{\mathbf{k}_2,V}a^\dagger_{\mathbf{k}_3,H}\right]|0\rangle \tag{24}$$

This shows that if the state before crossing the beam splitter is $\left|\Theta^B_{2,3}\right\rangle_{(-1)}$, the two photons are still in two different exit channels after the crossing.

If now the state before crossing the beam splitter is $\left|\Theta^B_{2,3}\right\rangle_{(+1)}$, we must replace (22) by:

$$\begin{aligned} a^\dagger_{\mathbf{k}_2,H}a^\dagger_{\mathbf{k}_3,V} + a^\dagger_{\mathbf{k}_2,V}a^\dagger_{\mathbf{k}_3,H} &\Rightarrow \frac{1}{2}\left[\left(a^\dagger_{\mathbf{k}_2,H} + ia^\dagger_{\mathbf{k}_3,H}\right)\left(ia^\dagger_{\mathbf{k}_2,V} + a^\dagger_{\mathbf{k}_3,V}\right)\right. \\ &\qquad \left. + \left(a^\dagger_{\mathbf{k}_2,V} + ia^\dagger_{\mathbf{k}_3,V}\right)\left(ia^\dagger_{\mathbf{k}_2,H} + a^\dagger_{\mathbf{k}_3,H}\right)\right] \\ &= i\left[a^\dagger_{\mathbf{k}_2,H}\ a^\dagger_{\mathbf{k}_2,V} + a^\dagger_{\mathbf{k}_3,H}\ a^\dagger_{\mathbf{k}_3,V}\right] \end{aligned} \tag{25}$$

which means that the two photons always exit the beam splitter through the same channel. In the same way, for the state $\left|\Phi^B_{2,3}\right\rangle_{(\pm1)}$, we get the operator:

$$\begin{aligned} a^\dagger_{\mathbf{k}_2,H}a^\dagger_{\mathbf{k}_3,H} \pm a^\dagger_{\mathbf{k}_2,V}a^\dagger_{\mathbf{k}_3,V} &\Rightarrow \frac{1}{2}\left[\left(a^\dagger_{\mathbf{k}_2,H} + ia^\dagger_{\mathbf{k}_3,H}\right)\left(ia^\dagger_{\mathbf{k}_2,H} + a^\dagger_{\mathbf{k}_3,H}\right)\right. \\ &\qquad \left. \pm \left(a^\dagger_{\mathbf{k}_2,V} + ia^\dagger_{\mathbf{k}_3,V}\right)\left(ia^\dagger_{\mathbf{k}_2,V} + a^\dagger_{\mathbf{k}_3,V}\right)\right] \\ &= i\left[\left(a^\dagger_{\mathbf{k}_2,H}\right)^2 + \left(a^\dagger_{\mathbf{k}_3,H}\right)^2 \pm \left(a^\dagger_{\mathbf{k}_2,V}\right)^2 \pm \left(a^\dagger_{\mathbf{k}_3,V}\right)^2\right] \end{aligned} \tag{26}$$

It shows again that for each term the photons exit through the same channel. The state $\left|\Theta^B_{2,3}\right\rangle_{(-1)}$ is therefore the only one that will lead to the photons exiting through different channels.

2-b. Discussion

In classical physics, it is also possible to obtain correlations between two objects initially totally independent, by sorting objects with which each of them is correlated. To underline the fundamental difference with entanglement swapping, we now discuss such a classical experiment. Imagine that two independent sources emit pairs of correlated objects, numbered 1 and 2 for the first source, 3 and 4 for the second, as in Figure 2. Each time the experiment is performed, each source emits two classical objects sharing a common property (such as, for example, the same color, or opposite angular momenta, etc.). The two sources are nevertheless totally uncorrelated (the objects emitted by two different sources present no correlations between their colors, their angular momenta, etc.). If, however, one selects particular experiments where particles 2 and 3 present a certain correlation (for example identical colors, or else parallel or antiparallel angular momenta, etc.), it is clear that the particles 1 and 4 will also be correlated, even if they never interacted in the past and if they are very far apart. It is a mere consequence of the selection performed in a classical probability distribution, and could be called "classical correlation swapping".

Note, however, that this selection remains purely classical; no entanglement can be produced by this method. Should a Bell experiment be performed on the objects 1 and 4, the correlations obtained will necessarily obey Bell's inequalities since we are in a classical physics context. The entanglement swapping method, however, allows creating by selection a true entanglement leading to strong violations of Bell's inequalities. This method is a way of producing stronger correlations than classical correlation swapping, and has been demonstrated in several experiments [87, 88].

Conclusion

The two examples we discussed illustrate the variety of situations where quantum entanglement produces significant physical effects, even when the entangled quantum systems are arbitrarily far from each other. In each situation, it is essential to follow the basic rules of quantum mechanics, and perform the computations with a global state vector, including all the physical systems under study. Any attempt to perform separate computations in different regions of space, and then add correlations using classical probability calculations, will necessarily lead to predictions ignoring numerous non-local quantum effects, in contradiction with experimental results.

Complement C_{XXI}

Measurement induced relative phase between two condensates

1 Probabilities of single, double, etc. position measurements 2239
- 1-a Single measurement (one particle) . . . 2239
- 1-b Double measurement (two particles) . . . 2240
- 1-c Generalization: measurement of any number of positions . . . 2241

2 Measurement induced enhancement of entanglement 2242
- 2-a Measuring the single density $\mathcal{P}(x_1)$ . . . 2243
- 2-b Entanglement between the two modes after the first detection 2243
- 2-c Measuring the double density $\mathcal{P}(x_2, x_1)$ . . . 2244
- 2-d Discussion . . . 2244

3 Detection of a large number Q of particles 2245
- 3-a Probability of a multiple detection sequence . . . 2245
- 3-b Discussion; emergence of a relative phase . . . 2247

Introduction

Referring to Bose-Einstein condensation (Complement C_{XV}, a system of identical bosons, all occupying the same individual state $|u_i\rangle$ is called a "condensate". It is described by a Fock state such as the one given by (A-17) of Chapter XV, where all the occupation numbers are zero, except for n_i, whose value N can be very large:

$$|u_i : N\rangle = \frac{1}{\sqrt{N!}} \left[a^\dagger_{u_i}\right]^N |0\rangle \tag{1}$$

The operator $a^\dagger_{u_i}$ is the creation operator of a particle in the individual state $|u_i\rangle$, and $|0\rangle$ is the vacuum state (for which all the occupation numbers are zero). In a similar way, a "double condensate" is described by a Fock state where N_1 particles are in the individual state $|u_i\rangle$ and N_2 particles in the individual state $|u_j\rangle$; its normalized state is written as:

$$|\Phi_0\rangle = |u_i : N_1; u_j : N_2\rangle = \frac{1}{\sqrt{N_1!N_2!}} \left[a^\dagger_{u_i}\right]^{N_1} \left[a^\dagger_{u_j}\right]^{N_2} |0\rangle \tag{2}$$

We shall focus on the case where the individual states u_i and u_j are states with well defined but opposite wave vectors $\pm k$:

$$|\Phi_0\rangle = |+k : N_1; -k : N_2\rangle = \frac{1}{\sqrt{N_1!N_2!}} \left[a^\dagger_{+k}\right]^{N_1} \left[a^\dagger_{-k}\right]^{N_2} |0\rangle \tag{3}$$

In such a state, while the occupation numbers are perfectly well defined, the relative phase between the two condensates is completely undetermined; we will confirm this

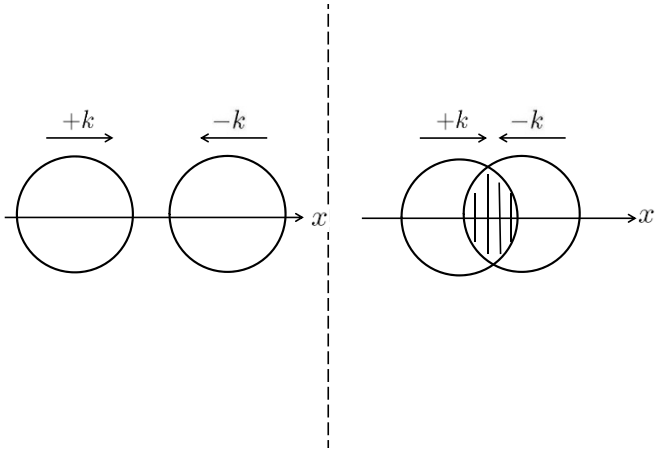


Figure 1: The left-hand side of the figure represents two groups of particles prepared independently. The first one is composed of a large number of particles, N_1, all in the same individual state with momentum $+k$ along the Ox axis, and propagating towards the right; the second group includes N_2 particles, in the other individual state with opposite $-k$ momentum, propagating towards the left. Each of these groups of independent particles is in a "condensate". The right-hand side of the figure shows that, after a certain time, the two condensates overlap in space; this allows measuring the positions of the particles in the overlap region. For clarity, the computations are limited to one dimension, taking into account only the x coordinate.
The first position measurement is totally random, but as measurements continue, there appear a periodic bunching of the observed positions, progressively forming a sharper fringe pattern. These fringes result from the emergence of a relative phase between the two condensates, which can only be a consequence of the position measurements, as it was totally absent at the beginning of the experiment.
If the whole process is repeated from the beginning, fringes appear with a position generally different from the first experiment: the phase appearing in each new experiment is totally independent of the one observed in previous experiments.

later (§ 2-a) by showing that measuring the position of a single particle with such a state does not lead to any observable interference fringes.

Now, recent experiments [89] have shown that when the positions of many particles are measured, interference fringes can indeed be observed in the region where the two condensates overlap (Figure 1). This remains true even if the condensates have been created in a totally independent way. This fringe pattern corresponds to a well defined value of the relative phase of the two condensates; one may then wonder about the origin of this observed phase. The object of this complement is to study the mechanism responsible for the emergence of this relative phase. We will show that it results from the successive detections of particles, which progressively modifies the initial state: as more and more particles are detected, it produces a progressively increasing entanglement between the two condensates, defining their relative phase in a more and more precise way.

During the course of one experiment, the position of the fringes is determined. However, should one repeat the experiment, preparing the condensates in exactly the same way, a new relative phase will progressively appear during the successive particle detections; its value is, in general, completely different from the one previously obtained. This means that if one averages observations over a large number of independent successive experiments, the fringes will be blurred and eventually completely disappear. The emergence of the phase is clearly observable only in the course of one specific experiment.

We first compute, in § 1, the probability of measurements concerning the positions of one, two, and more particles; we will show that these probabilities are proportional to spatio-temporal correlation functions of field operators of the various particles. Starting, in § 2, from two condensates in an initial state described by a simple juxtaposition of two Fock states, we will see how the successive particles' position measurements create an increasing entanglement between the two condensates. A more general study of the system's evolution is presented in § 3, showing, in particular, how this growing entanglement leads to a better and better definition of the relative phase between the two condensates. The computations presented in this complement are limited to the case where the number of measured positions remains small compared to the total number of particles in the condensates. Complement D_{XXI}, will go a step further and relax this hypothesis.

1. Probabilities of single, double, etc. position measurements

As we start the successive measurements of the particles' positions, we begin by computing the probability of finding a first particle in an interval[1] of infinitesimal width Δ around position $x = x_1$, then a second particle in the interval of width Δ around position $x = x_2$, etc. The computations we present here are valid for any state $|\Phi_0\rangle$ of the identical particle system. They are, actually, the equivalent of those encountered in the general study of correlation functions in § B-3 of Chapter XVI; nevertheless, we will go through them again in the specific context of the present complement. The results will be applied, in § 2, to the particular case where $|\Phi_0\rangle$ is a double Fock state.

1-a. Single measurement (one particle)

With a measurement of the position yielding a result included in the interval $D_1 = \left[x_1 - \frac{\Delta}{2}, x_1 + \frac{\Delta}{2}\right]$, we can associate the Hermitian operator:

$$P_\Delta(x_1) = \int_{x_1-\frac{\Delta}{2}}^{x_1+\frac{\Delta}{2}} \mathrm{d}x\ \Psi^\dagger(x)\,\Psi(x) \tag{4}$$

where $\Psi(x)$ is the field operator destroying a particle at point x, and $\Psi^\dagger(x)$ its Hermitian conjugate, creating a particle. The average value of $P_\Delta(x_1)$ yields the average particle number in the interval D_1. In what follows, we shall, most of the time, assume that Δ is small enough compared to the other dimensions of the problem to justify the approximation:

$$P_\Delta(x_1) \simeq \Delta \times \Psi^\dagger(x_1)\,\Psi(x_1) \tag{5}$$

[1] To keep the notation simple, we consider a one-dimensional problem and note x_1, x_2,..,x_p, the particles' positions. Generalizing to three dimensions only requires replacing all the x_i by the vectors $\mathbf{r}_i$.

Operator $P_\Delta(x_1)$ is a symmetric one-particle operator of the type described in relation (B-1) in Chapter XV. It can also be written as:

$$P_\Delta(x_1) = \sum_{q=1}^{N} \int_{x_1-\frac{\Delta}{2}}^{x_1+\frac{\Delta}{2}} \mathrm{d}x \ |q:x\rangle\langle q:x| \tag{6}$$

which is the sum over all the $q = 1, 2, .., N$ particles of the projectors into the interval D_1 of the positions of each of them. As all these projectors commute with each other, and since they each have eigenvalues 1 and 0, the eigenvalues of $P_\Delta(x_1)$ are equal to 0, 1, 2, .. N. Now, if Δ is small enough, there can be no more than one particle in the interval D_1; this means that the only accessible eigenvalues are 0 and 1, so that $P_\Delta(x_1)$ becomes the projector associated with the measurement of a particle's presence in the interval D_1:

$$[P_\Delta(x_1)]^2 = P_\Delta(x_1) \qquad \text{if } \Delta \to 0 \tag{7}$$

Suppose now the system is in state $|\Phi_0\rangle$. The probability $\mathcal{P}_{D_1}$ of finding a particle in the infinitesimal interval D_1 of length Δ is:

$$\begin{aligned}\mathcal{P}_{D_1}(x_1) &= \langle\Phi_0| P_\Delta(x_1) |\Phi_0\rangle \\ &= \Delta \ \langle\Phi_0| \Psi^\dagger(x_1)\Psi(x_1)|\Phi_0\rangle\end{aligned} \tag{8}$$

Right after the detection of this first particle, the system is now, according to relation (E-39) of Chapter III (postulate of wave packet reduction), in the normalized state:

$$|\Phi_0'\rangle = \frac{1}{\sqrt{\mathcal{P}_{D_1}(x_1)}} P_\Delta(x_1)|\Phi_0\rangle \tag{9}$$

1-b. Double measurement (two particles)

Let us now focus on the probability $\mathcal{P}_{D_1D_2}(x_2, x_1)$ of detecting a first particle in an interval of width Δ around point x_1, then a second one in an interval Δ around point x_2; we assume the system does not have time to evolve in between the two measurements.

We start by computing the conditional probability[2] $\mathcal{P}_{D_1D_2}(x_2/x_1)$ of detecting a particle in the interval $\left[x_2-\frac{\Delta}{2}, x_2+\frac{\Delta}{2}\right]$ noted D_2, knowing that a particle has been detected in the interval $\left[x_1-\frac{\Delta}{2}, x_1+\frac{\Delta}{2}\right]$ already noted D_1. This probability equals:

$$\begin{aligned}\mathcal{P}_{D_1D_2}(x_2/x_1) &= \langle\Phi_0'| P_\Delta(x_2)|\Phi_0'\rangle \\ &= \frac{1}{\mathcal{P}_{D_1}(x_1)}\langle\Phi_0| P_\Delta(x_1)P_\Delta(x_2)P_\Delta(x_1)|\Phi_0\rangle\end{aligned} \tag{10}$$

where, in the second line, we have used (9); the projector $P_\Delta(x_2)$ may be obtained by replacing x_1 by x_2 in expression (4). We assume that the two detection intervals do

[2]Note the different notation used for the conditional probability $\mathcal{P}_{D_1D_2}(x_2/x_1)$, with a fraction bar between the variables, and the simple probability (a priori probability) $\mathcal{P}_{D_1D_2}(x_2, x_1)$ of obtaining the two results. These two probabilities are related by expression (14).

not overlap in space, so that all the operators appearing in $P_\Delta(x_1)$ commute with those appearing in $P_\Delta(x_2)$. We then have, taking (7) into account:

$$\begin{aligned} P_\Delta(x_1) P_\Delta(x_2) P_\Delta(x_1) &= [P_\Delta(x_1)]^2 P_\Delta(x_2) \\ &= P_\Delta(x_1) P_\Delta(x_2) \end{aligned} \tag{11}$$

If Δ is small enough, we can replace $P_\Delta(x_1)$ and $P_\Delta(x_2)$ by their expressions (5); this leads to:

$$\begin{aligned} P_\Delta(x_1) P_\Delta(x_2) &= \Delta^2 \ \Psi^\dagger(x_1) \Psi(x_1) \Psi^\dagger(x_2) \Psi(x_2) \\ &= \Delta^2 \ \Psi^\dagger(x_1) \Psi^\dagger(x_2) \Psi(x_2) \Psi(x_1) \end{aligned} \tag{12}$$

where, in the second line, we again used the fact that field operators defined in non-overlapping regions of space commute with each other. Inserting this result in (10), we get:

$$\mathcal{P}_{D_1 D_2}(x_2/x_1) = \frac{\Delta^2}{\mathcal{P}_{D_1}(x_1)} \langle\Phi_0| \Psi^\dagger(x_1) \Psi^\dagger(x_2) \Psi(x_2) \Psi(x_1) |\Phi_0\rangle \tag{13}$$

Now, the probability of detecting a particle at point x_1, then a particle at point x_2, is the product of the probability $\mathcal{P}_{D_1}(x_1)$ of detecting a particle at point x_1 and the conditional probability $\mathcal{P}_{D_1 D_2}(x_2/x_1)$ of detecting a particle at point x_2 knowing that a particle has been detected at point x_1:

$$\mathcal{P}_{D_1 D_2}(x_2, x_1) = \mathcal{P}_{D_1}(x_1) \times \mathcal{P}_{D_1 D_2}(x_2/x_1) \tag{14}$$

Taking (13) into account, this leads to:

$$\begin{aligned} \mathcal{P}_{D_1 D_2}(x_2, x_1) &= \Delta^2 \langle\Phi_0| \Psi^\dagger(x_1) \Psi^\dagger(x_2) \Psi(x_2) \Psi(x_1) |\Phi_0\rangle \\ &= \langle\Phi_{21} \ |\Phi_{21}\rangle \end{aligned} \tag{15}$$

where $|\Phi_{21}\rangle$ is the non-normalized state:

$$|\Phi_{21}\rangle = \Delta \times \Psi(x_2) \Psi(x_1) |\Phi_0\rangle \tag{16}$$

The probability we are looking for is simply the squared norm of the ket obtained by destroying in the initial state a particle at point x_1, and a second one at point x_2, multiplied by the width Δ of the infinitesimal measurement interval.

1-c. Generalization: measurement of any number of positions

The previous computations deal with simple and double density measurements; we now generalize them to measurements of higher order densities. From now on and to simplify the notation, we shall omit the D subscript in the probabilities $\mathcal{P}$.

To compute the probability associated with a triple measurement, we start from the expression of the state vector right after the detection of the second particle at x_2. Taking (10) into account, and similarly as for (9), this normalized state is written:

$$|\Phi_0''\rangle = \frac{1}{\sqrt{\langle\Phi_0'| P_\Delta(x_2) |\Phi_0'\rangle}} P_\Delta(x_2) |\Phi_0'\rangle = \frac{1}{\sqrt{\mathcal{P}(x_2/x_1)}} P_\Delta(x_2) |\Phi_0'\rangle \tag{17}$$

or else, if we insert (9) and use (14):

$$|\Phi_0''\rangle = \frac{1}{\sqrt{\mathcal{P}(x_2/x_1)\mathcal{P}(x_1)}}\, P_\Delta(x_2)\, P_\Delta(x_1)\,|\Phi_0\rangle$$
$$= \frac{1}{\sqrt{\mathcal{P}(x_2,x_1)}}\, P_\Delta(x_2)\, P_\Delta(x_1)\,|\Phi_0\rangle \tag{18}$$

The probability of the third measurement at x_3, knowing that the first two measurements gave results at x_1 and x_2, is thus:

$$\mathcal{P}(x_3/x_2,x_1) = \langle\Phi_0''|\, P_\Delta(x_3)\,|\Phi_0''\rangle$$
$$= \frac{1}{\mathcal{P}(x_2,x_1)}\,\langle\Phi_0|\, P_\Delta(x_1)\, P_\Delta(x_2)\, P_\Delta(x_3)\, P_\Delta(x_2)\, P_\Delta(x_1)\,|\Phi_0\rangle \tag{19}$$

As before, we consider that the position measurement zones do not overlap, so that all the projection operators commute with each other:

$$\mathcal{P}(x_3/x_2,x_1) = \frac{1}{\mathcal{P}(x_2,x_1)}\,\langle\Phi_0|\, P_\Delta(x_3)\, P_\Delta(x_2)\, P_\Delta(x_1)\,|\Phi_0\rangle$$
$$= \frac{\Delta^3}{\mathcal{P}(x_2,x_1)}\,\langle\Phi_0|\,\Psi^\dagger(x_1)\,\Psi^\dagger(x_2)\,\Psi^\dagger(x_3)\,\Psi(x_3)\,\Psi(x_2)\,\Psi(x_1)\,|\Phi_0\rangle \tag{20}$$

In the second line, we assumed Δ was small enough to use the approximate relation (5).

As the law of conditional probabilities indicates that the probability of the three measurements at x_1, x_2 and x_3 is given by:

$$\mathcal{P}(x_3,x_2,x_1) = \mathcal{P}(x_2,x_1)\times\mathcal{P}(x_3/x_2,x_1) \tag{21}$$

we simply get:

$$\mathcal{P}(x_3,x_2,x_1) = \Delta^3\,\langle\Phi_0|\,\Psi^\dagger(x_1)\,\Psi^\dagger(x_2)\,\Psi^\dagger(x_3)\,\Psi(x_3)\,\Psi(x_2)\,\Psi(x_1)\,|\Phi_0\rangle \tag{22}$$

which is a direct generalization of (15).

The same line of reasoning allows showing that the probability associated with the measurement of Q positions is proportional to the average value in the system's state of a product of $2Q$ field operators $\Psi^\dagger$ and Ψ arranged in normal order, and evaluated at x_1, x_2,...x_Q. The probabilities are therefore equal to the spatio-temporal correlation functions of the field operators arranged in normal order (and multiplied by Δ^Q).

2. Measurement induced enhancement of entanglement

We have reasoned until now in a general way, without specifying the initial state $|\Phi_0\rangle$ of the system under study. We now assume we are dealing with a double condensate, as in (3), and for simplicity we shall take $N_1 = N_2 = N$ (actually the computation that follows only requires the hypothesis $N_1 \simeq N_2$). We thus have N particles occupying the individual state $|k\rangle$ with a well-defined momentum $\hbar k$, and an equal number of particles occupying the state $|-k\rangle$ with opposite momentum:

$$|\Phi_0\rangle = |+k:N;-k:N\rangle = \frac{1}{N!}\left[a^\dagger_{+k}\right]^N\left[a^\dagger_{-k}\right]^N|0\rangle \tag{23}$$

We propose studying the interference signals that may occur in the single and double counting rates measured on such a state. We shall need the probabilities calculated above as well as expressions (A-3) and (A-6) of Chapter XVI for the field operators:

$$\Psi(x) = \frac{1}{L^{1/2}}\left[e^{ikx}a_k + e^{-ikx}a_{-k} + ..\right]$$
$$\Psi^\dagger(x) = \frac{1}{L^{1/2}}\left[e^{-ikx}a_k^\dagger + e^{ikx}a_{-k}^\dagger + ..\right] \tag{24}$$

where a_k (or a_{-k}) and $a_k^\dagger$ (or $a_{-k}^\dagger$) are the annihilation and creation operators of a particle in mode k (or $-k$), and where L is the edge of the box used to normalize the plane waves. The dots on the right-hand side of these formulas stand for the other terms present in the field operator expansions of these operators. Because of the choice of the initial state (23), these additional terms do not play any role in the following calculations, as will be shown below.

2-a. Measuring the single density $\mathcal{P}(x_1)$

Relation (8) now becomes:

$$\mathcal{P}(x_1) = \Delta\ \langle +k:N;-k:N|\,\Psi^\dagger(x_1)\Psi(x_1)\,|+k:N;-k:N\rangle \tag{25}$$

Using expressions (24) for the field operators, and the fact that the cross terms $a_{-k}^\dagger a_k$ and $a_k^\dagger a_{-k}$ have a zero average value in the double Fock state (23), we get:

$$\mathcal{P}(x_1) = \frac{\Delta}{L}\langle +k:N;-k:N|\,a_k^\dagger a_k + a_{-k}^\dagger a_{-k}\,|+k:N;-k:N\rangle = \frac{2N\Delta}{L} \tag{26}$$

This means that there is no interference in the single density measurement signal. This was to be expected since the initial double Fock state includes no phase that could help determine the eventual position of such fringes.

2-b. Entanglement between the two modes after the first detection

Relation (9) yields the ket $|\Phi_0'\rangle$, right after the first measurement. It can be written as:

$$|\Phi_0'\rangle = \frac{1}{\sqrt{\mathcal{P}(x_1)}}\int_{x_1-\frac{\Delta}{2}}^{x_1+\frac{\Delta}{2}} \mathrm{d}x\ \Psi^\dagger(x)\,\Psi(x)\ \ |\Phi_0\rangle$$
$$= \frac{1}{L\sqrt{\mathcal{P}(x_1)}}\int_{x_1-\frac{\Delta}{2}}^{x_1+\frac{\Delta}{2}} \mathrm{d}x\ \left[e^{-ikx}a_k^\dagger + e^{ikx}a_{-k}^\dagger + ..\right]\left[e^{ikx}a_k + e^{-ikx}a_{-k} + ..\right]|\Phi_0\rangle \tag{27}$$

Taking (23) into account, and for an infinitesimal Δ, we get:

$$|\Phi_0'\rangle \propto \Big\{2N\,|\Phi_0\rangle + \sqrt{N(N+1)}\left[e^{-2ikx_1}\,|k:N+1;-k:N-1\rangle\right.$$
$$\left.+e^{2ikx_1}\,|k:N-1;-k:N+1\rangle\right] + ..\Big\} \tag{28}$$

where +.. stands for the components of $|\Phi_0'\rangle$ where a particle occupies an individual state other than k and $-k$; these components do not play any role in what follows. Relation (28) shows that the entanglement of state $|\Phi_0'\rangle$ has increased as a result of the detection of the first particle. This state now contains a linear superposition of the initial state and two additional states of the global system, $|N-1:+k;N:-k\rangle$ and $|N:+k;N-1:-k\rangle$; the coefficients of this superposition, and in particular their relative phase, depend on the point x_1 where the first particle has been detected.

2-c. Measuring the double density $\mathcal{P}(x_2,x_1)$

We now compute the probability $\mathcal{P}(x_2,x_1)$ associated with a double density measurement. Relations (15) and (16) show that the probability is the squared norm of the ket:

$$|\Phi_{21}\rangle = \Delta \times \Psi(x_2)\,\Psi(x_1)\,|N:+k;N:-k\rangle \tag{29}$$

Inserting in this equality the first relation (24), the terms symbolized by the dots .. disappear (as they involve annihilation operators yielding zero when acting on the $+k$ and $-k$ states, the only initially populated states). We obtain:

$$|\Phi_{21}\rangle = \frac{\Delta}{L} \times \left[e^{ikx_2}a_k + e^{-ikx_2}a_{-k}\right]\left[e^{ikx_1}a_k + e^{-ikx_1}a_{-k}\right]|N:+k;N:-k\rangle \tag{30}$$

or else:

$$\begin{aligned}|\Phi_{21}\rangle = \frac{\Delta}{L} \times \Big\{ \sqrt{N(N-1)}\Big[& e^{ik(x_1+x_2)}\,|N-2:+k;N:-k\rangle \\ & + e^{-ik(x_1+x_2)}\,|N:+k;N-2:-k\rangle\Big] \\ + N\left[e^{ik(x_2-x_1)} + e^{ik(x_1-x_2)}\right]&|N-1:+k;N-1:-k\rangle\Big\}\end{aligned} \tag{31}$$

The squared norm of this state vector yields the probability:

$$\mathcal{P}(x_2,x_1) = \frac{\Delta^2}{L^2}\left\{2N(N-1) + 4N^2\cos^2\left[k(x_2-x_1)\right]\right\} \tag{32}$$

The presence of the cosine of $k(x_2-x_1)$ reveals the existence of a spatial dependence, contrary to what happened for $\mathcal{P}(x_1)$: once a first particle is detected at x_1, the most probable positions x_2 for the second detection are those for which $k(x_2-x_1)$ is a multiple of π. In other words, fringes appear in the double density measurement.

2-d. Discussion

One may wonder which objects interfere in the double counting signal. They are not waves but transition amplitudes associated with two different paths leading the system from the initial state (23) to the same final state $|N-1,+k;N-1,-k\rangle$, where each of the two modes has lost one particle. In the above computation, the first path corresponds to the term $\left(e^{-ikx_2}a_{-k}\right)\left(e^{ikx_1}a_k\right)$, where one particle with momentum $+k$ disappears as it is detected at x_1 and the $-k$ particle disappears as it is detected at x_2; the second path corresponds to the term $\left(e^{ikx_2}a_k\right)\left(e^{-ikx_1}a_{-k}\right)$ where it is now the

particle with momentum $-k$ that disappears as it is detected at x_1 and the $+k$ particle that disappears as it is detected at x_2.

The double counting signal observed on a double condensate is very similar to the double photodetection signal obtained, in § 2-c-γ of Complement E_{XX}, in the study of a product of two one-photon wave packets. In both cases, the signal spatial dependence comes from a quantum interference between the amplitudes of two different paths between the same initial and final states. The difference between the paths comes from a different "switching" between one of the two components of the initial state and one of the two components of the final state.

3. Detection of a large number Q of particles

We now extend the previous reasoning to the case where any number Q of particles' position measurements are performed; we shall limit ourselves to the case where Q remains much smaller than the total particle number N of each condensate.

3-a. Probability of a multiple detection sequence

Generalizing relation (22) allows writing the probability $\mathcal{P}(x_Q, .., x_2, x_1)$ for detecting a particle at x_1, a particle at x_2, .. a particle at x_Q, in the form:

$$\mathcal{P}(x_Q, .., x_2, x_1) = \Delta^Q \ \langle N : +k; N : -k| \, \Psi^\dagger(x_1)\Psi^\dagger(x_2).. \, \Psi^\dagger(x_Q) \\ \Psi(x_Q).. \, \Psi(x_2)\Psi(x_1) \, |N : +k; N : -k\rangle \tag{33}$$

As before, we use relations (24) to replace the field operators and their adjoints by linear combinations of annihilation operators a_k, a_{-k} and creation operators $a_k^\dagger$, $a_{-k}^\dagger$. We then get:

$$\mathcal{P}(x_Q, ..., x_2, x_1) = \frac{\Delta^Q}{L^Q} \ \langle N : +k; N : -k| \, (a_k^\dagger e^{-ikx_1} + a_{-k}^\dagger e^{ikx_1})... \\ ...(a_k^\dagger e^{-ikx_Q} + a_{-k}^\dagger e^{ikx_Q})(a_k e^{ikx_Q} + a_{-k} e^{-ikx_Q})... \\ ... \, (a_k e^{ikx_1} + a_{-k} e^{-ikx_1}) \, |N : +k; N : -k\rangle \tag{34}$$

α. Simplifying hypothesis

When several annihilation operators act successively on the right on the initial ket, each of them introduces a varying factor $\sqrt{N-p}$; this factor depends on the number p of particles already annihilated by the other operators. In the same way, when the creation operators act on the left on the initial bra, they also introduce varying factors. To keep things simple, we shall ignore these variations, assuming that the total detection number Q is always very small compared to the total particle number N in each individual state:

$$Q \ll N \tag{35}$$

One can then replace all the factors $\sqrt{N-p}$ by $\sqrt{N}$:

$$\sqrt{N-p} \simeq \sqrt{N} \tag{36}$$

Apart from multiplication by a fixed factor $\sqrt{N}$, the only effect of each operator is to vary by one unit the occupation number; this result does not depend on the previous actions undergone by the state vector (all the operators commute, once the above approximation has been made). One can then freely move the annihilation and creation operators in the product of operators appearing in (34). Regrouping all the operators associated with the same values of x_i, we get the operator:

$$\begin{aligned} D(x_j) &= (a_k^\dagger e^{-ikx_j} + a_{-k}^\dagger e^{ikx_j})(a_k e^{ikx_j} + a_{-k} e^{-ikx_j}) \\ &= \left[a_k^\dagger a_k + a_{-k}^\dagger a_{-k} + a_k^\dagger a_{-k} e^{-2ikx_j} + a_{-k}^\dagger a_k e^{2ikx_j} \right] \end{aligned} \tag{37}$$

and expression (34) becomes:

$$\mathcal{P}(x_Q, ...x_2, x_1) = \frac{\Delta^Q}{L^Q} \ \langle N:+k; N:-k| \prod_{j=1}^{Q} D(x_j) \, |N:+k; N:-k\rangle \tag{38}$$

We are then left with the computation of the average value in the initial state of a product of operators $D(x_j)$. Expanding each of them according to the second line of (37), we get the sum of 4^Q products, most of them having, nevertheless, a zero average value in the double Fock state $|\Phi_0\rangle$. This is because the only products having a non-zero average value are those for which the repeated effect of the annihilation operators a_k is exactly balanced by the effect of an equal number of operators $a_k^\dagger$ (the particle number in the individual $-k$ state is then also constant, since the total number of particles must be conserved).

Consider then one of those non-zero products. Still using approximation (36), the contribution of each operator $D(x_j)$ will be one of the three factors $NF_{q_j}(x_j)$, with $q_j = 0, \pm 1$ and:

$$\begin{aligned} F_0(x_j) &= 2 \\ F_{\pm 1}(x_j) &= e^{\pm 2ikx_j} \end{aligned} \tag{39}$$

The contribution of F_0 leaves the particle numbers unchanged, the contribution of F_{+1} replaces a $+k$ particle by a $-k$ particle, and finally that of F_{-1} performs the opposite substitution. Relation (38) then becomes:

$$P(x_Q, ..., x_2, x_1) = \frac{\Delta^Q}{L^Q} \ \prod_{j=1}^{Q} \Big[\sum_{q_j = 0, \pm 1} NF_{q_j}(x_j) \Big] \qquad \text{with } S = 0 \tag{40}$$

where, when we expand the product in the right hand side of the equation, we retain only the terms for which the sum of all q_j vanishes:

$$S = \sum_{j=1}^{Q} q_j = 0 \tag{41}$$

This constraint simply expresses the conservation of the particle number in each individual state.

β. Simple expression for the multiple detection probability

An easy way to impose constraint (41) on all the q_j is to introduce a Kronecker delta $\delta_{S,0}$, and write:

$$P(x_Q, ..x_2, x_1) = \frac{\Delta^Q}{L^Q} \sum_{q_1, q_2, .., q_Q} \delta_{S,0} \prod_{j=1}^{Q} N F_{q_j}(x_j) \tag{42}$$

where the summation over the q_j is now free of constraint. We can then use relation:

$$\int_0^{2\pi} \frac{\mathrm{d}\xi}{2\pi}\, e^{iS\xi} = \delta_{S,0} \tag{43}$$

which amounts to multiplying each term $F_{q_j}(x_j)$ in (40) by $\exp(iq_j\xi)$ and summing over $\mathrm{d}\xi$. We then get:

$$P(x_Q, ..x_2, x_1) = \frac{\Delta^Q}{L^Q} \int_0^{2\pi} \frac{\mathrm{d}\xi}{2\pi} \prod_{j=1}^{Q} \sum_{q_j} N e^{iq_j\xi} F_{q_j}(x_j) \tag{44}$$

Each summation over q_j then yields the quantity:

$$\begin{aligned} &N\left[F_0(x_j) + F_+(x_j)e^{i\xi} + F_-(x_j)e^{-i\xi}\right] \\ &\qquad = N\left[2 + \exp(2ikx_j + i\xi) + \exp(-2ikx_j - i\xi)\right] \end{aligned} \tag{45}$$

which can be written as:

$$2N\left[1 + \cos(2kx_j + \xi)\right] = 4N\cos^2\left(kx_j + \frac{\xi}{2}\right) \tag{46}$$

Finally, we obtain the following simple analytical expression for the multiple detection rate:

$$P(x_Q, ..x_2, x_1) = \left(\frac{4\Delta N}{L}\right)^Q \int_0^{2\pi} \frac{\mathrm{d}\xi}{2\pi} \prod_{j=1}^{Q} \cos^2\left(kx_j + \frac{\xi}{2}\right) \tag{47}$$

3-b. Discussion; emergence of a relative phase

Relation (47) allows understanding how the successive measurements enable the progressive emergence of a relative phase.

α. Detecting the first particles

Let us start with the very first detection at $x = x_1$. Equation (47) yields the probability of such an event:

$$P(x_1) \propto \int_0^{2\pi} \frac{\mathrm{d}\xi}{2\pi} \cos^2(kx_1 + \frac{\xi}{2}) \tag{48}$$

The term in cosine appearing in the integral yields the fringes that one expects from the interference between two waves, with wave vectors $+k$ and $-k$ along the x axis, and a

phase shift ξ. However, the summation over ξ indicates that the interference pattern must be averaged over all the possible values of ξ, uniformly distributed between 0 and 2π: this means that the fringes are completely blurred out.

The double detection rate at x_1 and x_2 is obtained by keeping the terms $j = 1$ and $j = 2$ in (47). It is equal to:

$$P(x_2, x_1) \propto \int_0^{2\pi} \frac{\mathrm{d}\xi}{2\pi} \cos^2(kx_2 + \frac{\xi}{2}) \cos^2(kx_1 + \frac{\xi}{2}) \tag{49}$$

As the first detection has already occurred, x_1 is fixed in this equation, and the product of the two cosines yields the probability of finding the second particle at x_2. But the integral over $\mathrm{d}\xi$, which yields the x_2 dependence of the probability, is no longer over a phase ξ uniformly distributed between 0 and 2π, because of the presence of the $\cos^2(kx_1 + \xi/2)$ associated with the first detection; the blurring of the fringes is not as radical as before. For this second detection, the function $\cos^2(kx_1 + \xi/2)$ actually plays the role of an x_1 dependent phase distribution; the two detections are no longer independent. This confirms the qualitative discussion of § 2.

This mechanism can be generalized to higher order measurements. As an example, the triple detection rate at x_1, x_2 and x_3 is equal to:

$$P(x_3, x_2, x_1) \propto \int_0^{2\pi} \frac{\mathrm{d}\xi}{2\pi} \cos^2(kx_3 + \frac{\xi}{2}) \cos^2(kx_2 + \frac{\xi}{2}) \cos^2(kx_1 + \frac{\xi}{2}) \tag{50}$$

Once the first two detections have been made at x_1 and x_2, the relative phase distribution that comes into play for the third detection is the product of two cosine functions $\cos^2(kx_2+\xi/2) \cos^2(kx_1+\xi/2)$ – and no longer a single one as was the case before. As the product of two cosine functions yields a sharper curve than a single cosine function, the relative phase is better defined for the third detection than for the second. The process continues in the same way with the following detections, and the phase is more and more precisely defined. This means that it is the first detections that determine the positions of the fringes appearing in the following detections, each of them contributing to a more and more precise definition of the relative phase distribution.

This argument is of course only valid for a given experiment. If one performs a new experiment with the same experimental conditions, the first detections will not, in general, happen at the same places as in the first experiment. Consequently, after a large number of detections, a fringe pattern will appear, shifted with respect to the pattern observed in the first experiment. Finally, if one adds up the positions measured in a large number of successive experiments, the fringes average to practically zero, and one gets a quasi-uniform position distribution.

β. Emergence of a well-defined relative phase after a large number of detections

After a large number of detections, Q, the relative phase distribution for the $(Q + 1)^{th}$ detection is given by the product of a large number Q of cosine functions, yielding a very narrow phase distribution, centered at a value ξ_M. One can then replace in (47) all the $[1 + \cos(2kx_j + \xi)]$ by $[1 + \cos(2kx_j + \xi_M)]$, so that the probability becomes a product: the detections are now independent, the interference pattern becomes stable with a sharper and sharper contrast. These predictions have been confirmed by numerical simulations based on equation (47).

Narrowing of the relative phase distribution

The narrowing of the relative phase distribution can be explained by an analytical calculation. Let us assume that after Q detections this distribution can be approximated by a Gaussian curve centered at ξ_M, and with a width σ_Q:

$$W_Q(\xi) \simeq e^{-(\xi-\xi_M)^2/\sigma_Q^2} \tag{51}$$

After the $(Q+1)^{th}$ detection, the new distribution will be given by:

$$W_{Q+1}(\xi) \simeq e^{-(\xi-\xi_M)^2/\sigma_Q^2}\,[1+\cos(2kx_{Q+1}+\xi)] \tag{52}$$

As the function $1+\cos(2kx_{Q+1}+\xi)$ is much broader than $W_{Q+1}(\xi)$, it can be expanded in powers of $\xi-\xi_M$ in the vicinity of $\xi=\xi_M$ where the distribution $W_Q(\xi)$ takes on significant values:

$$\begin{aligned}1+\cos(2kx_{Q+1}+\xi) &= 1+\cos(2kx_{Q+1}+\xi_M)\\ &-(\xi-\xi_M)\sin(2kx_{Q+1}+\xi_M)-\frac{1}{2}(\xi-\xi_M)^2\cos(2kx_{Q+1}+\xi_M)\end{aligned} \tag{53}$$

One can also expand $W_Q(\xi)$ in the vicinity of $\xi=\xi_M$:

$$e^{-(\xi-\xi_M)^2/\sigma_Q^2} = 1-\frac{(\xi-\xi_M)^2}{\sigma_Q^2} \tag{54}$$

We then multiply (53) and (54) and obtain an expansion for $W_{Q+1}(\xi)$:

$$\begin{aligned}W_{Q+1}(\xi) &\simeq 1+\cos(2kx_{Q+1}+\xi_M)-(\xi-\xi_M)\sin(2kx_{Q+1}+\xi_M)\\ &-(\xi-\xi_M)^2\left[\frac{1}{2}\cos(2kx_{Q+1}+\xi_M)+\frac{1}{\sigma_Q^2}\Big(1+\cos(2kx_{Q+1}+\xi_M)\Big)\right]\end{aligned} \tag{55}$$

We note that $W_{Q+1}(\xi)$ depends on the position x_{Q+1} of the $(Q+1)^{th}$ detection. We can obtain an average value for $W_{Q+1}(\xi)$ by weighting $W_{Q+1}(\xi)$ by the probability $[1+\cos(2kx_{Q+1}+\xi_M)]$ for the $(Q+1)^{th}$ detection to occur at $x=x_{Q+1}$, and integrate x_{Q+1} over a spatial period of the interference pattern:

$$\overline{W_{Q+1}(\xi)} = \int_0^{2\pi/2k} \mathrm{d}x_{Q+1}\,[1+\cos(2kx_{Q+1}+\xi_M)]\,W_{Q+1}(\xi). \tag{56}$$

Since:

$$\begin{aligned}\overline{\cos(2kx_{Q+1}+\xi_M)} &= \overline{\sin(2kx_{Q+1}+\xi_M)}\\ &= \overline{\cos(2kx_{Q+1}+\xi_M)\sin(2kx_{Q+1}+\xi_M)} = 0\end{aligned} \tag{57}$$

and:

$$\overline{\cos^2(2kx_{Q+1}+\xi_M)} = \frac{1}{2} \tag{58}$$

we finally obtain:

$$\overline{W_{Q+1}(\xi)} \simeq \frac{3}{2}\left[1-(\xi-\xi_M)^2\left(\frac{1}{\sigma_Q^2}+\frac{1}{6}\right)\right] \simeq \frac{3}{2}\,e^{-(\xi-\xi_M)^2)/\sigma_{Q+1}^2} \tag{59}$$

where:

$$\frac{1}{\sigma_{Q+1}^2} = \frac{1}{\sigma_Q^2} + \frac{1}{6}. \tag{60}$$

Equation (60) shows that $\sigma_{Q+1} < \sigma_Q$, meaning that the distribution curve becomes narrower after each detection. One can easily iterate equation (60) to obtain:

$$\frac{1}{\sigma_{Q+n}^2} = \frac{1}{\sigma_Q^2} + \frac{n}{6} \tag{61}$$

where n is a positive integer. This shows that if $n \gg 1/\sigma_Q^2$, the width of the relative phase distribution decreases as $1/\sqrt{n}$.

A similar computation can be made to study the position of $W_{Q+1}(\xi)$'s maximum when Q increases. One finds that the center of the relative phase distribution is shifted by a quantity proportional to $1/n$.

Finally, it is interesting to note the link between the uncertainty on the relative phase (which decreases as the detection number Q increases) and the uncertainty on the difference $N_+ - N_-$ between the numbers $N_\pm$ of particles in the $\pm k$ condensates (which, on the contrary, increases). At the beginning of the experiment (before the first detection), we have $N_+ = N_- = N$. After the first detection, we saw in § 2 that the state of the system contains a linear superposition of states $N_+ = N \pm 1$ and $N_- = N \mp 1$: the difference $N_+ - N_-$ between N_+ and N_- is no longer fixed and equal to zero, but can take on several values 0, ± 2. After the second detection, the state of the system contains a superposition of states having always the same value of $N_+ + N_-$, but values of $N_+ - N_-$ that can be equal to 0, ± 2, .. and so on. After Q detections, the values of $N_+ - N_-$ spread out between $-2Q$ and $+2Q$. This result is an illustration of the fact that the relative phase and the difference between the particle numbers between the two condensates are conjugate quantities.

Conclusion

This complement illustrates how successive measurements on a system having components on two individual states, each with N particles, can build up (from zero) a relative phase between these components; for this to happen, the measurements must depend on the relative phase between those two components. Mathematically, relation (47) shows that the results obtained for an ensemble of Q position measurements (with $Q \ll N$) are exactly the same as if an initial well defined phase ξ had existed from the beginning of the experiment, even though it was totally unknown (and could have taken on any value uniformly distributed between 0 and 2π). The measurements did indeed introduce entanglement and its associated relative phase, but the quantum predictions are equivalent to those obtained by assuming that the measurements only reveal a preexisting phase (as in quantum theories with so-called "additional variables").

The process we have discussed is, however, of an essentially different nature: it is indeed each individual measurement that contributes to a better and better definition of the relative phase for the measurements to come; these will occur at points x whose probability distributions depend on the results of all the previously performed measurements. We shall see in Complement D_{XXI} that if, instead of measuring a fraction of

all the particles, they each undergo a measurement, the phase properties can no longer be understood as that of a classical preexisting (but unknown) phase; these properties clearly become quantum, as shown by the possibility of violations of Bell's inequalities.

Complement D_{XXI}

Emergence of a relative phase with spin condensates; macroscopic non-locality and the EPR argument

1 Two condensates with spins 2254

1-a Spin 1/2: a brief review 2254

1-b Projectors associated with the measurements 2255

2 Probabilities of the different measurement results 2255

2-a A first expression for the probability 2256

2-b Introduction of the phase and the quantum angle 2258

3 Discussion 2259

3-a Measurement number $Q \ll 2N$ 2259

3-b Macroscopic EPR argument 2261

3-c Measuring all the spins, violation of Bell's inequalities 2263

This complement continues the discussion of Complement C_{XXI} on the measurement induced emergence of a relative phase between condensates, but in a more general case. We established in C_{XXI} that as more and more measurement results are obtained, their number still remaining smaller than the total particle number, the relative phase of the two condensates becomes better defined. It soon reaches a classical regime where it is (almost) perfectly determined. This necessarily comes with large fluctuations of the numbers of particles occupying the two individual states (or more precisely of their difference), as required by the uncertainty relation between phase and occupation numbers. In the present complement, a first important difference is that we no longer assume that the number of measurements remains small compared to the total particle number. This will enable us to follow the evolution of the phase properties during the whole series of measurements, including the last moments when the number of particles remaining to be measured is just a few units. For these few remaining particles, the fluctuations in the difference of the occupation numbers is necessarily limited to a few units, meaning that the phase can no longer be precisely determined. The phase then comes back to a quantum regime, where one can no longer interpret the measurement results in a classical context (preexisting but totally unknown phase). Another difference with Complement C_{XXI} is that we now assume the two condensates correspond to different individual spin states. Instead of position measurements yielding continuous results, we can now perform measurements on the spin directions, which yield discrete results. This will make it easier to discuss the quantum effects, which can lead to violations of Bell's inequalities (Chapter XXI, § F-3). Another advantage of dealing with spins is that we can go back to the EPR argument (Chapter XXI, § F-1) in a case where the elements of reality, introduced by EPR, are macroscopic and have, in addition, a simple physical interpretation (spin angular momentum).

1. Two condensates with spins

We now assume that the two individual states populated in the condensates are the two states $|u,\pm\rangle$ corresponding to two different internal states noted $|\pm\rangle$, but to the same orbital state $|u\rangle$:

$$|u,\pm\rangle = |u\rangle \otimes |\pm\rangle \tag{1}$$

If $(a_{u,+})^\dagger$ and $(a_{u,-})^\dagger$ are the creation operators associated with these states, the state $|\Phi_0\rangle$ of the system formed by the juxtaposition of the two condensates can be written as:

$$|\Phi_0\rangle = \frac{1}{N!}\left[(a_{u,+})^\dagger\right]^N \quad \left[(a_{u,-})^\dagger\right]^N \quad |0\rangle \tag{2}$$

which replaces relation (23) of Complement C_{XXI}; the total particle number is $2N$.

By commodity, we will often call "spin states" the two states $|\pm\rangle$, and reason as if they were indeed the two accessible states of a spin-1/2 particle. This is just a manner of speaking: according to the spin-statistic theorem (Chapter XIV, § C-1), bosons cannot be half-integer spin particles. The system we consider is actually an ensemble of bosons that have access to only two internal states; these can be, for example, the two $m = 0$ and $m = 1$ states of a spin equal to 1, or not necessarily spin states.

1-a. Spin 1/2: a brief review

For the reasoning that follows, it may be useful to recall a few relations (Chapter IV, § A-2) concerning a spin 1/2 (with no orbital variables). As pointed out above, we are dealing with a fictitious spin, whose operators act on any two internal states, noted $|\pm\rangle$ by pure commodity. Operator σ_z, associated with the first Pauli matrix (Complement A_{IV}), is the difference between the projector onto the state $|+\rangle$ and the projector onto the state $|-\rangle$:

$$\sigma_z = |+\rangle\langle+| \;-\; |-\rangle\langle-| \tag{3}$$

whereas operators σ_x and σ_y are expressed as a function of the two non-diagonal operators $|+\rangle\langle-|$ and $|-\rangle\langle+|$ as:

$$\begin{aligned}\sigma_x &= \;\; |+\rangle\langle-| \;+\; |-\rangle\langle+| \\ \sigma_y &= -i\,|+\rangle\langle-| \;+\; i\,|-\rangle\langle+|\end{aligned} \tag{4}$$

As for the fictitious spin component along a direction $O\varphi$ in the xOy plane, making an angle φ with the Ox axis, the corresponding operator is written:

$$\sigma_\varphi = \cos\varphi\;\sigma_x + \sin\varphi\;\sigma_y = e^{-i\varphi}\,|+\rangle\langle-| \;+\; e^{i\varphi}\,|-\rangle\langle+| \tag{5}$$

Its eigenvalues are $\eta = \pm 1$, and its eigenvectors can be expressed as:

$$\begin{aligned}|\psi_{\eta=+1}\rangle &= \frac{1}{\sqrt{2}}\left[e^{-i\varphi/2}\,|+\rangle + e^{i\varphi/2}\,|-\rangle\right] \\ |\psi_{\eta=-1}\rangle &= \frac{1}{\sqrt{2}}\left[-e^{-i\varphi/2}\,|+\rangle + e^{i\varphi/2}\,|-\rangle\right]\end{aligned} \tag{6}$$

as can be easily checked by applying operator (5) to these relations. The projector onto the ket with eigenvalue η can thus be written:

$$\begin{aligned} P_\eta(\varphi) &= \frac{1}{2}\left[1+\eta\sigma_\varphi\right] \\ &= \frac{1}{2}\left[1+\eta\left(e^{-i\varphi}\left|+\right\rangle\left\langle-\right| \;+\; e^{i\varphi}\left|-\right\rangle\left\langle+\right|\right)\right] \end{aligned} \tag{7}$$

1-b. Projectors associated with the measurements

For an ensemble of identical bosons having orbital variables, we note $\Psi_\pm(\mathbf{r})$ the two field operators associated with the two internal states $\pm$. The operator associated with the total particle density at point $\mathbf{r}$ is the sum of the local densities, $\Psi_+^\dagger(\mathbf{r})\Psi_+(\mathbf{r})$ corresponding to the + spin state, and $\Psi_-^\dagger(\mathbf{r})\Psi_-(\mathbf{r})$ corresponding to the − state:

$$n(\mathbf{r}) = \Psi_+^\dagger(\mathbf{r})\Psi_+(\mathbf{r}) + \Psi_-^\dagger(\mathbf{r})\Psi_-(\mathbf{r}) \tag{8}$$

As for the operator associated with the σ_z spin component along the Oz quantization axis, relation (3) indicates that it is the difference:

$$\sigma_z(\mathbf{r}) = \;\Psi_+^\dagger(\mathbf{r})\Psi_+(\mathbf{r}) - \Psi_-^\dagger(\mathbf{r})\Psi_-(\mathbf{r}) \tag{9}$$

According to (5), the operator associated with a measurement performed along a direction $O\varphi$ of the xOy plane making an angle φ with the Ox axis is written as:

$$\sigma_\varphi(\mathbf{r}) = e^{-i\varphi}\Psi_+^\dagger(\mathbf{r})\Psi_-(\mathbf{r}) + \;e^{i\varphi}\Psi_-^\dagger(\mathbf{r})\Psi_+(\mathbf{r}) \tag{10}$$

The measurements we are interested in pertain both to the position of the particles and their spin: for each measurement, the position is measured in an infinitesimal volume Δ centered at point $\mathbf{r}_j$ and, when measuring the direction of the spin along the $O\varphi_j$ direction, we obtain the result $\eta_j = \pm 1$. By analogy with (7), the projector associated with such a measurement can be written:

$$\begin{aligned} P_{\eta_j}(\mathbf{r}_j,\varphi_j) &= \frac{1}{2}\int_\Delta \mathrm{d}^3r' \;\left[n(\mathbf{r}') + \eta_j\;\sigma_{\varphi_j}(\mathbf{r}')\right] \\ &\simeq \frac{\Delta}{2}\left[n(\mathbf{r}_j) + \eta_j\;\sigma_{\varphi_j}(\mathbf{r}_j)\right] \end{aligned} \tag{11}$$

where $n(\mathbf{r})$ and $\sigma_\varphi(\mathbf{r})$ are given by (8) and (10). Operator $P_{\eta_j}(\mathbf{r}_j,\varphi_j)$ projects both the orbital variables onto this small domain and the spins onto the eigenstate of the component along the $O\varphi_j$ axis, with eigenvalue $\eta_j = \pm 1$.

2. Probabilities of the different measurement results

Consider now a system of $2N$ bosons in the state (2). Spin measurements are performed in a series of regions of space, which cover the whole extension of the orbital wave function $u(\mathbf{r})$ without overlapping. The measurements are supposed to be ideal, so that every particle is detected. The regions are supposed to be sufficiently small to obtain a negligible probability of double detection in any of them. Those where a particle is actually detected are centered at $\mathbf{r}_j$ (with $j = 1, 2, .., 2N$), as illustrated in Figure 1.

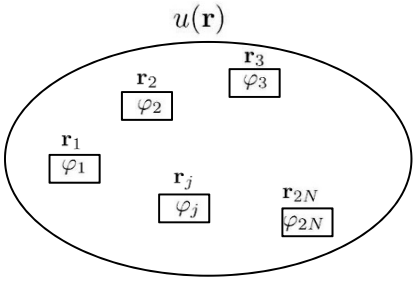


Figure 1: Two condensates, each having N particles, the first one with + spins, and the other with − spins, share the same orbital wave function $u(\mathbf{r})$, represented by the oval in the figure. Measurements of the transverse direction of the spins are performed in $2N$ non overlapping regions of space, centered at points $\mathbf{r}_j$ (with $j = 1, 2, .., 2N$). In each region, the measurement is performed along a transverse direction (perpendicular to the quantization axis), defined by an angle φ_j, which may depend on j; the corresponding result is $\eta_j = \pm 1$.

In each of these regions, one performs a measurement of the spin component along a transverse direction (perpendicular to the quantization axis), defined by the angle φ_j, and the measurement result is $\eta_j = \pm 1$. We now calculate the probability of getting a series of results $\eta_j = \pm 1$ in those $2N$ regions.

2-a. A first expression for the probability

The associated projectors $P_{\eta_j}(\mathbf{r}_j, \varphi_j)$ all commute with each other, since they contain field operators (and their adjoints) at different points in space (we assume that all measurements are done simultanously, or separated by a very short time). The probability $\mathcal{P}_{2N}$ of a result is therefore the average value in the state $|\Phi_0\rangle$ of the product of projectors:

$$\begin{aligned}\mathcal{P}_{2N}\left(\{\eta_i,\varphi_i\}\right) &= \langle\Phi_0| \prod_{j=1}^{2N} P_{\eta_j}(\mathbf{r}_j,\varphi_j)\,|\Phi_0\rangle \\ &= \left(\frac{\Delta}{2}\right)^{2N} \langle\Phi_0| \prod_{j=1}^{2N} \Big\{ \Psi_+^\dagger(\mathbf{r}_j)\Psi_+(\mathbf{r}_j) + \Psi_-^\dagger(\mathbf{r}_j)\Psi_-(\mathbf{r}_j) + \eta_j e^{-i\varphi_j}\Psi_+^\dagger(\mathbf{r}_j)\Psi_-(\mathbf{r}_j) \\ &\qquad + \eta_j e^{i\varphi_j}\Psi_-^\dagger(\mathbf{r}_j)\Psi_+(\mathbf{r}_j)\Big\}\,|\Phi_0\rangle \end{aligned} \tag{12}$$

where $\{\eta_i, \varphi_i\}$ symbolizes the ensemble of the variables $(\eta_1, \varphi_1, .., \eta_N, \varphi_N)$. As the operators commute, we can also move all the field operators $\Psi_\pm(\mathbf{r}_j)$ towards the right, and their adjoints $\Psi_\pm^\dagger(\mathbf{r}_j)$ towards the left. We now introduce a basis $u_i(\mathbf{r})$ for the wave

functions, its first element being the wave function of the populated states (1). This basis allows expanding the field operators according to relation (A-14) of Chapter XVI, and we can write:

$$\Psi_{\pm}(\mathbf{r}_j) = \sum_i u_i(\mathbf{r}_j)\ a_{u_i,\pm} \tag{13}$$

where $a_{u_i,\pm}$ is the annihilation operator of a particle in the individual state $|u_i,\pm\rangle$. Now the only term that actually plays a role in this expansion is the $i=1$ term: all the other $i \neq 1$ terms yield zero when acting on the state $|\Phi_0\rangle$, which only contains particles in the orbital state $u_1(\mathbf{r}) = u(\mathbf{r})$. One can simply replace $\Psi_{\pm}(\mathbf{r}_j)$ by $u(\mathbf{r}_j)\, a_{u,\pm}$. The same is true for the adjoint of the field operators which, acting on the bra placed on their left, can only destroy particles in a state previously populated; consequently, we can also replace $\Psi^{\dagger}_{\pm}(\mathbf{r}_j)$ by $u^*(\mathbf{r}_j)\, a^{\dagger}_{u,\pm}$. Once these replacements have been performed, we get an expression that can be written, in a symbolic way, as:

$$\begin{aligned}\mathcal{P}_{2N}(\{\eta_i,\varphi_i\}) = \\ \left(\frac{\Delta}{2}\right)^{2N} \langle\Phi_0| : \prod_{j=1}^{2N} |u(\mathbf{r}_j)|^2 \Big[a^{\dagger}_{u,+}a_{u,+} + a^{\dagger}_{u,-}a_{u,-} \\ +\eta_j e^{-i\varphi_j} a^{\dagger}_{u,+}a_{u,-} + \eta_j e^{i\varphi_j} a^{\dagger}_{u,-}a_{u,+}\Big] : |\Phi_0\rangle\end{aligned} \tag{14}$$

The two dots surrounding the product over j in this symbolic writing express the following convention, which originates from the rearrangement of the operators $\Psi_{\pm}(\mathbf{r}_j)$ and $\Psi^{\dagger}_{\pm}(\mathbf{r}_j)$ mentioned above: in each of the 4^{2N} terms of the product of sums, all the annihilation operators $a_{u,\pm}$ are regrouped towards the right, and all the creation operators $a^{\dagger}_{u,\pm}$ towards the left. To obtain the probability we are looking for, we have to compute the average values in the state $|\Phi_0\rangle$ of 4^{2N} products, in normal order (Complement $\mathrm{B_{XVI}}$, § 1-a-α), of creation and annihilation operators in the two states $|u_{i=1},\pm\rangle$.

The situation now becomes very similar to that leading to relation (37) in Complement $\mathrm{C_{XXI}}$. The computation that follows is indeed similar, except for the fact that we no longer use the approximation (35) of that complement (number of measurements small compared to the particle number): we now assume that all the particles are measured. Actually, most of the terms of the product over j appearing in (14) have a zero average value in the state $|\Phi_0\rangle$. The only relevant terms are those that contain exactly N annihilation operators $a_{u,+}$ and N other annihilation operators $a_{u,-}$, in which case their action yields the vacuum, hence a normalized ket; if this is not the case, the result is zero. For a similar reason, they must also contain exactly N creation operators $a^{\dagger}_{u,+}$ and N other $a^{\dagger}_{u,-}$, otherwise the result is zero. All these non-zero terms have the same average values, since the product of operators in the normal order introduces each time the same factor $\left(\sqrt{N!}\times\sqrt{N!}\right)^2$, that is $(N!)^2$; we also get the product of $2N$ coefficients F_{q_j,q'_j}, which can take one of these 4 values:

$$F_{+1,+1} = F_{-1,-1} = 1 \tag{15}$$

and:

$$F_{+1,-1} = \eta_j e^{-i\varphi_j} \qquad F_{-1,+1} = \eta_j e^{i\varphi_j} \tag{16}$$

Now $F_{+1,+1}$ corresponds to a term associated with a particle destruction in $|u,+\rangle$, followed by a particle creation in that same state. In the same way, $F_{-1,-1}$ corresponds to an annihilation-creation in the individual state $|u,-\rangle$. Finally, $F_{+1,-1}$ corresponds to an annihilation in state $|u,-\rangle$ followed by a creation in state $|u,+\rangle$, and the opposite for $F_{-1,+1}$. All the non-zero terms therefore correspond to products of $2N$ numbers F_{q_j,q'_j} such that the sum of the q_i and the sum of the q'_j are both zero; this condition automatically ensures the conservation of the particle number in each individual state. The final result is:

$$\mathcal{P}_{2N}\left(\{\eta_i,\varphi_i\}\right) = \left(\frac{\Delta}{2}\right)^{2N} (N!)^2 \prod_{j=1}^{2N} \Big[\, |u(\mathbf{r}_j)|^2 \sum_{q_j,q'_j=\pm 1} F_{q_j,q'_j} \Big] \quad \text{with } S=S'=0 \tag{17}$$

where, in the right hand side, we retain only the terms satisfying the double condition:

$$\begin{aligned} &\sum_j q_j = S = 0 \\ &\sum_j q'_j = S' = 0 \end{aligned} \tag{18}$$

2-b. Introduction of the phase and the quantum angle

Because of the summation constraints, expression (17) is not easy to handle. This is why we introduce two delta functions $\delta_{S,0}$ and $\delta_{S',0}$, which obey:

$$\begin{aligned} \delta_{S,0} &= \int_{-\pi}^{+\pi} \frac{\mathrm{d}\xi}{2\pi}\, e^{iS\xi} \\ \delta_{S',0} &= \int_{-\pi}^{+\pi} \frac{\mathrm{d}\xi'}{2\pi}\, e^{iS'\xi'} \end{aligned} \tag{19}$$

This amounts to multiplying in (17) each F_{q_j,q'_j} by $e^{i\left(q_j\xi+q'_j\xi'\right)}$ and integrating over the two angles ξ and ξ'. This enables us to write the probability in the form:

$$\begin{aligned} &\mathcal{P}_{2N}\left(\{\eta_i,\varphi_i\}\right) = \\ &\quad \left(\frac{\Delta}{2}\right)^{2N} (N!)^2 \int_{-\pi}^{+\pi} \frac{\mathrm{d}\xi}{2\pi} \int_{-\pi}^{+\pi} \frac{\mathrm{d}\xi'}{2\pi} \prod_{j=1}^{2N} |u(\mathbf{r}_j)|^2 \sum_{q_j,q'_j} e^{i\left(q_j\xi+q'_j\xi'\right)} F_{q_j,q'_j} \end{aligned} \tag{20}$$

where the summations over q_j and q'_j are now independent, thanks to relations (19) that automatically ensure the constraints are obeyed. For each value of j, each sum contributes the factor:

$$\begin{aligned} &F_{+1,+1}\, e^{i(\xi+\xi')} + F_{-1,-1}\, e^{-i(\xi+\xi')} + F_{+1,-1}\, e^{i(\xi-\xi')} + F_{-1,+1}\, e^{i(\xi'-\xi)} \\ &\qquad = 2\cos\left(\xi+\xi'\right) + 2\eta_j \cos\left(\xi-\xi'-\varphi_j\right) \end{aligned} \tag{21}$$

We finally make the change of variables[1]:

$$\begin{aligned}\Lambda &= \xi + \xi' \\ \lambda &= \xi - \xi'\end{aligned} \tag{22}$$

This leads to the simpler expression:

$$\mathcal{P}_{2N}\left(\{\eta_i, \varphi_i\}\right) = \\ (\Delta)^{2N}\,(N!)^2 \int_{-\pi}^{+\pi} \frac{\mathrm{d}\Lambda}{2\pi} \int_{-\pi}^{+\pi} \frac{\mathrm{d}\lambda}{2\pi} \prod_{j=1}^{2N} \left|u(\mathbf{r}_j)\right|^2 \left[\cos\Lambda + \eta_j \cos\left(\lambda - \varphi_j\right)\right] \tag{23}$$

The following discussion is entirely based on this result. For reasons that will be explained below, λ is called the phase, whereas Λ is called the "quantum angle".

Comment:

It will be useful in what follows to note that the right-hand side of the above equality stays the same if we make the change:

$$\int_{-\pi}^{+\pi} \frac{\mathrm{d}\Lambda}{2\pi} \Rightarrow 2 \int_{-\pi/2}^{+\pi/2} \frac{\mathrm{d}\Lambda}{2\pi} \tag{24}$$

To show this, we can decompose the integral over $\mathrm{d}\Lambda$ in a sum $I = I_1 + I_2$, where I_1 is the integral between $-\pi/2$ and $+\pi/2$, and I_2, the integral between $\pi/2$ and $3\pi/2$ (as the period of the function to be integrated is 2π, any integration domain covering the entire circle is equivalent). Now I_2 is just equal to I_1. This is because the function to be integrated is multiplied by $(-1)^{2N} = 1$ when one changes Λ to $\Lambda' = \Lambda - \pi$ as well as λ to $\lambda' = \lambda - \pi$. Consequently, changing the integration variables Λ, λ to Λ', λ' allows giving to I_2 the same integration domain as I_1.

3. Discussion

Let us examine first the case where the number of measurements is negligible compared to the particle number in each condensate; this will enable us to compare the results obtained with those of Complement $\mathrm{C_{XXI}}$.

3-a. Measurement number $Q \ll 2N$

We first recall a general property of quantum mechanics concerning compatible observables (Chapter III, § C-6-a). When several operators A, B, C, etc. commute with each other, one can build a basis with their common eigenvectors. The scalar product of these eigenvectors and the system state vector yields the probability amplitude for finding the system in each of these eigenvectors. If the eigenvalues are non-degenerate, the squared modulus of this amplitude yields the probability of finding the corresponding eigenvalues upon a series of simultaneous measurements associated with all the operators

[1]The Jacobian of this change of variables is equal to 2, and this factor should be introduced in the denominator. Nevertheless, since the integrated function is periodic, this factor can be taken into account by reducing the integration domain of the two variables λ and Λ to the interval $-\pi$, $+\pi$, which reduces the area of the integration domain by a factor 2.

A, B, C, etc. If they are degenerate, we just have to sum the probabilities over all the orthogonal eigenkets. This is the rule we have followed until now in this complement.

Imagine now that we ignore the measurement results associated with one or several operators of the series, for example B and C; the probabilities of the measurement results we still consider relevant are then simply the sum over the possible results associated with the ignored measurements (sum of the probabilities of exclusive events). One can also imagine another situation where the quantities associated with B and C are actually never measured. The possible series of eigenvalues of the measured operators are then less numerous than in the previous case (since a smaller number of measurements is performed), which increases the degree of degeneracy of these eigenvalues. As for the eigenvalues of B and C, even though they do not correspond to actual measurements, they can still be used as indices to distinguish between the different orthogonal eigenvectors associated with measurement results of operators A, D, etc. Consequently we still have to sum the probabilities on these different eigenvalues, just as we did in the case where these measurements were ignored. This means that quantum mechanics yields the same probabilities whether we assume that the measurement results of B and C are ignored or have never been measured.

Let us now compute the probability $\mathcal{P}_Q$ of obtaining results $\eta_1, .., \eta_Q$ when performing Q measurements on the spins. As we already know the probability (23) corresponding to the case $Q = 2N$, we can consider that all the $2N$ measurements have been performed, but that we ignore the results of $2N-Q$ among them. As we just discussed, this amounts to summing in (23) the probabilities of the two possible results for each of these $2N-Q$ ignored measurements, i.e. the probabilities associated with two opposite values of the η_j. It follows that in the product over j in (23), $\cos(\lambda - \varphi_j)$ will disappear from all the $2N-Q$ terms, leaving only $\cos\Lambda$. We get for $\mathcal{P}_Q$ the following expression (omitting from now on the numerical factors, which are not relevant for our discussion):

$$\mathcal{P}_Q(\{\eta_i, \varphi_i\}) \propto \int_{-\pi/2}^{+\pi/2} \frac{\mathrm{d}\Lambda}{2\pi} [\cos\Lambda]^{2N-Q} \int_{-\pi}^{+\pi} \frac{\mathrm{d}\lambda}{2\pi} \prod_{j=1}^{Q} |u(\mathbf{r}_j)|^2 [\cos\Lambda + \eta_j \cos(\lambda - \varphi_j)] \tag{25}$$

In this expression, the notation $(\{\eta_i, \varphi_i\})$ now stands for Q pairs of variables (η, φ), instead of $2N$ as before. The integral over $\mathrm{d}\Lambda$ contains the function $\cos\Lambda$ to the power $2N - Q$; if $Q \ll 2N$, this power is very high, and the function becomes a very narrow peak centered at $\Lambda = 0$. This allows us to write:

$$\mathcal{P}_Q(\{\eta_i, \varphi_i\}) \propto \int_{-\pi}^{+\pi} \frac{\mathrm{d}\lambda}{2\pi} \prod_{j=1}^{Q} |u(\mathbf{r}_j)|^2 [1 + \eta_j \cos(\lambda - \varphi_j)] \tag{26}$$

This result is very similar to the one obtained in relation (47) of Complement C_{XXI}, namely a product of two positive individual probabilities[2]:

$$p_j(\lambda, \varphi_j) = \frac{1}{2} |u(\mathbf{r}_j)|^2 [1 + \eta_j \cos(\lambda - \varphi_j)] \tag{27}$$

[2]Thanks to the factor 1/2, the sum of the two probabilities ($\eta_j = \pm 1$) is normalized to the probability of presence of a particle in the detection volume.

This product is then averaged over the angle λ, which can take on any value between $-\pi$ and $+\pi$. As we show below, probability (27) is actually the probability of finding the result η_j when measuring the component of a single spin along an axis with direction φ_j, assuming that spin was initially polarized along a direction defined by the angle λ.

Demonstration:

We call Oz the spin quantization axis, and consider a spin polarized in the (transverse) direction in the plane xOy, making an angle λ with the Ox axis. Relation (6) indicates that its state is then:

$$|\psi_\lambda\rangle = \frac{1}{\sqrt{2}}\left[e^{-i\lambda/2}\,|+\rangle + e^{+i\lambda/2}\,|-\rangle\right] \tag{28}$$

When measuring the spin component along an axis defined by the angle φ, the state associated with the $\eta = +1$ measurement result is:

$$|\psi_\varphi\rangle = \frac{1}{\sqrt{2}}\left[e^{-i\varphi/2}\,|+\rangle + e^{+i\varphi/2}\,|-\rangle\right] \tag{29}$$

Consequently, the probability of that result is:

$$\mathcal{P}_{\eta=+1} = |\langle\psi_\varphi|\psi_\lambda\rangle|^2 = \frac{1}{4}\left[e^{i(\lambda-\varphi)/2} + e^{-i(\lambda-\varphi)/2}\right]^2 = \cos^2\frac{\lambda-\varphi}{2} \tag{30}$$

$$= \frac{1}{2}\left[1+\cos(\lambda-\varphi)\right] \tag{31}$$

As for the probability of the $\eta = -1$ result, it is simply the complementary probability, obtained by changing the sign in front of $\cos(\lambda-\varphi)$. In both cases, we get the probability given by (27).

This means that λ can be interpreted as the relative phase between the two condensates. In this case, the predictions of quantum mechanics are identical to the predictions of a theory where the phase would be considered as a classical quantity, perfectly determined but as yet unknown at the beginning of the experiment. From such a point of view, this phase would be revealed more and more precisely by the successive measurements, instead of being created as assumed in the standard quantum mechanics interpretation. Therein lies a link to the heart of the EPR argument.

3-b. Macroscopic EPR argument

The EPR argument was presented in § F-1 of Chapter XXI. It is based on the double hypothesis of reality and locality, as well as on the assumption that all quantum mechanical predictions are correct. The conclusion of the argument is that quantum mechanics is necessarily incomplete; to render it complete, "elements of reality" must be added to it. In an EPRB experiment, involving two spins in a singlet state, these elements of reality can be spin directions, well defined even before any measurements has been performed.

Such an addition necessarily falls outside the framework of standard quantum mechanics. Bohr was opposed to it; he argued that the concept of elements of reality proposed by EPR could not be relevant for microscopic systems, since it was meaningless to try and dissociate them, conceptually, from their experimental surrounding. As we

discussed in Chapter XXI, this position is logically sound; it allows invalidating the conclusions of the EPR argument. However, we are going to show that the double condensates offer another context for applying the EPR reasoning, particularly interesting as it involves macroscopic quantities (as well as the conservation of angular momentum). These physical quantities can, in principle, be on our scale, thereby making it more difficult to deny them an independent physical reality.

We consider a physical system in a quantum state similar to (2), where the two $|\pm\rangle$ internal states of the particles are eigenstates of the spin components along the Oz quantization axis, for example the $m = 0$ and $m = 1$ states of a spin $S = 1$. For the clarity of the discussion, we will assume that the orbital wave functions of each condensate are distinct but overlap in two regions of space, as schematized in Figure 2. In each of these two regions (which may be separated by an arbitrarily large distance), two observers, Alice and Bob, perform measurements of the spin components along transverse directions[3], N_A measurements for Alice, N_B for Bob. For each measurement performed, each of the observers chooses an arbitrary direction defined by an angle φ_i; Alice's choices are completely independent of those of Bob, and vice versa. A first series of measurements ($1 \leq i \leq N_A$) is performed by Alice in a first region of space; right after that, Bob performs another series ($N_A + 1 \leq i \leq N_A + N_B$) in his own laboratory, located very far away.

Now we saw that, as soon as Alice has measured the spins of a few particles, the relative phase of the two condensates in the entire space is fixed with a fairly good precision (the larger the number of measurements, the better the precision). These measurements also fix the transverse direction of the spins. Alice cannot, however, decide what this direction will be, as it is fixed in a totally random way in the measurement process. Standard quantum mechanics then predicts that when Bob will perform his own measurements, it is practically certain (within a negligible error) that he will find the same relative phase. As he can perform a large number of measurements, he can find out, practically instantaneously, the spin direction created and observed by Alice. The EPR argument underlines that, as no interaction had time to propagate from Alice to Bob, it is not possible for this transverse orientation to have been created by Alice's measurements: it necessarily existed prior any measurements.

What is new in our case, compared to the two-spin case, is that Bob's observations may concern an arbitrarily large number of spins; his experiment then amounts to measuring the angular momentum direction of a macroscopic spin system, which has an arbitrarily large angular momentum. As we are now dealing with macroscopic quantities, one can no longer argue, as Bohr did, that the microscopic world is accessible neither to human experimentation nor to human language description. In our present case, it seems more artificial to refuse, as suggested by Bohr, to consider separately the physical properties of systems located in distinct regions of space. The EPR argument becomes harder to refute. Reference [90] contains a discussion of this unexpected situation in terms of conservation of angular momentum.

[3]The longitudinal direction is the direction of the spin polarization in the initial state (2), the transverse directions are all the perpendicular directions.

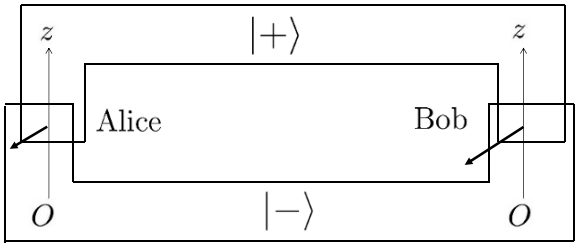


Figure 2: Scheme of an experiment on a double spin condensate, one in a + internal state, the other in a − state. The two condensates have distinct orbital wave functions, overlapping in two regions of space where two observers, Alice and Bob, perform measurements. These two regions can be separated by an arbitrarily large distance. Alice and Bob measure the spins one by one, choosing each time a transverse component (perpendicular to the Oz quantization axis) defined by an angle φ. Whereas, initially, the two condensates have no relative phase, the first measurements performed by Alice create one. The paradox is that this phase propagates instantaneously to the remote region where Bob performs his measurements. This is reminiscent of the EPR argument, but in a more striking case where the EPR "elements of reality" can be macroscopic.

3-c. Measuring all the spins, violation of Bell's inequalities

Imagine now that we measure all the spins; the selective effect around the $\Lambda = 0$ value does not occur any longer. The interpretation in terms of probabilities of individual events is no longer possible: the factor $[\cos\Lambda + \eta_j \cos(\lambda - \varphi_j)]$ in (23) can sometimes take on negative values, which rules out its possibility to be considered as a standard probability. Actually, it is not unusual in quantum mechanics that purely quantum effects arise through "negative probabilities ", as in this case. The angle Λ is called the "quantum angle ", which underlines that its role is to introduce such quantum effects, such as non-locality effects and violations of the Bell's inequalities.

To prove that such violations occur for any value of N requires using relation (23), which involves many parameters (all the measuring angles, which are arbitrary); it is easier to perform a numerical calculation as explained in the second reference of [90]. Our objective here is to simply show that the phase does not always behave in a classical way. This is why, without presenting the numerical calculation, we shall study the behavior of expression (23) for the simple case of two measurements on two spins ($Q = 2$ and $N = 1$). This will enable us to show that this expression does predict the existence of violations of the inequalities, for certain cases (more general cases are treated in the above reference). Clearly, in the case of two spins we could have carried out this computation in a simpler and more direct way.

Using definition (A-7) of Chapter XV for the Fock states, the state (2) can be

written in the form:

$$|\Psi\rangle = \left(a_{u,+}\right)^{\dagger}\left(a_{u,-}\right)^{\dagger}|0\rangle = \frac{1}{\sqrt{2}}\left[|1:u,+;2:u,-\rangle + |1:u,-;2:u,+\rangle\right]$$
$$= |1:u;2:u\rangle \otimes \frac{1}{\sqrt{2}}\left[|1:+;2:-\rangle + |1:-;2:+\rangle\right] \tag{32}$$

This leads to an entangled spin state, very similar to the one considered in § B of Chapter XXI. The only difference is the + sign in the present spin state, instead of the − in the singlet state considered in that chapter, but this difference is of no great consequence (we shall come back to this point more precisely – see note 4). Such a state can be expected to lead to significant quantum effects, as for example to situations violating Bell's inequalities.

We can also go back to the general relation (23) to show that it indeed leads to violations of Bell's inequalities in this simple case. For $N = 1$, this relation becomes:

$$\mathcal{P}_2\left(\eta_1, \varphi_1; \eta_2, \varphi_2\right)$$
$$\propto \int_{-\pi}^{+\pi} \frac{\mathrm{d}\Lambda}{2\pi} \int_{-\pi}^{+\pi} \frac{\mathrm{d}\lambda}{2\pi} \left[\cos\Lambda + \eta_1 \cos\left(\lambda - \varphi_1\right)\right]\left[\cos\Lambda + \eta_2 \cos\left(\lambda - \varphi_2\right)\right]$$
$$= \frac{1}{2}\left[1 + \eta_1\eta_2 \ \cos\varphi_1 \cos\varphi_2 + \eta_1\eta_2 \ \sin\varphi_1 \sin\varphi_2\right] \tag{33}$$

where we have used the fact that the average value on the circle of cosine squared or sine squared is equal to 1/2, whereas the average value of the product of cosine and sine is zero. We obtain:

$$\mathcal{P}_2\left(\eta_1, \varphi_1; \eta_2, \varphi_2\right) \propto \frac{1}{2}\left[1 + \eta_1\eta_2 \ \cos\left(\varphi_1 - \varphi_2\right)\right] \tag{34}$$

Normalizing to unity the sum of the 4 probabilities obtained for $\eta_1 = \pm 1$ and $\eta_2 = \pm 1$, we finally get:

$$\mathcal{P}_2\left(+1,+1\right) = \mathcal{P}_2\left(-1,-1\right) = \frac{1}{2}\cos^2\left(\frac{\varphi_1 - \varphi_2}{2}\right)$$
$$\mathcal{P}_2\left(+1,-1\right) = \mathcal{P}_2\left(-1,+1\right) = \frac{1}{2}\sin^2\left(\frac{\varphi_1 - \varphi_2}{2}\right) \tag{35}$$

These relations are very similar to equalities (B-10) of Chapter XXI, if we make the change[4]:

$$\varphi_1 \Rightarrow \theta_A + \pi$$
$$\varphi_2 \Rightarrow \theta_B \tag{36}$$

The angles φ_1 and φ_2 now play the role of the analyzer orientation angles in Figure 2 of that chapter. Now we know (Chapter XXI, § F-3) that these equalities lead to significant violations of Bell's inequalities (by a factor $\sqrt{2}$), hence to marked non-local quantum effects; such effects should therefore be expected in our present case.

[4]This change results from the + sign in the spin state (32), instead of the − sign in the singlet state.

In the general case where N can take on any value, measuring the totality of the spins may lead to strong violations of Bell's inequalities, provided the measurement angles are judiciously chosen (for large N, the angular domain leading to such violations decreases as the inverse of the square root of that number [90]). Note however that these violations will disappear as soon as certain spins are no longer measured (or their corresponding results no longer taken into account, which amounts to the same thing).

Conclusion

This complement is an illustration of the limits of the phenomenon studied in the previous complement: the process of successive measurements builds up a phase that has all the properties of a classical phase, but only up to a certain point. If all the particles are measured, and for certain particular choices of the measuring angles, in an ideal experiment the phase should exhibit some distinctly quantum properties.

Furthermore, one could have expected that the extreme quantum properties, as for example their non-local aspects discussed in §§ F-1 and F-3 of Chapter XXI, would only concern systems with a small particle number, or in singlet spin states (these having a special status among all the states accessible to a physical system). The present complement shows that this is not at all the case: in principle, the same properties should exist for systems composed of a very large number of particles, in a fairly simple quantum state (a double condensate).

Appendix IV

Feynman path integral

1 Quantum propagator of a particle 2267
1-a Expressing the propagator as a sum of products 2268
1-b Calculation of the matrix elements 2268
2 Interpretation in terms of classical histories 2272
2-a Expressing the propagator as a function of classical actions . 2272
2-b Generalization: several particles interacting via a potential . 2274
3 Discussion; a new quantization rule 2274
3-a Analogy with classical optics 2274
3-b A new quantization rule 2275
4 Operators 2276
4-a One single operator 2276
4-b Several operators 2279

In Chapter III, we introduced the postulates of quantum mechanics using the Hamiltonian approach, with quantization rules applied to conjugate Hamiltonian variables. It is however possible to introduce quantum mechanics and its quantization rules in an entirely different way, starting from a classical Lagrangian, and using Feynman path integrals. This approach provides an interesting insight to the relationship between classical and quantum physics, reminiscent of the connections between geometric and wave optics. Furthermore, this approach is preferable in a certain number of cases, in particular for situations where we know the classical Lagrangian but not the conjugate variables necessary to define a Hamiltonian[1].

This appendix is an elementary introduction to Feynman path integrals, and some of their properties, without too much concern for mathematical rigor. We first study in § 1 the quantum propagator of a particle, and then show in § 2 how to express it as the sum of contributions coming from different classical histories (possible evolutions) of the physical system. Once these results have been established, we discuss in § 3 how to take the inverse point of view, and start from these classical histories to derive the usual form of quantum mechanics, its quantization rules, its propagators and, in § 4, its operators. For the sake of simplicity, we shall only consider an ensemble of particles interacting via a position dependent potential; for the study of more general cases (vector potential, commutative or non-commutative gauge invariance), the reader may consult references [91], [92], or [93].

1. Quantum propagator of a particle

Consider a spinless particle. The propagator $G(\mathbf{r}', t'; \mathbf{r}, t)$ of this particle is defined as:

$$G(\mathbf{r}', t'; \mathbf{r}, t) = \langle \mathbf{r}' | U(t', t) | \mathbf{r} \rangle \tag{1}$$

[1] This happens, for example, if the Lagrangian does not depend on the time derivative of a coordinate q_i, in which case one cannot define the conjugate momentum p_i.

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.
© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

where the kets $|\mathbf{r}\rangle$ are the eigenkets of the position operator, and $U(t',t)$ the evolution operator between the initial time t and the final time t' (Complement $\mathrm{F_{III}}$). This propagator gives the probability amplitude for the particle, starting at time t from a state localized at point $\mathbf{r}$, to be found at a later time t' at point $\mathbf{r}'$.

1-a. Expressing the propagator as a sum of products

We can split the time interval $[t,t']$ into $\mathcal{N}$ equal smaller segments by setting:

$$\delta t = \frac{t'-t}{\mathcal{N}} \tag{2}$$

We therefore introduce $\mathcal{N}-1$ intermediate times $t_1,\ t_2,\ ..,\ t_k,\ ..,t_{\mathcal{N}-1}$ (see Figure 1), which are equal to:

$$t_k = t + k\ \delta t \qquad k = 1,2,..,\mathcal{N}-1 \tag{3}$$

It will also be useful for what follows to set $t_0 = t$ and $t_{\mathcal{N}} = t'$. We can then express $U(t',t)$ as a product of $\mathcal{N}$ terms:

$$U(t',t) = U(t_{\mathcal{N}},t_{\mathcal{N}-1})\ ...\ U(t_k,t_{k-1})\ ...\ U(t_2,t_1)\,U(t_1,t_0) \tag{4}$$

Between each evolution operators, we introduce a closure relation in the $\{|\mathbf{r}\rangle\}$ basis:

$$\begin{aligned} U(t',t) = \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2 .. \int \mathrm{d}^3 r_k .. \int \mathrm{d}^3 r_{\mathcal{N}-1} \\ U(t_{\mathcal{N}},t_{\mathcal{N}-1})\,|\mathbf{r}_{\mathcal{N}-1}\rangle\,\langle\mathbf{r}_{\mathcal{N}-1}|\,U(t_{\mathcal{N}-1},t_{\mathcal{N}-2})\,|\mathbf{r}_{\mathcal{N}-2}\rangle\ ... \\ \langle\mathbf{r}_k|\,U(t_k,t_{k-1})\,|\mathbf{r}_{k-1}\rangle\ ...\ \langle\mathbf{r}_1|\,U(t_1,t_0) \end{aligned} \tag{5}$$

Inserting this equality in (1), the propagator is expressed as:

$$\begin{aligned} G(\mathbf{r}',t';\mathbf{r},t) = \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2 .. \int \mathrm{d}^3 r_k .. \int \mathrm{d}^3 r_{\mathcal{N}-1} \\ \langle\mathbf{r}'|\,U(t_{\mathcal{N}},t_{\mathcal{N}-1})\,|\mathbf{r}_{\mathcal{N}-1}\rangle\,\langle\mathbf{r}_{\mathcal{N}-1}|\,U(t_{\mathcal{N}-1},t_{\mathcal{N}-2})\,|\mathbf{r}_{\mathcal{N}-2}\rangle\ ... \\ ...\ \langle\mathbf{r}_k|\,U(t_k,t_{k-1})\,|\mathbf{r}_{k-1}\rangle\ ...\ \langle\mathbf{r}_1|\,U(t_1,t_0)\,|\mathbf{r}\rangle \end{aligned} \tag{6}$$

We now let δt tend towards zero (or, equivalently, $\mathcal{N}$ towards infinity). The number of integrations over the $\mathrm{d}^3 r_k$ will then tend towards infinity but the matrix elements are now those of the evolution operator over an infinitesimal time δt.

1-b. Calculation of the matrix elements

We now compute the matrix elements $\langle\mathbf{r}_k|\,U(t_k,t_{k-1})\,|\mathbf{r}_{k-1}\rangle$ of the evolution operators, with $t_k = t_{k-1} + \delta t$. The particle's Hamiltonian H is the sum of the kinetic energy and the energy associated with an external potential:

$$H = \frac{\mathbf{P}^2}{2m} + V(\mathbf{R}) \tag{7}$$

where $\mathbf{P}$ and $\mathbf{R}$ are, respectively, the momentum and position operators of the particle; m is its mass.

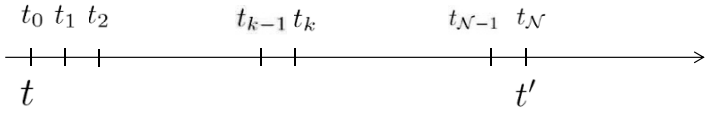


Figure 1: To express the evolution operator as a product of operators, one introduces between the initial time t and the final time t' a whole series of intermediate times t_1, t_2, .., t_k, ..,$t_{\mathcal{N}-1}$; an evolution operator is associated with each of the N time intervals. Introducing at each time t_k a closure relation then leads to a summation over all the possible positions $\mathbf{r}_k$ of the particle at that time.

α. *Free particle*

For a free particle, the Hamiltonian is simply the kinetic energy. Introducing a closure relation on the momentum eigenstates, we can write:

$$\langle \mathbf{r}_k |\, e^{-i\mathbf{P}^2\delta t/2m\hbar}\, |\mathbf{r}_{k-1}\rangle = \frac{1}{(2\pi)^3}\int \mathrm{d}^3k\; e^{i\mathbf{k}\cdot(\mathbf{r}_k-\mathbf{r}_{k-1})}\; e^{-i\hbar k^2\delta t/2m} \tag{8}$$

We show below that this propagator can be computed exactly, whether the time interval δt is infinitesimal or finite; its expression is:

$$\langle \mathbf{r}_k |\, e^{-i\mathbf{P}^2\delta t/2m\hbar}\, |\mathbf{r}_{k-1}\rangle = \left(\frac{m}{2\pi\hbar\;\delta t}\right)^{3/2} e^{-3i\pi/4}\; e^{im(\mathbf{r}_k-\mathbf{r}_{k-1})^2/2\hbar\delta t} \tag{9}$$

Imagine for a moment that in the argument of the last exponential function the factor i is replaced by a simple minus sign (this amounts to switching to an imaginary time). Within a numerical factor, the propagator of the free particle becomes a Gaussian function, decreasing rapidly as soon as the distance $|\mathbf{r}_k - \mathbf{r}_{k-1}|$ becomes large compared to its width $\sqrt{2\hbar\delta t/m}$.

Comment:

This result is not surprising since if one replaces, in Schrödinger's equation, the real time t by an imaginary time it, one gets the diffusion equation whose propagator is indeed a Gaussian. The width of that Gaussian goes to zero as $\delta t \to 0$: as expected, the shorter the time δt, the smaller the distance covered by the particle. It is worth noting, however, that this distance is not proportional to δt, but to the square root of this time; in other words, during very short times, the particle may propagate much further than if it had a constant velocity. This situation is characteristic of a random motion whose correlation time and mean free path tend simultaneously to zero, in a way where the diffusion coefficient remains constant: such a process lead to the classical diffusion equation, and its propagator therefore has the same property.

As for Schrödinger's equation itself (without switching to an imaginary time), when the distance $|\mathbf{r}_k - \mathbf{r}_{k-1}|$ increases, instead of a decrease of the propagator's modulus, we

get an oscillation that gets faster as the distance gets greater; these oscillations are also faster as δt gets shorter. We discuss below how all these phases interfere, as well as the particular role played by the phases associated with classical paths.

Demonstration of relation (9):

We modify the exponent of the function to be integrated in (8) to introduce a perfect square:

$$\frac{\hbar\ \delta t}{2m}\mathbf{k}^2 + \mathbf{k}\cdot(\mathbf{r}_k - \mathbf{r}_{k-1}) = \frac{\hbar\ \delta t}{2m}\left[\mathbf{k} + \frac{m}{\hbar\ \delta t}(\mathbf{r}_k - \mathbf{r}_{k-1})\right]^2 - \frac{m}{2\hbar\ \delta t}(\mathbf{r}_k - \mathbf{r}_{k-1})^2 \tag{10}$$

Making the change of variables:

$$\mathbf{q} = \mathbf{k} + \frac{m}{\hbar\ \delta t}(\mathbf{r}_k - \mathbf{r}_{k-1}) \tag{11}$$

leads us to:

$$\langle \mathbf{r}_k |\, e^{-i\mathbf{P}^2\delta t/2m\hbar}\, |\mathbf{r}_{k-1}\rangle = \frac{1}{(2\pi)^3}\left[\int \mathrm{d}^3q\ e^{-i\hbar q^2\delta t/2m}\right] \times e^{im\left(\mathbf{r}_k - \mathbf{r}_{k-1}\right)^2/2\hbar\delta t} \tag{12}$$

The integral that still remains between the brackets on the right-hand side is now a number independent of $\mathbf{r}_k$ and $\mathbf{r}_{k-1}$. As the three components of the vector $\mathbf{q}$ yield the same contribution, this number is the cube of the integral:

$$I = \int_{-\infty}^{+\infty} \mathrm{d}q_x\ e^{-i\hbar q_x^2\ \delta t/2m} \tag{13}$$

Following a classical procedure, we compute the square of that integral I using the polar coordinates where $\rho = \sqrt{q_x^2 + s_x^2}$:

$$\begin{aligned} I^2 &= \int_{-\infty}^{+\infty} \mathrm{d}q_x \int_{-\infty}^{+\infty} \mathrm{d}s_x\ e^{-i\hbar\left(q_x^2+s_x^2\right)\delta t/2m} \\ &= 2\pi\int_0^{\infty} \rho\ d\rho\ e^{-i\hbar\rho^2\delta t/2m} = \pi\frac{2im}{\hbar\ \delta t}\left[e^{-i\infty} - 1\right] \end{aligned} \tag{14}$$

Now, writing the number $e^{-i\infty}$ is mathematically meaningless: as $M \to \infty$, the imaginary exponential e^{iM} oscillates indefinitely between $+1$ and -1, around a zero average value. We shall simply take this number equal to its zero average value[2], so that:

$$I = \sqrt{\frac{2\pi m}{\hbar\ \delta t}}e^{-i\pi/4} \tag{15}$$

Replacing by I^3 the integral over d^3q in (12), we get relation (9).

[2]One can reach this same result by adding a small imaginary part $-i\epsilon$ to the coefficient $\hbar^2\delta t/2m$ of ρ^2 in the exponent. Thanks to this imaginary part, which can be arbitrarily small, the oscillating term does disappear.

β. Effects of the external potential

We must now compute the matrix elements of the exponential operator $e^{-iH\delta t/\hbar}$ when H is the sum of two non-commuting operators. We shall only consider the case where δt is infinitesimal. In that case, the exponential of a sum of non-commuting operators A and B can be written:

$$e^{-i(A+B)\delta t} = 1 - i\left[A+B\right]\delta t - \left[A^2 + AB + BA + B^2\right]\frac{\delta t^2}{2} + 0\left(\delta t^3\right) \tag{16}$$

We can also expand the following product of exponentials as:

$$e^{-iA\delta t/2}\ e^{-iB\delta t}\ e^{-iA\delta t/2} = \left[1 - iA\frac{\delta t}{2} - A^2\frac{\delta t^2}{8} + ..\right]\times \left[1 - iB\delta t - B^2\frac{\delta t^2}{2} + ..\right]\left[1 - iA\frac{\delta t}{2} - A^2\frac{\delta t^2}{8} + ..\right] \tag{17}$$

This leads to:

$$\begin{aligned} &e^{-iA\delta t/2}\ e^{-iB\delta t}\ e^{-iA\delta t/2} \\ &\quad = 1 - i\left[A+B\right]\delta t - \left[\frac{A^2}{2} + \frac{A^2}{2} + \frac{B^2}{2} + \frac{B^2}{2} + AB + BA\right]\frac{\delta t^2}{2} + 0\left(\delta t^3\right) \\ &\quad = e^{-i(A+B)\delta t} + 0\left(\delta t^3\right) \end{aligned} \tag{18}$$

Setting $A = \mathbf{P}^2/2m\hbar$ and $B = V\left(\mathbf{R}\right)/\hbar$ we get, assuming that δt is infinitesimal so that we can neglect the δt^3 terms:

$$\begin{aligned} \left\langle \mathbf{r}_k\right| U\left(t_k, t_{k-1}\right)\left|\mathbf{r}_{k-1}\right\rangle &= \left\langle \mathbf{r}_k\right| e^{-iV(\mathbf{R})\delta t/2\hbar}\ e^{-i\mathbf{P}^2\delta t/2m\hbar}\ e^{-iV(\mathbf{R})\delta t/2\hbar}\left|\mathbf{r}_{k-1}\right\rangle \\ &= e^{-i[V(\mathbf{r}_k)+V(\mathbf{r}_{k-1})]\delta t/2\hbar}\ \left\langle \mathbf{r}_k\right| e^{-i\mathbf{P}^2\delta t/2m\hbar}\left|\mathbf{r}_{k-1}\right\rangle \end{aligned} \tag{19}$$

This result is simply the product of an exponential including the potential and a term that is the propagator of the free particle. Taking (9) into account, we can write:

$$\begin{aligned} &\left\langle \mathbf{r}_k\right| U\left(t_{k-1}, t_k\right)\left|\mathbf{r}_{k-1}\right\rangle \\ &\quad = \left(\frac{m}{2\pi\hbar\ \delta t}\right)^{3/2} e^{-3i\pi/4}\ e^{im(\mathbf{r}_k - \mathbf{r}_{k-1})^2/2\hbar\delta t} \times e^{-i[V(\mathbf{r}_k)+V(\mathbf{r}_{k-1})]\delta t/2\hbar} \end{aligned} \tag{20}$$

The effect of the external potential is straightforward: it adds an exponential including half the sum of the two potential energies associated, one with the position of the bra, the other with the position of the ket.

γ. Final expression

Inserting (20) into (5), we get the following expression for the propagator:

$$G\left(\mathbf{r}', t'; \mathbf{r}, t\right) = \left(\frac{e^{-i\pi/2}\ m}{2\pi\hbar\ \delta t}\right)^{3\mathcal{N}/2} \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2 .. \int \mathrm{d}^3 r_k .. \int \mathrm{d}^3 r_{\mathcal{N}-1} \exp\left\{ i \sum_{k=1}^{\mathcal{N}-1} \left[\frac{m\left(\mathbf{r}_k - \mathbf{r}_{k-1}\right)^2}{2\hbar\delta t} - \frac{V\left(\mathbf{r}_k\right) + V\left(\mathbf{r}_{k-1}\right)}{2\hbar}\delta t\right]\right\} \tag{21}$$

This equality is valid in the limit where $\delta t \to 0$ (or $\mathcal{N} \to \infty$) since, to establish (19), we neglected the δt^3 terms.

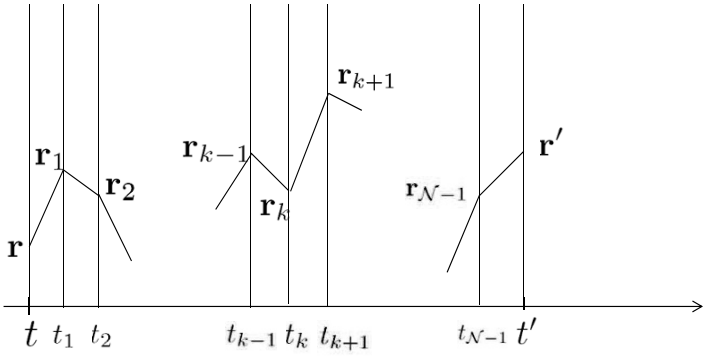


Figure 2: A Feynman path is obtained by associating a position $\mathbf{r}_k$ *with each time* t_k*. The path thus obtained is continuous, but looks in general like a zigzag (the velocity is discontinuous at each time* t_k*). Nevertheless, one can associate a classical action with each of these paths, and get the quantum propagator by summing exponentials of the actions along all these classical paths: while keeping the two end positions* $\mathbf{r}$ *and* $\mathbf{r}'$ *fixed, the summation is performed on all the intermediate positions* $\mathbf{r}_1$, $\mathbf{r}_2$,...,$\mathbf{r}_k$,...$\mathbf{r}_{N-1}$.

2. Interpretation in terms of classical histories

The probability amplitude (6) is obtained by inserting, as many times as necessary, the propagator matrix elements (20). It thus contains as many sums over intermediate positions as there are intermediate times t_k, and this number goes to infinity as $\delta t \to 0$. This expression, somewhat complicated at first sight, actually has a simple interpretation in terms of classical paths the particles can follow between the initial and final times.

2-a. Expressing the propagator as a function of classical actions

Let us go back to classical physics and consider for a moment all the $\mathbf{r}_k$ as being fixed, and a particle that goes through these successive positions at the different times t_k. In between two consecutive times, we assume the particle keeps a constant velocity equal to:

$$\mathbf{v}_k = \frac{\mathbf{r}_k - \mathbf{r}_{k-1}}{\delta t} \tag{22}$$

This defines a classical path Γ, with a linear interpolation for times between the discrete instants t_k (*cf.* Figure 2). The particle's Lagrangian $\mathcal{L}$ is written:

$$\mathcal{L} = \frac{1}{2} m \mathbf{v}^2 - V(\mathbf{r}) \tag{23}$$

For any classical path Γ followed by the particle between the initial time t and the final time t', the position and velocity are both functions of time, and so is the Lagrangian $\mathcal{L}(t)$. The corresponding action is written (Appendix III, § 5-b):

$$S_\Gamma = \int_t^{t'} \mathrm{d}t''\ \mathcal{L}(t'') \tag{24}$$

Let us compute this integral using Riemann's method by introducing, between the times t and t', $\mathcal{N}$ time intervals of length δt. We consider that during an infinitesimal time interval δt, the potential energy can be approximated by half the sum of its values at each end of the interval. We then get:

$$S_\Gamma \simeq \sum_{k=1}^{\mathcal{N}} \delta t \left\{ \left[\frac{m\left(\mathbf{r}_k - \mathbf{r}_{k-1}\right)^2}{2\ \delta t^2} \right] - \frac{V\left(\mathbf{r}_{k-1}\right) + V\left(\mathbf{r}_k\right)}{2} \right\} \tag{25}$$

This approximate equality becomes exact in the limit $\delta t \to 0$.

On the right-hand side of this equality, we find (to within a factor $i/\hbar$) the argument of the exponential appearing in expression (21) for the quantum propagator. This means that this quantum propagator contains the exponential of an approximate value of the classical action, multiplied by $i/\hbar$. When $\delta t \to 0$ (and hence $\mathcal{N} \to \infty$), the approximate value becomes exact[3], and the sums over d^3r_1, d^3r_2,.., $\mathrm{d}^3r_{\mathcal{N}-1}$ on the right-hand side of (21) introduce a summation over all the paths going from $\mathbf{r}, t$ to $\mathbf{r}', t'$:

$$\sum_{\text{paths } \Gamma} \exp\left[\frac{iS_\Gamma}{\hbar}\right] \tag{26}$$

For this summation over the paths to be meaningful, we must now choose a "path density"; we therefore assume that the number of paths in a "path interval", determined by the set of d^3r_k, is given by the product $C^{3\mathcal{N}/2} \times \mathrm{d}^3r_1 \mathrm{d}^3r_2 .. \mathrm{d}^3r_k .. \mathrm{d}^3r_{\mathcal{N}-1}$, where C is the constant (inverse of a length):

$$C = e^{-i\pi/4} \sqrt{\frac{m}{2\pi\hbar\ \delta t}} \tag{27}$$

This allows us to write:

$$\boxed{G\left(\mathbf{r}', t'; \mathbf{r}, t\right) = \left\langle \mathbf{r}' \right| U\left(t', t\right) \left| \mathbf{r} \right\rangle = \sum_{\text{paths } \Gamma} \exp\left[\frac{iS_\Gamma}{\hbar}\right]} \tag{28}$$

In the limit $\delta t \to 0$, the sum over the paths is no longer discrete. This is why, instead of (28), one often writes:

$$\left\langle \mathbf{r}' \right| U\left(t', t\right) \left| \mathbf{r} \right\rangle = \int \mathcal{D}\left[\mathbf{r}\left(t\right)\right]\ \exp\left[\frac{iS_\Gamma}{\hbar}\right] \tag{29}$$

where the notation $\mathcal{D}\left[\mathbf{r}\left(t\right)\right]$ symbolizes the limit of a sum over the paths:

$$\mathcal{D}\left[\mathbf{r}\left(t\right)\right] = \lim_{\mathcal{N}\to\infty} \left[e^{-i\pi/4} \sqrt{\frac{m}{2\pi\hbar\ \delta t}} \right]^{3\mathcal{N}/2} \int \mathrm{d}^3r_1 \int \mathrm{d}^3r_2 .. \int \mathrm{d}^3r_k .. \int \mathrm{d}^3r_{\mathcal{N}-1} \tag{30}$$

[3]Remember that this complement does not pretend to be mathematically rigorous. This would entail a more careful study of the classical as well as quantum expressions, and of the effects of the simultaneous limits δt goes to zero and $\mathcal{N}$ (number of terms in the products) goes to infinity, keeping in mind that these approximations are done on functions that are arguments of exponentials.

2-b. Generalization: several particles interacting via a potential

The previous considerations can be directly generalized to a system with several particles interacting via a potential depending on their positions. Since the system Hamiltonian is, as above, the sum of two non-commuting terms, we again use the approximate expression (18) for the evolution operator. But rather than inserting the closure relations for a single position, one must now use a basis involving the positions of all the particles: each integral over $\mathrm{d}^3 r$ is then demultiplied into as many integrals as there are particles in the system.

We will not include here the case where the particles are charged and subjected to a magnetic field, which would include terms such as $q\mathbf{v}\cdot\mathbf{A}(\mathbf{r})$, where $\mathbf{A}(\mathbf{r})$ is the vector potential. Even though it is an interesting case, in particular for its relation to gauge invariance, it will not be considered here for the sake of brevity. The interested reader should consult the references given in the introduction.

3. Discussion; a new quantization rule

The path integral approach is particularly fit for developing an analogy with classical optics. In addition, it allows building new quantization rules. These two points will be discussed successively.

3-a. Analogy with classical optics

Relations (28) and (29) allow making a link between two a priori unrelated quantities: the classical mechanics paths with their associated actions, and the quantum propagator. Knowing the wave function $\Psi(\mathbf{r},t)$ at a given time t, it is this quantum propagator that enables computing that wave function at a later time t' as follows:

$$\begin{aligned}\Psi(\mathbf{r}',t') &= \langle \mathbf{r}' \,|\Psi(t')\rangle = \langle \mathbf{r}'|\, U(t',t)\,|\Psi(t)\rangle \\ &= \int \mathrm{d}^3 r \;\langle \mathbf{r}'|\, U(t',t)\,|\mathbf{r}\rangle\,\langle \mathbf{r}\,|\Psi(t)\rangle \end{aligned} \tag{31}$$

Taking into account definition (1) of the propagator, the above expression can be rewritten as:

$$\Psi(\mathbf{r}',t') = \int \mathrm{d}^3 r\; G(\mathbf{r}',t';\mathbf{r},t)\;\Psi(\mathbf{r},t) \tag{32}$$

The propagator is thus the kernel of the integral equation expressing the temporal propagation of the wave function. Now, equality (28) shows that this propagator is equal to a sum of exponentials of the actions corresponding to all the classical paths. Therefore, whereas in classical mechanics a single path (or in certain cases a finite number of paths) is selected by the stationarity condition of the action, in quantum mechanics all the paths come into play to determine the propagation amplitude, each one with its particular phase. In a manner of speaking, one can say that, in quantum mechanics, the particle goes through all the possible intermediate positions $\mathbf{r}_k$ and hence follows all the possible histories between the two end points.

It is worth noting that all these histories (even highly unlikely histories involving totally arbitrary positions) contribute with the same amplitude. On the other hand, the

phases associated with the histories are different from each other and allow understanding how the different histories come into play. This situation can be analyzed in terms of stationary phase conditions. It is easy to understand that in the summation, the histories corresponding to a stationary action will play a particular role since all the neighboring histories will add their contribution in a coherent way. On the other hand, in the vicinity of histories for which the action varies rapidly, the phase oscillates quickly and the corresponding contributions will cancel out through destructive interference. Consequently, classical histories play a privileged role, which becomes more prominent as the phase oscillations become more rapid. In the limit where $\hbar \to 0$, these oscillations become infinitely rapid and only the classical histories prevail.

Finally, this situation reminds us of classical optics and the Huyghens-Fresnel principle, where a light wave is computed as the sum of waves radiated from each point of an intermediate surface, taking into account the phases linked to the propagation along each path. In the geometrical approximation, where the wavelength tends towards zero, the trajectories of the light rays correspond to paths having a stationary phase, i.e. which does not change for infinitely close paths. Geometrical optics is the analog of classical mechanics, whereas Huyghens wave optics is the analog of quantum mechanics. The Feynman integral path is therefore a useful tool to study the link between classical and quantum mechanics, and in particular the semiclassical limit of quantum mechanics (WKB approximation, etc.).

Comment:

The preceding analogy is well founded for a single particle. For a system with N particles, the histories no longer propagate in the ordinary three-dimensional space, but in a $3N$-dimensional configuration space. The analogy with optics described above is no longer as adequate, since in classical optics, electromagnetic waves propagate in ordinary $3D$ space.

3-b. A new quantization rule

At this stage, we can invert the approach. Until now, starting from the rules of Hamiltonian quantum mechanics, we deduced an equivalent expression for the propagator, i.e. another way for finding the solutions of Schrödinger's equation. It is also possible to consider this equivalent expression as the starting point and postulate that the propagator is defined ab initio by a sum over all the classical paths Γ, each contributing an exponential $\exp(iS_\Gamma/\hbar)$. This yields another method for the quantization of a physical system, which offers several advantages. First of all, as we just saw, it highlights the relation of quantum mechanics with classical mechanics, where only a single classical path exists (or sometimes a finite number of paths), as opposed to an infinite number of possible paths in quantum mechanics. Furthermore, it is remarkable that the probability amplitudes thus computed only depend on classical functions (involving only numbers and not operators), the only explicit quantum component being the presence of $\hbar$ in the denominator of the phase. We shall see in § 4 how the concept of operators can be introduced in this approach. In addition, the expressions involving path integrals are symmetric with respect to time and space, since both type of coordinates are integrated

in a similar way[4]. Reasoning directly in space-time makes it easier to include Einstein's relativity, since one can replace the time differential δt by a proper time differential $\delta\tau$. If now the Lagrangian $\mathcal{L}$ is a space-time scalar, so is the action, and the theory acquires relativistic invariance. Finally, Feynman's quantization method only requires the existence of a Lagrangian, with its associated variational principle. Now all the physical systems that have a Lagrangian do not necessarily have the conjugate variables permitting the definition of a Hamiltonian. For such systems, the Feynman path integral method is powerful and this is why it is so important in quantum field theory.

4. Operators

Feynman paths also permit computing the matrix elements of operators in the Heisenberg picture, where they are time-dependent. We shall mostly consider the simplest case where the operators are functions of the position operator $\mathbf{R}$.

4-a. One single operator

Let us insert any operator A "in the middle" of the evolution operator (1) by splitting the time interval $[t, t']$ into two adjacent intervals $[t, t'']$ and $[t'', t']$, with $t < t'' < t'$. We get the expression:

$$\langle \mathbf{r}' | \, U(t', t'') \;\; A \; U(t'', t) \, | \mathbf{r} \rangle = \langle \theta(\mathbf{r}', t'', t') | \, A \, | \theta(\mathbf{r}, t'', t) \rangle \tag{33}$$

with:

$$\begin{aligned} |\theta(\mathbf{r}, t'', t)\rangle &= U(t'', t) \, |\mathbf{r}\rangle \\ |\theta(\mathbf{r}', t'', t')\rangle &= U^{\dagger}(t', t'') \, |\mathbf{r}'\rangle = U(t'', t') \, |\mathbf{r}'\rangle \end{aligned} \tag{34}$$

In this matrix element of A, the ket $|\theta(\mathbf{r}, t'', t)\rangle$ is obtained by the evolution until the time t'' of a state localized at $\mathbf{r}$ at time t; the bra $\langle \theta(\mathbf{r}', t'', t')|$ corresponds to the ket $|\theta(\mathbf{r}', t'', t')\rangle$ which, as it evolves between t'' and t', becomes a ket localized at $\mathbf{r}'$:

$$U(t', t'') \, |\theta(\mathbf{r}', t'', t')\rangle = U(t', t'') \, U(t'', t') \, |\mathbf{r}'\rangle = |\mathbf{r}'\rangle \tag{35}$$

α. Operator function of the position

We now assume operator A is a function $A(\mathbf{R})$ of the particle's position operator. Inserting a closure relation on the positions, we can write the left-hand side of (33) as:

$$\langle \mathbf{r}' | \, U(t', t'') \;\; A(\mathbf{R}) \;\; U(t'', t) \, | \mathbf{r} \rangle = \int \mathrm{d}^3 r'' \, \langle \mathbf{r}' | \, U(t', t'') \, | \mathbf{r}'' \rangle \;\; A(\mathbf{r}'') \;\; \langle \mathbf{r}'' | \, U(t'', t) \, | \mathbf{r} \rangle \tag{36}$$

[4]The sum over all the paths introduces an integral over all the positions $\boldsymbol{r}_k$ in Figure 2, hence differentials of the three space coordinates; in addition, the integral over the times introduces a differential $\mathrm{d}t = t_k - t_{k-1}$. The product of three space differentials by a time differential thus allows one to introduce a differential of space-time volume.

Using the general relation (28), we can then write the two propagators appearing under the integral as the sum over all the paths Γ_1 or Γ_2:

$$\begin{aligned}\langle \mathbf{r}''| U(t'',t) |\mathbf{r}\rangle &= \sum_{\text{paths } \Gamma_1} \exp\left[\frac{iS_{\Gamma_1}}{\hbar}\right] \\ \langle \mathbf{r}'| U(t',t'') |\mathbf{r}''\rangle &= \sum_{\text{paths } \Gamma_2} \exp\left[\frac{iS_{\Gamma_2}}{\hbar}\right]\end{aligned} \tag{37}$$

where Γ_1 is a path linking the initial position $\mathbf{r}$ at time t to the intermediate position $\mathbf{r}''$ at time t'', and Γ_2 the path linking thereafter the intermediate position $\mathbf{r}''$ at time t'' to the final position $\mathbf{r}'$ at time t'. If t'' coincides with the intermediate time t_k, relation (36) becomes:

$$\begin{aligned}&\langle \mathbf{r}'| U(t',t_k)\ A(\mathbf{R})\ U(t_k,t) |\mathbf{r}\rangle \\ &\qquad = \int \mathrm{d}^3 r_k \sum_{\text{paths } \Gamma_1 \text{ and } \Gamma_2} \exp\left[\frac{iS_{\Gamma_2}}{\hbar}\right] A(\mathbf{r}_k) \exp\left[\frac{iS_{\Gamma_1}}{\hbar}\right]\end{aligned} \tag{38}$$

Now the product of the two exponentials yields a single exponential $\exp[iS_\Gamma/\hbar]$ associated with the action S_Γ of a path Γ consisting of the two paths Γ_1 and Γ_2 joined together end to end at $\mathbf{r}_k$. The sum over $\mathrm{d}^3 r_k$ reconstitutes the ensemble of all paths going from the initial position $\mathbf{r}$ at time t to the final position $\mathbf{r}'$ at time t', the only difference being that now each exponential $\exp[iS_\Gamma/\hbar]$ is multiplied by the value $A(\mathbf{r}_k)$ taken by the function A at the intermediate point at time t_k.

We finally obtain:

$$\langle \mathbf{r}'| U(t',t'')\ A(\mathbf{R})\ U(t'',t) |\mathbf{r}\rangle = \sum_{\text{paths } \Gamma} A(\mathbf{r}'')\ \exp\left[\frac{iS_\Gamma}{\hbar}\right] \tag{39}$$

where $A(\mathbf{r}'')$ is the value of $A(\mathbf{r})$ at position $\mathbf{r}''$ which the path Γ traverses at time t''. The matrix elements of the operator in the Heisenberg picture (special case $t' = t$) are thus given by the same summation over the histories as for the propagator, the only difference being that the contribution of each path is now multiplied by the value taken by the operator at the position $\mathbf{r}''$ at the intermediate time t''.

As before, we can now invert the approach and consider relation (39) as the definition of an operator in the framework of the Feynman path quantization method. Here again, it is remarkable that this relation involves only classical functions, without any operator.

β. Velocity operator; canonical commutation relations

In order to define an operator $\mathbf{W}$ associated with the particle's velocity at time t'' (we use the notation $\mathbf{W}$ to avoid any confusion with the potential V), and taking (22) into account, a natural extension of (38) leads to setting:

$$\begin{aligned}&\langle \mathbf{r}'| U(t',t'')\ \mathbf{W}\ U(t'',t) |\mathbf{r}\rangle \\ &\qquad = \int \mathrm{d}^3 r_k \sum_{\text{paths } \Gamma_2} \exp\left[\frac{iS_{\Gamma_2}}{\hbar}\right] \sum_{\text{paths } \Gamma_1} \left(\frac{\mathbf{r}_k - \mathbf{r}_{k-1}}{\delta t}\right) \exp\left[\frac{iS_{\Gamma_1}}{\hbar}\right]\end{aligned} \tag{40}$$

where the paths Γ_1 are all those going from the initial position $\mathbf{r}$ to the intermediate position $\mathbf{r}_k$ (the preceding intermediate position $\mathbf{r}_{k-1}$ does depend on the path), whereas the paths Γ_2 are all those going from $\mathbf{r}_k$ to the final position $\mathbf{r}'$. Introducing in the middle of the left-hand side a closure relation on the kets $\{|\mathbf{r}_k\rangle\}$, this relation becomes:

$$\int \mathrm{d}^3 r_k \;\; \langle \mathbf{r}' | \, U\left(t', t''\right) | \mathbf{r}_k \rangle \;\; \langle \mathbf{r}_k | \, \mathbf{W} \; U\left(t'', t\right) | \mathbf{r} \rangle$$
$$= \int \mathrm{d}^3 r_k \;\; \langle \mathbf{r}' | \, U\left(t', t''\right) | \mathbf{r}_k \rangle \;\; \Psi_W\left(\mathbf{r}_k, t''\right) \tag{41}$$

with:

$$\Psi_W\left(\mathbf{r}_k, t''\right) = \langle \mathbf{r}_k | \, \mathbf{W} \; U\left(t'', t\right) | \mathbf{r} \rangle = \sum_{\text{paths } \Gamma_1} \left(\frac{\mathbf{r}_k - \mathbf{r}_{k-1}}{\delta t} \right) \exp\left[\frac{i S_{\Gamma_1}}{\hbar} \right] \tag{42}$$

This wave function is the result of the action of operator $\mathbf{W}$ on the wave function at time t'', equal to:

$$\Psi\left(\mathbf{r}_k, t''\right) = \langle \mathbf{r}_k | \, U\left(t'', t\right) | \mathbf{r} \rangle \tag{43}$$

Let us compare the sum over the paths Γ_1 in relation (42) and in the relation (28) used to build the propagator between the times t and t''. In these two equalities, the actions are given by the summations (25) over the intermediate positions. Concerning the contributions of the paths between the initial time and the intermediate time t_{k-1} (the last time over which the summation runs), the two sums in (42) and (28) are identical. Actually, their only difference concerns the very last time interval (between t_{k-1} and $t_k = t''$), which in (42) is multiplied by the factor $\left(\mathbf{r}_k - \mathbf{r}_{k-1}\right)/\delta t$. Now this multiplicative factor can also be found by taking the derivative of (28) with respect to $\mathbf{r}_k$ since, using (25), we have:

$$\boldsymbol{\nabla}_{\mathbf{r}_k} \langle \mathbf{r}_k | \, U\left(t'', t\right) | \mathbf{r} \rangle = \frac{i \delta t}{\hbar} \sum_{\text{paths } \Gamma_1} \left[m \left(\frac{\mathbf{r}_k - \mathbf{r}_{k-1}}{\delta t^2} \right) + \frac{1}{2} \boldsymbol{\nabla}_{\mathbf{r}_k} V \right] \exp\left[\frac{i S_{\Gamma_1}}{\hbar} \right] \tag{44}$$

As $\mathbf{r}_k$ is a final fixed point, the term in $\boldsymbol{\nabla}_{\mathbf{r}_k} V$ on the right-hand side can be taken out of the summation over the paths. It yields a contribution in $\delta t \times \boldsymbol{\nabla}_{\mathbf{r}_k} V$ which goes to zero in the limit $\delta t \to 0$. We are left with:

$$\boldsymbol{\nabla}_{\mathbf{r}_k} \langle \mathbf{r}_k | \, U\left(t'', t\right) | \mathbf{r} \rangle = \frac{i m}{\hbar} \sum_{\text{paths } \Gamma_1} \left(\frac{\mathbf{r}_k - \mathbf{r}_{k-1}}{\delta t} \right) \exp\left[\frac{i S_{\Gamma_1}}{\hbar} \right] \tag{45}$$

so that relation (42) becomes:

$$\Psi_W\left(\mathbf{r}_k, t''\right) = \frac{\hbar}{i m} \boldsymbol{\nabla}_{\mathbf{r}_k} \langle \mathbf{r}_k | \, U\left(t'', t\right) | \mathbf{r} \rangle = \frac{\hbar}{i m} \boldsymbol{\nabla}_{\mathbf{r}_k} \Psi\left(\mathbf{r}_k, t''\right) \tag{46}$$

This means that the action of the velocity operator $\mathbf{W}$ is simply proportional to a derivative[5] with respect to the position $\mathbf{r}_k$, which is the variable of the wave function at the

[5]The demonstration pertains to a wave function at the instant t'' that is issued from a wave function localized at point $\mathbf{r}$ at time t. By linear superposition, it can be generalized to any wave function at time t, hence confirming the same derivation property.

instant t''. In other words, if $\mathbf{P} = m\mathbf{W}$ is the particle's momentum operator, its action on the wave function is $(\hbar/i)$ times the gradient with respect to the position. We have established a basic result of the usual quantum mechanics, starting from operators introduced in the path integral approach.

The canonical commutation relations between $\mathbf{R}$ and $\mathbf{P}$ are easily derived, since:

$$\frac{\partial}{\partial x}\left[x\Psi\left(x\right)\right] = x\frac{\partial}{\partial x}\left[\Psi\left(x\right)\right] + \Psi\left(x\right) \tag{47}$$

These commutation relations can also be considered as consequences of the path quantization rules.

4-b. Several operators

Feynman postulates also permit introducing products of several operators, acting at the same instant or at different times.

α. Several operators at different times

The previous argument can be generalized to several operators $A\left(\mathbf{R}\right)$, $B\left(\mathbf{R}\right)$, etc. acting at intermediate times t'', t''', etc. As before, we can split the evolution operator into several parts corresponding to the successive time intervals, and insert position closure relations at the intermediate times. Each operator introduces a factor dependent on the corresponding intermediate position, and the time propagation is a sum over histories between successive time intervals. For instance, for two operators, the same reasoning followed above leads to:

$$\begin{aligned}\left\langle \mathbf{r}'\right| U\left(t',t'''\right)\, B\left(\mathbf{R}\right)\, U\left(t''',t''\right)\, A\left(\mathbf{R}\right)\, U\left(t'',t\right)\left|\mathbf{r}\right\rangle \\ = \sum_{\text{paths }\Gamma} B\left(\mathbf{r}'''\right) A\left(\mathbf{r}''\right) \exp\left[\frac{iS_\Gamma}{\hbar}\right]\end{aligned} \tag{48}$$

where $A\left(\mathbf{r}''\right)$ is the value of A for the position $\mathbf{r}''$ crossed by the path Γ at time t'', $B\left(\mathbf{r}'''\right)$ the value of B for the position $\mathbf{r}'''$ crossed by the path Γ at a later time t'''. The result is easily generalized to any number of operators. Note the order in which the operators are arranged in the matrix element on the left-hand side: it corresponds to the order in which the times t'', t''', etc. are arranged in the classical histories used to calculate the actions. The quantum operators are automatically arranged in decreasing times from left to right, even if $A\left(\mathbf{r}''\right)$ and $B\left(\mathbf{r}'''\right)$ are numbers that commute in the right hand side of relation (48).

β. Position and velocity operators, symmetrization

Imagine now we want to introduce, for example, the operator corresponding to the product $\mathbf{R}\cdot\mathbf{P}$ of the position and momentum. We will then proceed as in (41) and (42), however with an added precaution: should we multiply $\left(\mathbf{r}_k - \mathbf{r}_{k-1}\right)/\delta t$ by $\mathbf{r}_k$ or by $\mathbf{r}_{k-1}$ (the order does not matter, since we are dealing with numbers). For the sake of symmetry, we multiply by half their sum, so that in (42) $\left(\mathbf{r}_k - \mathbf{r}_{k-1}\right)/\delta t$ is replaced by:

$$\frac{1}{\delta t}\frac{\mathbf{r}_k + \mathbf{r}_{k-1}}{2}\cdot\left(\mathbf{r}_k - \mathbf{r}_{k-1}\right) = \frac{1}{2\delta t}\left[\mathbf{r}_k\cdot\left(\mathbf{r}_k - \mathbf{r}_{k-1}\right) + \left(\mathbf{r}_k - \mathbf{r}_{k-1}\right)\cdot\mathbf{r}_{k-1}\right] \tag{49}$$

On the right-hand side, we wrote the two terms so that the times are always decreasing (or stationary) from left to right. Because of the order of the k indices, the first term on the right-hand side introduces the operator $\mathbf{R} \cdot \mathbf{W}$, whereas the second introduces $\mathbf{W} \cdot \mathbf{R}$, i.e. the same operators but in the inverse order. This is an example of how the path integral method leads quite naturally to a symmetrization of the operator order, which automatically ensures the hermiticity of their product.

Conclusion

To be able to use two complementary approaches, the Hamiltonian method and the path integral method, is often quite valuable in the study of numerous physical problems. As an example, path integrals play a fundamental role in field theory. They are for a large part at the base of the implementation of symmetry groups (Abelian or non-commutative) in this theory, which allows building a theory for elementary particles and their interactions. There are, however, other cases where the path integral formalism is very useful, as for example, in the computations of quantum interference with cold atoms. Conceptually, the path integral approach can shed new light on the relations between quantum mechanics and classical mechanics, as well as classical optics as we saw in § 3-a.

In this appendix, we only used path integrals as a method to compute the time propagator of a quantum physical system, which involves imaginary exponentials of the Hamiltonian. Path integrals can also be used in quantum statistical mechanics (Appendix VI), and involve real exponentials of the Hamiltonian (multiplied by the inverse of the temperature). It is the basic tool for many numerical calculations; the interested reader can consult Zinn-Justin's book [92], or reference [94] where, in particular, the PIMC (Path Integral Quantum Monte Carlo) methods are described.

Appendix V

Lagrange multipliers

1 Function of two variables 2281
2 Function of N variables 2283

When a function F depends on non-independent variables (i.e. which are related by constraints), its extrema (maxima or minima) can be found by the Lagrange multiplier method. A brief summary of this method is proposed in this appendix. The first part concerns functions of two variables, and the second part will generalize the concept to any number of variables.

1. Function of two variables

Consider first a real function $F(x_1, x_2)$ of two independent variables x_1 and x_2. We assume the fonction F to be regular, continuous, differentiable with continuous derivatives. The extrema of F correspond to values of the variables for which the two partial derivatives are zero:

$$\frac{\partial F(x_1, x_2)}{\partial x_1} = 0 \qquad ; \qquad \frac{\partial F(x_1, x_2)}{\partial x_2} = 0 \tag{1}$$

These two relations amount to stating that the gradient of F must be zero:

$$\overrightarrow{\nabla F} = 0 \tag{2}$$

Two equations with two unknowns x_1 and x_2 generally admit a finite number of solutions (pairs of values for x_1 and x_2); this number can even be zero if the function F does not present any extrema.

Let us now look for the extrema of F when the variables are no longer independent, but must obey a constraint:

$$E(x_1, x_2) = C \tag{3}$$

where C is a constant and $E(x_1, x_2)$ a regular function (continuous, differentiable, etc.). When this constraint is satisfied, the point M with coordinates x_1 and x_2 is forced to follow a curve in the plane (solid line in Figure 1). Imagine we place the point close to an arbitrary point M of the curve, and move it by varying slightly its coordinates by $\mathrm{d}x_1$ and $\mathrm{d}x_2$. For E to remain constant, $\mathrm{d}x_1$ and $\mathrm{d}x_2$ must necessarily obey:

$$\mathrm{d}E = \frac{\partial E(x_1, x_2)}{\partial x_1}\,\mathrm{d}x_1 + \frac{\partial E(x_1, x_2)}{\partial x_2}\,\mathrm{d}x_2 = \overrightarrow{\nabla E} \cdot \overrightarrow{\mathrm{d}M} = 0 \tag{4}$$

The point M therefore necessarily moves along the tangent to the curve, i.e. perpendicularly to its gradient $\overrightarrow{\nabla E}$, as shown in Figure 1. As for the variation of F, it is given by:

$$\mathrm{d}F = \overrightarrow{\nabla F} \cdot \overrightarrow{\mathrm{d}M} \tag{5}$$

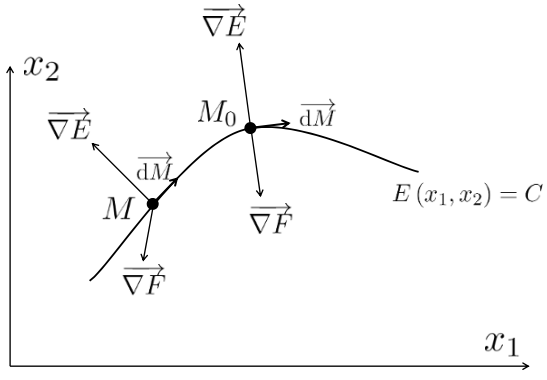


Figure 1: When the constraint $E(x_1, x_2) = C$ is satisfied, the point M, with coordinates x_1 and x_2, is forced to move along a curve in the plane, shown as a solid line. The tangent to this curve is perpendicular to the gradient $\overrightarrow{\nabla E}$ of the function E, meaning that any small displacement of point M along the curve must be perpendicular to this gradient. When the displacement starts from an arbitrary point M, the vectors $\overrightarrow{\nabla E}$ and $\overrightarrow{\nabla F}$ are not parallel, and the function F varies to first order in $dF = \overrightarrow{\nabla F} \cdot \overrightarrow{dM}$. However, if the variation starts from a point, such as M_0, where the two gradients are parallel, the function F is stationary. Geometrically, this parallelism means that the solid line is tangent to a contour line of the surface representing the function $F(x_1, x_2)$.

As in general the vectors $\overrightarrow{\nabla E}$ and $\overrightarrow{\nabla F}$ are not parallel, this scalar product is not zero. The function F thus varies to first order in $\overrightarrow{\mathrm{d}M}$, meaning it is not stationary at that point.

If, however, we start from a point M_0 on the curve where the two gradients are parallel (or antiparallel), condition (4) means that the variation (5) is zero, and stationarity is attained. In such a case, M moves (at constant E) along a curve that is tangent at M_0 to a contour line of the surface representing the function F. Geometrically, it is easy to understand that a displacement along a contour line keeps F constant to first order. Algebraically, imposing the gradients to be parallel amounts to writing that there exists a constant λ, called "Lagrange multiplier", such that:

$$\overrightarrow{\nabla F} = \lambda \, \overrightarrow{\nabla E} \tag{6}$$

which is equivalent to saying that the differential of the function $F - \lambda E$ is zero:

$$\begin{aligned} \mathrm{d}(F - \lambda E) &= \overrightarrow{\nabla}(F - \lambda E) \cdot \overrightarrow{\mathrm{d}M} \\ &= \frac{\partial [F - \lambda E]}{\partial x_1} \mathrm{d}x_1 + \frac{\partial [F - \lambda E]}{\partial x_2} \mathrm{d}x_2 \\ &= 0 \end{aligned} \tag{7}$$

This means that one must simply replace the function F by the fonction $F - \lambda E$ with an arbitrary Lagrange multiplier λ to obtain the stationarity of F when its variables obey the constraint (3).

When λ is fixed, we get as before two equations with two unknowns, so that the variables x_1 and x_2 are determined. Inserting them into (3) yields a value for C, which is thus fixed. If, however, the Lagrange multiplier λ is allowed to vary, the constant C becomes a function of λ, and can be adjusted by changing λ. As an example, when studying the canonical equilibrium (Appendix VI, § 1-b), one maximizes the value of the entropy S (which plays the role of the function F) while keeping the average energy value E constant. A Lagrange multiplier β is then introduced to impose the stationarity of $S - \beta E$; changing β allows controlling the value of E.

2. Function of N variables

We now consider a function $F(x_1, x_2, ..., x_N)$ of N supposedly independent variables x_1, x_2,... ,x_N. The extrema of F are obtained by annulling the N components of the gradient of F (each component being the partial derivative of F with respect to one of the variables):

$$\overrightarrow{\nabla F} = 0 \tag{8}$$

We get N equations to determine N unknown variables, yielding a finite number of extrema.

Imagine now that the N variables are no longer independent, but linked by $P < N$ conditions:

$$E_p(x_1, x_2, ..., x_N) = C_p \qquad \text{with } p = 1, 2, ..., P \tag{9}$$

Consider a point M in an N-dimensional space, with coordinates $x_1, x_2, ..., x_N$. If these N coordinates satisfy the P conditions (9), their infinitesimal variations obey the P relations:

$$\overrightarrow{\nabla E_p} \cdot \overrightarrow{\mathrm{d}M} = 0 \qquad \text{with } p = 1, 2, ..., P \tag{10}$$

For all the functions E_p to remain constant, the displacement $\overrightarrow{\mathrm{d}M}$ of point M in the N-dimensional space must be orthogonal to all the gradients $\overrightarrow{\nabla E_p}$. Two cases are then possible:

(i) either the gradient $\overrightarrow{\nabla F}$ belongs to the sub-space generated by the $\overrightarrow{\nabla E_p}$, in which case the orthogonality condition (10) implies that $\overrightarrow{\mathrm{d}M}$ is also orthogonal to $\overrightarrow{\nabla F}$. Consequently the variation $\mathrm{d}F = \overrightarrow{\nabla E} \cdot \overrightarrow{\mathrm{d}M}$ is zero and the stationarity is ensured.

(ii) or the gradient $\overrightarrow{\nabla F}$ is not contained in that subspace, and it possesses a non-zero component $\overrightarrow{\nabla F}_\perp$ orthogonal to that subspace. One can then choose $\overrightarrow{\mathrm{d}M}$ parallel to $\overrightarrow{\nabla F}_\perp$ and obtain a first order variation of F, while satisfying the P constraints.

In conclusion, the stationarity of F is equivalent to the condition that the gradient $\overrightarrow{\nabla F}$ be contained in the subspace generated by the $\overrightarrow{\nabla E_p}$. This amounts to stating that there exist Lagrange multipliers λ_p (with $p = 1, 2, ..., P$) such that:

$$\overrightarrow{\nabla F} = \lambda_1 \overrightarrow{\nabla E_1} + \lambda_2 \overrightarrow{\nabla E_2} + ... + \lambda_P \overrightarrow{\nabla E_P} \tag{11}$$

In an equivalent way, the stationarity condition can be obtained by annulling the differential:

$$\mathrm{d}\left(F - \lambda_1 E_1 - \lambda_2 E_2 - ... - \lambda_P E_P\right) = 0 \tag{12}$$

and then treating the variables $x_1, x_2, ..., x_N$ as if they were independent.

When the Lagrange multipliers λ_p are fixed, each component of relation (11) yields an equation, so that we have as many equations as variables $x_1, x_2, ..., x_N$. One therefore obtains for the function F a finite number of extrema linked by the constraints, yielding fixed values for the functions E_1, E_2, ..., E_P. A variation in the Lagrange multipliers will change the values of these functions, which can therefore be adjusted to a value that has been chosen in advance.

Appendix VI

Brief review of Quantum Statistical Mechanics

1 Statistical ensembles 2285
1-a Microcanonical ensemble 2285
1-b Canonical ensemble 2289
1-c Grand canonical ensemble 2291
2 Intensive or extensive physical quantities 2292
2-a Microcanonical ensemble 2293
2-b Canonical ensemble 2294
2-c Grand canonical ensemble 2295
2-d Other ensembles 2295

In quantum mechanics, as in classical mechanics, it is not possible to describe a system having a very large number of degrees of freedom (for example a system containing a number of particles that is of the order of the Avogadro number) with highest precision. Such a description would in particular include the value of many quantities, for instance many particle correlations, which fluctuate rapidly and are not necessarily of interest. A less detailed and more probabilistic description must be used, in which the state of the system in known only statistically. The system occupies one on a series of possible states, with a certain probability. One then says that the system is described by a "statistical ensemble". The use of a density operator (Complement E_{III}) to describe the physical system is particularly convenient in this case.

We do not attempt in this appendix to give a general introduction to statistical mechanics and its postulates. We simply summarize a number of quantum statistical mechanics results used in several complements. For example, most of the complements of Chapter XV, as well as B_{XVII} and D_{XVII}, use the concept of chemical potential μ or of "grand potential" Φ; their interpretation in the framework of the different statistical ensembles will be given in this appendix.

1. Statistical ensembles

Several "statistical ensembles" are commonly used to describe physical systems at equilibrium. We shall focus here on the three main ones: the microcanonical, the canonical and the grand canonical ensembles. The first of these ensembles provides the general setting for introducing the two others.

1-a. Microcanonical ensemble

Consider a physical system containing N particles in a box of volume $\mathcal{V}$. The energy E' of the system lies within an interval:

$$E - \Delta E/2 \leq E' \leq E + \Delta E/2 \tag{1}$$

with $\Delta E \ll E$. The system is isolated from its surroundings preventing any exchange of particles or energy. We note $|E', q\rangle$ the eigenstates of its Hamiltonian $\widehat{H}$, where q is an index reflecting the possible degeneracy of each eigenvalue E' of this Hamiltonian.

α. Density operator, entropy

The system is supposed to have the same probability of being in any state whose energy falls within the interval (1); no state is favored over any other. The microcanonical density operator of the system at equilibrium is then:

$$\rho_{\text{eq}} = \frac{1}{Z} P_{\Delta E}(E) \tag{2}$$

where $P_{\Delta E}(E)$ is the projector onto the subspace containing all the accessible states:

$$P_{\Delta E}(E) = \sum_{E'=E-\frac{\Delta E}{2}}^{E+\frac{\Delta E}{2}} \sum_{q} |E', q\rangle \langle E', q| \tag{3}$$

and where Z is the microcanonical partition function defined as:

$$Z = \text{Tr}\{P_{\Delta E}(E)\} \tag{4}$$

What relation (2) means is that the occupation probabilities of the states $|E', q\rangle$ are all equal to $1/Z$. Relation (3) shows that each of the projectors onto a state $|E', q\rangle$ contributes one unit to the trace of relation (4); the partition function Z is simply the number of terms in the summation (3), i.e. the number of levels in the energy interval (1). If $D(E)$ is the density of states, we can write:

$$Z = D(E) \times \Delta E \tag{5}$$

As in § 1 of Complement A_{XXI}, we define the entropy S as:

$$S = -k_B \text{ Tr}\{\rho_{\text{eq}} \ln \rho_{\text{eq}}\} \tag{6}$$

where k_B is the Boltzmann constant. The $|E', q\rangle$ are eigenvectors of ρ_{eq}, with an eigenvalue equal to $1/Z$ if E' belongs to the interval $[E - \frac{\Delta E}{2}, E + \frac{\Delta E}{2}]$, and equal to zero otherwise. If E' belongs to the interval, we have:

$$\rho_{\text{eq}} \ln \rho_{\text{eq}} |E', q\rangle = -\frac{\ln Z}{Z} |E', q\rangle \tag{7}$$

If E' does not belong to the interval, since $\lim_{x \to 0} x \ln x = 0$, we get:

$$\rho_{\text{eq}} \ln \rho_{\text{eq}} |E', q\rangle = 0 \tag{8}$$

We now multiply the two previous relations by the bra $\langle E', q|$ and sum over E' and q to get a trace. Only the bras whose energy falls within the interval will yield a non-zero contribution. As there are Z of them, we obtain:

$$\text{Tr}\{\rho_{\text{eq}} \ln \rho_{\text{eq}}\} = -\ln Z \tag{9}$$

The equilibrium value of the entropy is therefore:

$$S = k_B \ln Z \tag{10}$$

β. Temperature, chemical potential

Suppose we now change the equilibrium energy E by an infinitesimal amount $\mathrm{d}E$, keeping the volume $\mathcal{V}$ and the particle number N constant. Since no work is exchanged with the outside (the external walls are fixed), this amounts solely to a heat change:

$$\mathrm{d}Q = \mathrm{d}E = T\mathrm{d}S \tag{11}$$

where we have used the usual thermodynamic definition of the entropy $\mathrm{d}S = \mathrm{d}Q/T$. In the microcanonical ensemble, the temperature is thus defined as:

$$\frac{1}{T} = \left.\frac{\partial S}{\partial E}\right|_{N,\mathcal{V}} \tag{12}$$

where we have used the partial derivative notation to emphasize that the changes are made keeping both N and $\mathcal{V}$ constant.

Let us now change the particle number N, keeping the volume and the energy E constant. We then define the "chemical potential" μ (which has the dimension of an energy) as:

$$\mu = -T \left.\frac{\partial S}{\partial N}\right|_{E,\mathcal{V}} \tag{13}$$

At fixed temperature, the faster the entropy (which depends on the number of accessible levels in the energy band ΔE) grows with the particle number N, the larger the absolute value of μ. The chemical potential plays an essential role in the grand canonical equilibrium as we shall see (§ 1-c). The third partial derivative of S (with respect to the volume) will be determined in § 2-a.

Comment:

Let us insert (5) in relation (10), but first multiplying $D(E)$ by $k_B T$ and dividing ΔE by this same quantity (this has the advantage of providing dimensionless arguments for the logarithmic functions). This yields:

$$\frac{S}{k_B} = \ln\left[D\left(E\right) k_B T\right] + \ln\left[\frac{\Delta E}{k_B T}\right] \tag{14}$$

In a macroscopic system, the particle number is very large, of the order of the Avogadro number. Let us see then what happens when the particle number N goes to infinity. We assume that the energy E as well as the volume are proportional to N (thermodynamic limit). We then expect the entropy to also be proportional to N. This linear variation of the entropy cannot come from the second term in (14): even if the energy interval ΔE is proportional to N, it will only yield a much slower logarithmic variation. Most of the variation of S actually comes from the first term of (14), and from the fact that the density of states increases with N in an exponential way: as the exponent of $D(E)$ contains N, this variation is phenomenally rapid. In the limit of large systems, the first term in (14) largely dominates the second. This is why it is often said that the entropy characterizes the density of states of a physical system (or more precisely the number of its quantum energy levels in a microscopic energy interval, chosen here to be equal to $k_B T$).

γ. Entropy maximization

We now choose for ρ an arbitrary Hermitian density operator, with positive or zero eigenvalues whose sum is equal to 1. We denote $|\theta_n\rangle$ its eigenvectors and p_n its eigenvalues ($0 \leq p_n \leq 1$) which obey:

$$\sum_n p_n = 1 \tag{15}$$

We assume that ρ is restricted to the energy band (1): all the $|\theta_n\rangle$ for which $p_n \neq 0$ are arbitrary linear combinations of the eigenvectors $|E, q\rangle$ obeying (1).

An entropy can be associated with ρ:

$$S_\rho = -k_B \ \mathrm{Tr}\left\{\rho \ \ln\rho\right\} \tag{16}$$

where k_B is the Boltzmann constant. We are going to show that among all possible ρ operators, the equilibrium one, ρ_{eq}, maximizes this entropy. We can write:

$$\rho = \sum_n p_n \ \ |\theta_n\rangle \langle\theta_n| \qquad\qquad \ln\rho = \sum_n \ln p_n \ \ |\theta_n\rangle \langle\theta_n| \tag{17}$$

and therefore:

$$S_\rho = -k_B \ \mathrm{Tr}\left\{\sum_n p_n \ \ln p_n \ \ |\theta_n\rangle \langle\theta_n|\right\} = -k_B \sum_n p_n \ \ln p_n \tag{18}$$

Any variation of the p_n results in a variation of S written as:

$$\mathrm{d}S = -k_B \sum_n \left[1 + \ln p_n\right] \ \mathrm{d}p_n \tag{19}$$

However relation (15) requires the sum of the variations $\mathrm{d}p_n$ to be zero. To write that relation (19) is zero while taking into account this constraint, we use a Lagrange multiplier λ (Appendix V) and obtain the equation:

$$\sum_n \left[\lambda + 1 + \ln p_n\right] \ \mathrm{d}p_n = 0 \tag{20}$$

which must be satisfied for any $\mathrm{d}p_n$. Canceling the corresponding coefficients leads to:

$$\ln \ p_n = -\lambda - 1 \tag{21}$$

This means that all the non-zero p_n must be equal. Operator ρ is therefore proportional to the projector (3). Once we normalize its trace to 1, we get (2): the microcanonical density operator corresponds to an entropy extremum. As all the p_n are between 0 and 1, relation (18) shows that this extremum is positive. To find out if it is a maximum or a minimum, we consider another ρ operator, whose eigenvalues are all zero except one, equal to 1; its associated entropy S_ρ is zero. Consequently the extremum of S_ρ obtained for the microcanonical equilibrium is an absolute maximum.

This result proves an important theorem: the density operator that maximizes the entropy is the sum of the projectors onto all the accessible states, with equal eigenvalues (the probabilities of finding the physical system in each of these states).

1-b. Canonical ensemble

We now consider a physical system $\mathscr{S}$ no longer isolated but in contact with a reservoir $\mathcal{R}$ with which it exchanges energy; for example $\mathscr{S}$ and $\mathcal{R}$ could be coupled through a wall conducting heat but remaining fixed so that no work can be exchanged between them. Let us call $N_{\mathcal{R}}$, $E_{\mathcal{R}}$ and $\mathcal{V}_{\mathcal{R}}$ the particle number, the energy and the volume of the reservoir $\mathcal{R}$.

α. Density operator

Assuming the reservoir $\mathcal{R}$ to be much larger than the system $\mathscr{S}$, its temperature remains constant as it exchanges energy with $\mathscr{S}$. According to relation (12), this implies that its temperature $T_{\mathcal{R}}$, defined as:

$$\frac{1}{T_{\mathcal{R}}} = \left.\frac{\partial S_{\mathcal{R}}}{\partial E_{\mathcal{R}}}\right|_{N_{\mathcal{R}},\mathcal{V}_{\mathcal{R}}} \tag{22}$$

is a constant. It will be characterized by the constant β:

$$\beta = \frac{1}{k_B T_{\mathcal{R}}} = \frac{1}{k_B}\left.\frac{\partial S_{\mathcal{R}}}{\partial E_{\mathcal{R}}}\right|_{N_{\mathcal{R}},\mathcal{V}_{\mathcal{R}}} \tag{23}$$

As relation (10) showed that $S_{\mathcal{R}} = k_B \ln Z_{\mathcal{R}}$, where $Z_{\mathcal{R}}$ is the number of the reservoir's accessible levels in an energy band ΔE around $E_{\mathcal{R}}$, we deduce:

$$\left.\frac{\partial \ln Z_{\mathcal{R}}}{\partial E_{\mathcal{R}}}\right|_{N_{\mathcal{R}},\mathcal{V}_{\mathcal{R}}} = \beta \tag{24}$$

This means that this number of levels varies exponentially as a function of the energy (keeping $N_{\mathcal{R}}$ and $\mathcal{V}_{\mathcal{R}}$ constant):

$$Z_{\mathcal{R}} \propto e^{\beta E_{\mathcal{R}}} \tag{25}$$

The total system $\mathscr{S} + \mathcal{R}$ is described by an equilibrium microcanonical density operator. Its energy eigenvectors are the tensor product[1] of the energy eigenvectors $|E, q\rangle$ of the system $\mathscr{S}$ and the energy eigenvectors $|E_{\mathcal{R}}, q_{\mathcal{R}}\rangle$ of the system $\mathcal{R}$:

$$|E, q\rangle \otimes |E_{\mathcal{R}}, q_{\mathcal{R}}\rangle \tag{26}$$

The microcanonical density operator of $\mathscr{S} + \mathcal{R}$, with a total energy $E_{\text{tot}} = E' + E'_{\mathcal{R}}$ is given by:

$$\rho_{\mathscr{S}+\mathcal{R}} = \frac{1}{Z_{\mathscr{S}+\mathcal{R}}} \sum_{E'+E'_{\mathcal{R}}=E_{\text{tot}}-\frac{\Delta E}{2}}^{E_{\text{tot}}+\frac{\Delta E}{2}} \sum_{q,q_{\mathcal{R}}} \Big(|E', q\rangle \otimes |E'_{\mathcal{R}}, q_{\mathcal{R}}\rangle \Big) \Big(\langle E', q| \otimes \langle E'_{\mathcal{R}}, q_{\mathcal{R}}| \Big) \tag{27}$$

We get the density operator ρ_{eq} of the system $\mathscr{S}$ by taking a trace over the reservoir:

$$\rho_{\text{eq}} = \text{Tr}_{\mathcal{R}} \{\rho_{\mathscr{S}+\mathcal{R}}\} \tag{28}$$

[1] We assume that the coupling between $\mathscr{S}$ and $\mathcal{R}$ is weak, so that its contribution to the total energy is negligible.

In (27), the trace over $\mathcal{R}$ of each projector $|E'_{\mathcal{R}}, q_{\mathcal{R}}\rangle \langle E'_{\mathcal{R}}, q_{\mathcal{R}}|$ is just equal to one. The density operator ρ_{eq} is simply a sum of projectors onto the energy eigenstates $|E', q\rangle \langle E', q|$, multiplied by the number of levels of $\mathcal{R}$ with an energy $E_{\text{tot}} - E'$ within an energy band ΔE. Relation (25) shows that this number of levels varies exponentially as $e^{\beta E_{\mathcal{R}}} = e^{\beta\left(E_{\text{tot}} - E'\right)}$. Omitting the proportionality factors $1/Z_{\mathscr{S}+\mathcal{R}}$ and $e^{\beta E_{\text{tot}}}$, we get:

$$\rho_{\text{eq}} \propto \sum_{E'} \sum_{q} e^{-\beta E'} |E', q\rangle \langle E', q| = e^{-\beta \widehat{H}} \tag{29}$$

where $\widehat{H}$ is the Hamiltonian of the system $\mathscr{S}$. Normalizing the trace of ρ, we obtain:

$$\boxed{\rho_{\text{eq}} = \frac{1}{Z_c}\, e^{-\beta \widehat{H}}} \tag{30}$$

where Z_c is the "canonical partition function" defined as:

$$\boxed{Z_c = \text{Tr}\left\{e^{-\beta \widehat{H}}\right\}} \tag{31}$$

These two relations define the density operator of $\mathscr{S}$ in the canonical thermal equilibrium. Contrary to what happened in the microcanonical equilibrium, the energy of the system $\mathscr{S}$ is no longer restricted to a small interval ΔE, but may spontaneously fluctuate outside this energy band under the effect of the coupling with the reservoir.

The thermodynamic potential of the canonical equilibrium is defined by the function F called the "free energy":

$$F = \left\langle \widehat{H} \right\rangle - TS \tag{32}$$

At equilibrium, when ρ is given by (30), this free energy is equal to:

$$F = \left\langle \widehat{H} \right\rangle + k_B T\ \text{Tr}\left\{\rho_{\text{eq}}\ \left[-\ln Z_c - \beta \widehat{H}\right]\right\} = \left\langle \widehat{H} \right\rangle - \left\langle \widehat{H} \right\rangle - k_B T\ \ln Z_c \tag{33}$$

and we obtain:

$$\boxed{F = -k_B T\ \ln Z_c} \tag{34}$$

β. Minimization of the free energy

Starting from an arbitrary density operator ρ of unit trace, let us show that its associated free energy will be minimal when ρ is equal to its value at the canonical equilibrium (30). We first compute the variation of F:

$$\mathrm{d}F = \text{Tr}\left\{\left[\widehat{H} + k_B T\,(1 + \ln\rho)\right] \mathrm{d}\rho\right\} \tag{35}$$

This variation is zero for any $\mathrm{d}\rho$ only if the operator between the inner brackets is zero, which means:

$$\ln\ \rho = -1 - \beta \widehat{H} \tag{36}$$

This indicates that $\rho \propto e^{-\beta \widehat{H}}$, which is the canonical equilibrium operator. Finally, if we choose for ρ the projector onto a state having a large positive energy, S will be zero, $\left\langle \widehat{H} \right\rangle$ arbitrarily very large, and consequently F will be very large as well. It is thus clear that the extremum of F, which occurs when ρ takes the equilibrium value, is a minimum.

1-c. Grand canonical ensemble

We now assume that the physical system $\mathscr{S}$ can exchange not only energy but also particles with the reservoir $\mathcal{R}$: $\mathscr{S}$ and $\mathcal{R}$ must be coupled through an interface the particles can cross. As above, this reservoir is supposed large enough for its temperature to remain constant when it exchanges energy with $\mathscr{S}$. We also assume it contains a very large number of particles, which barely changes in relative value during the particle exchanges with $\mathscr{S}$. Its chemical potential therefore remains constant:

$$\mu = -T \left.\frac{\partial S_{\mathcal{R}}}{\partial N}\right|_{E_{\mathcal{R}},\mathcal{V}_{\mathcal{R}}} = \text{constant} \tag{37}$$

α. Density operator

As the particle numbers N of system $\mathscr{S}$ and N_R of the reservoir are no longer constant, their state spaces are now Fock spaces (Chapter XV). We must add indices N and $N_{\mathcal{R}}$ respectively to the kets $|E, q\rangle$ and $|E_{\mathcal{R}}, q_{\mathcal{R}}\rangle$ and there are now two summations over these indices in the expression for the microcanonical density operator $\rho_{\mathscr{S}+\mathcal{R}}$ written in (27). The microcanonical density operator of the total system is then:

$$\rho_{\mathscr{S}+\mathcal{R}} = \frac{1}{Z_{\mathscr{S}+\mathcal{R}}} \sum_{E'+E'_{\mathcal{R}}=E_{\rm tot}-\frac{\Delta E}{2}}^{E_{\rm tot}+\frac{\Delta E}{2}} \sum_{N'+N'_{\mathcal{R}}=N_{\rm tot}} \sum_{q,q_{\mathcal{R}}} \Big(|E', N', q\rangle \otimes |E'_{\mathcal{R}}, N'_{\mathcal{R}}, q_{\mathcal{R}}\rangle \Big) \Big(\langle E', N', q| \otimes \langle E'_{\mathcal{R}}, N'_{\mathcal{R}}, q_{\mathcal{R}}| \Big) \tag{38}$$

where $E_{\rm tot}$ and $N_{\rm tot}$ are respectively the energy and the particle number of the total system $\mathscr{S} + \mathcal{R}$. The argument then follows the same lines as in § 1-b-α. A partial trace over the reservoir leads to the density operator of $\mathscr{S}$, which is a linear combination of projectors:

$$|E', N', q\rangle \langle E', N', q| \tag{39}$$

with weights corresponding to the number of states of the reservoir in an energy band centered around $E'_{\mathcal{R}} = E_{\rm tot} - E'$, the number of particles in the reservoir being $N'_{\mathcal{R}} = N_{\rm tot} - N'$. Two reservoir variables change simultaneously, instead of one for the canonical equilibrium. As μ and β remain constant in (24) and (37), the entropy $S_{\mathcal{R}}$ varies linearly with respect to these variables:

$$S_{\mathcal{R}} = S^0_{\mathcal{R}} + \frac{1}{T} E'_{\mathcal{R}} - \frac{\mu}{T} N'_{\mathcal{R}} = S^0_{\mathcal{R}} + k_B \beta \left(E'_{\mathcal{R}} - \mu N'_{\mathcal{R}}\right) \tag{40}$$

where $S^0_{\mathcal{R}}$ is a constant that is of no importance in what follows. Using again relation (10) to relate the reservoir entropy to the number of states $Z_{\mathcal{R}}$ accessible to this reservoir, we get:

$$Z_{\mathcal{R}} = e^{S_{\mathcal{R}}/k_B} \propto e^{\beta\left(E'_{\mathcal{R}} - \mu N'_{\mathcal{R}}\right)} = e^{\beta(E_{\rm tot} - \mu N_{\rm tot})} \times e^{-\beta\left(E' - \mu N'\right)} \tag{41}$$

where μ and β remain constant. The same argument as above shows that the trace over the reservoir variables lead to the following density operator ρ for the system $\mathscr{S}$:

$$\boxed{\rho_{\text{eq}} = \frac{1}{Z_{gc}}\ e^{-\beta\left(\widehat{H}-\mu\widehat{N}\right)}} \tag{42}$$

with:

$$\boxed{Z_{\text{gc}} = \text{Tr}\left\{e^{-\beta\left(\widehat{H}-\mu\widehat{N}\right)}\right\}} \tag{43}$$

where $\widehat{N}$ is total particle number operator of $\mathscr{S}$.

β. Grand potential

The thermodynamic potential for the grand canonical ensemble is the "grand potential" Φ defined as:

$$\Phi = \langle\widehat{H}\rangle - TS - \mu\langle\widehat{N}\rangle \tag{44}$$

Following the same demonstration as for F in the canonical ensemble, we can show that the equilibrium value of this potential is:

$$\boxed{\Phi = -k_B T\ \ln Z_{gc}} \tag{45}$$

If we let ρ vary, we can show, as above, that this potential reaches a minimum when ρ is equal to (42); a detailed demonstration is given in § 1 of Complement G_{XV}.

2. Intensive or extensive physical quantities

Take a macroscopic physical system $\mathscr{S}$ at equilibrium, and divide it into two subsystems of equal sizes $\mathscr{S}'$ and $\mathscr{S}''$; one can imagine that a wall separates $\mathscr{S}'$ from $\mathscr{S}''$. Certain physical quantities associated with $\mathscr{S}'$ or $\mathcal{S}''$, taken separately, are half of what they were for $\mathscr{S}$: the volumes, the energies, the particle numbers, the entropies, etc. Such quantities are said to be "extensive". Inversely, other physical quantities do not change upon this division: the particle number per unit volume, the temperature, the chemical potential, etc. Such quantities are said to be "intensive". In a general way, when a macroscopic physical system of volume $\mathcal{V}$ is divided into several macroscopic parts of volumes $\mathcal{V}_1$, $\mathcal{V}_2$, etc., the physical quantities measured in each part and which are proportional to their respective volume are said to be extensive, and those which remain constant are said to be intensive.

As for the ensembles studied above, their description involves a mixture[2] of extensive and intensive variables:

(i) In the microcanonical ensemble, the three independent variables describing the physical system at equilibrium are the three extensive variables $\mathcal{V}$, N and the system's energy E; the other physical quantities (temperature, entropy, chemical potential, etc.) are

[2]including at least one extensive variable, otherwise the system's size would not be determined.

considered functions of these variables. The thermodynamic potential is the entropy S, extensive and directly related to the logarithm of the microcanonical partition function.

(ii) In the canonical ensemble, the three independent variables include two extensive variables $\mathcal{V}$ and N as well as an intensive variable T (or β). The thermodynamic potential is the free energy F, an extensive function directly related to the logarithm of the canonical partition function.

(iii) In the grand canonical ensemble, there is only one independent extensive variable, the volume $\mathcal{V}$, and two intensive variables T and μ. The thermodynamic potential is the function Φ, extensive and directly related to the logarithm of the grand canonical partition function.

For a macroscopic system, the three ensembles are generally considered equivalent. The statistical descriptions are, however, different. In the canonical equilibrium, for example, the energy is not restricted to an interval ΔE but can fluctuate and take on values outside this interval. However, for a macroscopic system, the fluctuations in energy are very small compared to its average value. Assuming that the system's energy is confined within a fixed band ΔE is a valid approximation and allows taking for $\langle \hat{H} \rangle$ the microcanonical energy E:

$$\langle \hat{H} \rangle = E \tag{46}$$

Another example, in the grand canonical ensemble, is the particle number, which fluctuates around its average value $\langle \hat{N} \rangle$. For a macroscopic system, the relative value of these fluctuations is in general[3] very small, and the average value $\langle \hat{N} \rangle$ is practically equal to the particle number N of the microcanonical or canonical ensembles:

$$\langle \hat{N} \rangle = N \tag{47}$$

2-a. Microcanonical ensemble

Relations (12) and (13) give the partial derivatives of the entropy S with respect to the variables E and N; we now compute that derivative with respect to the volume $\mathcal{V}$.

Let us change the physical system volume $\mathcal{V}$ by a small quantity $\mathrm{d}\mathcal{V}$, keeping the particle number N constant, and without any heat exchange (the system is surrounded by isolating walls). The system, having an internal pressure P, is doing the work $P\mathrm{d}\mathcal{V}$, which means that its internal energy varies as:

$$\mathrm{d}E = -P\mathrm{d}\mathcal{V} \tag{48}$$

As there is no heat exchange, $\mathrm{d}Q = 0$, and the thermodynamic relation $\mathrm{d}Q = T\mathrm{d}S$ means that the entropy does not change either:

$$\mathrm{d}S = \left.\frac{\partial S}{\partial \mathcal{V}}\right|_{N,E} \mathrm{d}\mathcal{V} + \left.\frac{\partial S}{\partial E}\right|_{N,\mathcal{V}} \mathrm{d}E = 0 \tag{49}$$

[3]There are exceptions to this rule: for a Bose-condensed ideal gas, the grand canonical fluctuations of the particles' number remain large for a macroscopic system. This is a very special system for which the canonical and grand canonical ensembles are not equivalent for certain physical properties.

Inserting relation (12) in this result and multiplying by T, we obtain:

$$T \left.\frac{\partial S}{\partial \mathcal{V}}\right|_{N,E} \mathrm{d}\mathcal{V} + \mathrm{d}E = 0 \tag{50}$$

As relation (48) shows that the pressure P is given by $-\mathrm{d}E/\mathrm{d}\mathcal{V}$ and taking (10) into account, we finally get:

$$P = T \left.\frac{\partial S}{\partial \mathcal{V}}\right|_{N,E} = k_B T \left.\frac{\partial \ln Z}{\partial \mathcal{V}}\right|_{N,E} \tag{51}$$

which defines the pressure in the microcanonical ensemble.

We already studied, in § 1-a-β, the entropy changes due to variations of either E (keeping $\mathcal{V}$ and N constant) or N (keeping E and $\mathcal{V}$ constant). The present calculation is the last step for obtaining the three partial derivatives of the microcanonical thermodynamic potential, and we can express its total derivative as:

$$\boxed{\mathrm{d}S = \frac{1}{T}\mathrm{d}E + \frac{P}{T}\mathrm{d}\mathcal{V} - \frac{\mu}{T}\mathrm{d}N} \tag{52}$$

2-b. Canonical ensemble

For a macroscopic system, we just saw that $\langle H \rangle$ could be replaced by the microcanonical energy E in the definition (32) of the free energy. Taking the differential of (32) then leads to:

$$\mathrm{d}F = \mathrm{d}E - T\mathrm{d}S - S\mathrm{d}T \tag{53}$$

Using (52) in the $T\mathrm{d}S$ term of this equation, the $\mathrm{d}E$ terms cancel out and we get:

$$\boxed{\mathrm{d}F = -S\mathrm{d}T - P\mathrm{d}\mathcal{V} + \mu \mathrm{d}N} \tag{54}$$

This is the total differential of the thermodynamic potential in the canonical ensemble.

This relation allows a physical interpretation of the chemical potential: it is the gain in free energy when one particle is added to the system[4], keeping constant the temperature and the volume of the system. As for the pressure P, it is given by:

$$P = -\left.\frac{\partial F}{\partial \mathcal{V}}\right|_{N,T} \tag{55}$$

or, using (34):

$$P = k_B T \left.\frac{\partial \ln Z_c}{\partial \mathcal{V}}\right|_{N,T} \tag{56}$$

which is similar to (51). We have obtained the pressure of the physical system as a function of its volume and its temperature, i.e. its "equation of state".

[4]When the temperature is zero, the free energy is just the energy E, and μ is the increase of energy when one particle is added.

To compute the average energy $\left\langle \widehat{H} \right\rangle$, we can use relations (30) and (31), which yields:

$$\left\langle \widehat{H} \right\rangle = \frac{1}{Z_c} \mathrm{Tr} \left\{ \widehat{H}\ e^{-\beta \widehat{H}} \right\} = \frac{-1}{Z_c} \frac{\partial}{\partial \beta} \mathrm{Tr} \left\{ e^{-\beta \widehat{H}} \right\} \tag{57}$$

where the partial derivative is taken keeping N and $\mathcal{V}$ constant. We then have:

$$\left\langle \widehat{H} \right\rangle = -\frac{1}{Z_c} \frac{\partial Z_c}{\partial \beta} = -\frac{\partial \ln Z_c}{\partial \beta} \tag{58}$$

.

2-c. Grand canonical ensemble

In a macroscopic system, as the particle number generally fluctuates very little in relative value, we can replace $\langle N \rangle$ by N in the definition (44) of the thermodynamic grand potential Φ. This leads to:

$$\mathrm{d}\Phi = \mathrm{d}F - \mu \mathrm{d}N - N \mathrm{d}\mu \tag{59}$$

Using (54) in this relation, the $\mu \mathrm{d}N$ terms cancel out and we are left with:

$$\boxed{\mathrm{d}\Phi = -S\mathrm{d}T - P\mathrm{d}\mathcal{V} - N\mathrm{d}\mu} \tag{60}$$

In this ensemble, the volume $\mathcal{V}$ is the only extensive variable. For a fixed temperature and chemical potential, and for a large volume, we get a macroscopic system whose energy, entropy and particle number are proportional to $\mathcal{V}$. We simply get:

$$\Phi = -P\mathcal{V} \tag{61}$$

The grand potential divided by $\mathcal{V}$ yields the pressure directly, without any partial derivative.

Taking (45) into account, the average particle number and the pressure obey:

$$\begin{aligned} N &= -\left. \frac{\partial \Phi}{\partial \mu} \right|_{\mathcal{V},T} = k_B T \left. \frac{\partial \ln Z_{gc}}{\partial \mu} \right|_{\mathcal{V},T} \\ P &= -\frac{\Phi}{\mathcal{V}} = \frac{k_B T}{\mathcal{V}} \ln Z_{gc} \end{aligned} \tag{62}$$

Using these two equalities to eliminate the chemical potential μ, we get the particle number in a given volume as a function of the pressure P and the temperature T (equation of state for the physical system).

2-d. Other ensembles

We have studied the three most commonly used statistical ensembles, but there are others, as for example the isothermal-isobaric ensemble. In this ensemble, the system $\mathscr{S}$ is coupled with a reservoir allowing exchanges of energy and volume, but not particles; the number N remains fixed. The only extensive variable is precisely this variable N,

the other two, the temperature T and the pressure P, being intensive. The thermodynamic potential associated with this isothermal-isobaric ensemble is the Gibbs function G defined as:

$$G = E + P\mathcal{V} - TS \tag{63}$$

As before, we take the differential of this function and note, using (52), that the terms $\mathrm{d}E$ and $P\mathrm{d}\mathcal{V}$ cancel out. We then get:

$$\mathrm{d}G = \mathcal{V}\mathrm{d}P - S\mathrm{d}T + \mu\mathrm{d}N \tag{64}$$

The function G is extensive. It increases as the particle number N gets larger (for fixed pressure and temperature), and for a macroscopic system it is proportional to the system size:

$$G = \mu N \tag{65}$$

Varying both the temperature and the pressure of an ensemble of N particles, we can get the resulting variation of the chemical potential μ by dividing (64) by N and then setting $\mathrm{d}N = 0$. This yields the Gibbs-Duhem relation:

$$\boxed{\mathrm{d}\mu = \frac{\mathcal{V}}{N}\,\mathrm{d}P - \frac{S}{N}\,\mathrm{d}T} \tag{66}$$

This ensemble is particularly useful in the study of a two-phase equilibrium such as a liquid and its vapor, both at the same pressure and temperature.

We have presented a brief review of the general principles of statistical mechanics. For more details, the reader may consult, for example, the following references [95, 96, 97, 3].

Appendix VII

Wigner transform

1 Delta function of an operator 2299

2 Wigner distribution of the density operator (spinless particle) 2299

2-a Definition of the distribution, Weyl operators 2300

2-b Expressions for the Wigner transform 2301

2-c Reality, normalization, operator form 2304

2-d Gaussian wave packet 2305

2-e Semiclassical situations 2306

2-f Quantum situations where the Wigner distribution is not a probability distribution 2309

3 Wigner transform of an operator 2310

3-a Average value of a Hermitian operator (observable) 2311

3-b Special cases 2312

3-c Wigner transform of an operator product 2313

3-d Evolution of the density operator 2316

4 Generalizations 2318

4-a Particle with spin 2318

4-b Several particles 2319

5 Discussion: Wigner distribution and quantum effects 2319

5-a An interference experiment 2319

5-b General discussion; "ghost" component 2322

Introduction

In classical mechanics, it is possible to specify with an arbitrary precision both the position $\mathbf{r}$ and the momentum $\mathbf{p}$ (and hence the velocity) of a particle. If the state of the particle is defined in a statistical way, one describes the classical particle by a distribution $\rho_{\text{cl}}(\mathbf{r}, \mathbf{p})$ in the phase space (Appendix III, § 3-a) , which can be any positive function, normalized to unity. This distribution can, for example, include correlations between the particle's position and velocity. In quantum mechanics, the situation is different. It is true that one often uses two representations, one in position space, the other in momentum space, and that we go from one to the other via a Fourier transformation. But in quantum mechanics these two representations are exclusive: in the position representation, one loses all information on the particle's momentum, and conversely, in the momentum representation one loses all information on the particle's position; consequently, no information on an eventual correlation between position and momentum can be obtained.

It is interesting to introduce a quantum point of view intermediate between these two extremes, to keep at the same time information about position and momentum,

while obeying the general rules of quantum mechanics; these rules impose a limitation on the information precision. The Wigner transformation offers this intermediate point of view as it introduces a quantum mechanical function $\rho_W(\mathbf{r},\mathbf{p})$ that allows computing average values in the same way as with a classical distribution $\rho_{\rm cl}(\mathbf{r},\mathbf{p})$. Historically, this transformation was introduced[1] in 1932 by Wigner [98, 99] as he was working on the quantum corrections to thermal equilibrium, but it turned out to be a much more general tool. It yields very naturally semiclassical expansions in powers of $\hbar$ while studying, for instance, the temporal evolution of a quantum system. We shall also show that it provides a quantization method, leading in particular to correctly symmetrized expressions for quantum operators starting from classical functions of the positions and momenta. There are numerous domains in physics where the Wigner transform has proven useful, and sometimes indispensable.

This appendix will show how to associate with any quantum density operator a *Wigner distribution* $\rho_W(\mathbf{r},\mathbf{p})$, sometimes called a "semiclassical distribution", and we will discuss a certain number of its properties. In a similar way, to any operator (observable) acting in the state space, we can associate a Wigner transform $A_W(\mathbf{r},\mathbf{p})$ that is a simple function of $\mathbf{r}$ and $\mathbf{p}$.

In classical or semiclassical situations (i.e. when the spatial variations of physical quantities occur over large enough distances), the function $\rho_W(\mathbf{r},\mathbf{p})$ possesses all the properties of a classical distribution: it is a positive function that, multiplied by $A_W(\mathbf{r},\mathbf{p})$ and then integrated over the variables $\mathbf{r}$ and $\mathbf{p}$, does yield the average value of the operator A. In this case, we will show that the Wigner distribution simply describes the flow of the probability fluid (Chapter III, § D-1-c-β). As $\rho_W(\mathbf{r},\mathbf{p})$ allows keeping track of the correlations between position and momentum, this function is particularly useful in a number of cases, such as in the theory of quantum transport, with its numerous applications.

In the general case where the quantum effects are important (rapid spatial variations), the classical relations between the distribution and the average values are still valid, meaning the Wigner distribution continues to be quite useful. This distribution shows, however, a significant difference with a probability distribution: the function $\rho_W(\mathbf{r},\mathbf{p})$ can sometimes take on negative values. Furthermore, as we shall point out later, it can sometimes present "ghost" components where it is different from zero at points where the probability of finding the particle is zero. It is thus not possible to interpret the product $\rho_W(\mathbf{r},\mathbf{p})\mathrm{d}^3r\mathrm{d}^3p$ as the probability for the particle to occupy an infinitesimal cell $\mathrm{d}^3r\mathrm{d}^3p$ of the phase space centered at $(\mathbf{r},\mathbf{p})$; moreover, when dealing with infinitesimal volumes such an interpretation would be in clear contradiction with Heisenberg's uncertainty relations. Therefore, one should consider $\rho_W(\mathbf{r},\mathbf{p})$ to be a "quasi-distribution", a tool for computing all the average values without being a real probability distribution, even though we shall use the word distribution, following common usage.

This appendix introduces the tools necessary for the study of these different situations. We shall obtain, in particular, gradient expansions directly yielding expansions in powers of $\hbar$. We first introduce in § 1 a convenient form for the delta function of an operator. This form will be useful, in § 2, for defining the Wigner distribution of a spinless particle density operator. Several of its characteristics will be studied, in particular

[1] Wigner does mention in his article that the same transformation had already been used by L. Szilard, but in another context.

when dealing with a Gaussian wave packet. We then focus in § 3 on the *Wigner transform* of operators and show how it can be associated with the Wigner distribution of the density operator, leading to computations similar to those corresponding to a classical distribution. An important step will be the computation of the Wigner transform associated with a product of operators. Generalization of these concepts to take into account the spin, as well as the possible presence of several interacting particles, is discussed in § 4. Finally, we focus in § 5 on the physical properties of the Wigner transformation, using it to analyze a quantum interference experiment. We will show, in particular, how "ghost components" of the Wigner transform can appear, rapidly change sign, and are the signature of quantum effects.

1. Delta function of an operator

Consider a Hermitian operator A having a continuous spectrum, and whose eigenvalues are noted α. We define the operator $D_A(a)$, depending on a real continuous parameter a, as:

$$D_A(a) = \frac{1}{2\pi} \int \mathrm{d}x \; e^{ix(A-a)} \tag{1}$$

It is, in a way, a "delta function of an operator" associated with the difference between operator A and the constant a. We note $|\alpha, s\rangle$ the eigenvectors of A; the index s (assumed to be discrete) accounts for the possible degeneracy of the eigenvalue α. In a quantum state defined by the density operator ρ, the average value of this operator is:

$$\begin{aligned}
\langle D_A(a) \rangle &= \frac{1}{2\pi} \int \mathrm{d}x \; \mathrm{Tr}\left\{ \rho \; e^{ix(A-a)} \right\} \\
&= \frac{1}{2\pi} \int \mathrm{d}x \int \mathrm{d}\alpha \sum_s \langle \alpha, s | \, e^{ix(A-a)} \, | \alpha, s \rangle \, \langle \alpha, s | \, \rho \, | \alpha, s \rangle \\
&= \frac{1}{2\pi} \int \mathrm{d}x \int \mathrm{d}\alpha \sum_s e^{ix(\alpha-a)} \;\; \langle \alpha, s | \, \rho \, | \alpha, s \rangle
\end{aligned} \tag{2}$$

The integral over $\mathrm{d}x$ yields $2\pi\delta(\alpha - a)$, and we obtain:

$$\langle D_A(a) \rangle = \sum_s \langle a, s | \, \rho \, | a, s \rangle = \mathcal{P}(a) \tag{3}$$

where $\mathcal{P}(a)\,\mathrm{d}a$ is the probability, in a measurement associated with operator A, of finding a result in the interval $[a, a + \mathrm{d}a]$.

2. Wigner distribution of the density operator (spinless particle)

Imagine now that the physical system under study is a spinless particle, described by a density operator ρ. Our purpose is to introduce a function $\rho_W(\mathbf{r}, \mathbf{p})$ yielding simultaneously information on the probability of measurement results on either the position operator $\mathbf{R}$, or the momentum operator $\mathbf{P}$. As these are incompatible observables, we should not expect to make highly accurate predictions concerning both types of measurements: the precision is limited by Heisenberg's uncertainty principle. This function

will have to make a compromise between the two types of information, resulting from an unavoidable quantum uncertainty. As already mentioned in the introduction, this function is actually a "quasi-distribution", even though it is commonly called the "Wigner distribution".

2-a. Definition of the distribution, Weyl operators

By analogy with (1), we can introduce the "Weyl operator", which will be noted $W(\mathbf{r},\mathbf{p})$. It is, in a way, a "delta function" of an operator, associated at the same time with the operators $\mathbf{R}-\mathbf{r}$ and $\mathbf{P}-\mathbf{p}$. We set:

$$W(\mathbf{r},\mathbf{p}) = \frac{1}{\hbar^3}\int \frac{\mathrm{d}^3\kappa}{(2\pi)^3}\int \frac{\mathrm{d}^3x}{(2\pi)^3}\ \mathrm{e}^{i[\boldsymbol{\kappa}\cdot(\mathbf{R}-\mathbf{r})+\mathbf{x}\cdot(\mathbf{P}-\mathbf{p})/\hbar]} \tag{4}$$

where the integration variable $\boldsymbol{\kappa}$ has the dimension of a wave vector (the inverse of a length), and $\mathbf{x}$ the dimension of a length. This operator is Hermitian, as can be shown by changing the sign of the two integration variables. For each quantum state, the Wigner distribution $\rho_W(\mathbf{r},\mathbf{p};t)$ is defined as the average value in that state of the Weyl operator:

$$\boxed{\rho_W(\mathbf{r},\mathbf{p};t) = \langle W(\mathbf{r},\mathbf{p})\rangle} \tag{5}$$

If the system under study is characterized by a density operator $\rho(t)$ whose trace equals 1, this definition amounts to:

$$\rho_W(\mathbf{r},\mathbf{p};t) = \mathrm{Tr}\left\{\rho(t)\,W(\mathbf{r},\mathbf{p})\right\} \tag{6}$$

On the other hand, if the system is described by a normalized state vector $|\Psi\rangle$, this definition becomes:

$$\rho_W(\mathbf{r},\mathbf{p};t) = \langle\Psi(t)|\,W(\mathbf{r},\mathbf{p})\,|\Psi(t)\rangle \tag{7}$$

Note that the density operator as well as the state vector are dimensionless. Taking into account the factors $\hbar$ introduced in (4), $W(\mathbf{r},\mathbf{p})$ and $\rho_W(\mathbf{r},\mathbf{p})$ have the dimensions of $\hbar^{-3}$; the product $\rho_W(\mathbf{r},\mathbf{p})\mathrm{d}^3r\mathrm{d}^3p$ is thus dimensionless, as a probability should be. We are now going to show that this distribution has numerous useful properties.

Later, we shall need the matrix elements $\langle\mathbf{r}_1|\,W(\mathbf{r},\mathbf{p})\,|\mathbf{r}_2\rangle$ of the operator $W(\mathbf{r},\mathbf{p})$ in the position representation. As demonstrated below, they are written:

$$\langle\mathbf{r}_1|\,W(\mathbf{r},\mathbf{p})\,|\mathbf{r}_2\rangle = \frac{1}{(2\pi\hbar)^3}\ e^{-i\mathbf{p}\cdot(\mathbf{r}_2-\mathbf{r}_1)/\hbar}\ \delta\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2}-\mathbf{r}\right) \tag{8}$$

Demonstration:

The demonstration uses relation (63) of Complement $\mathrm{B_{II}}$, which expresses the exponential of the sum of two operators A and B as a product of exponentials, provided both operators commute with their commutator $[A,B]$:

$$e^{A+B} = e^{A}\ e^{B}\ e^{-\frac{1}{2}[A,B]} \tag{9}$$

We choose:

$$A = i\,(\mathbf{R}-\mathbf{r})\cdot\boldsymbol{\kappa} \qquad \text{and} \qquad B = i\,(\mathbf{P}-\mathbf{p})\cdot\mathbf{x}/\hbar \tag{10}$$

The commutator of these two operators:

$$[A,B] = -i\;\boldsymbol{\kappa}\cdot\mathbf{x} \tag{11}$$

is a number, so that both A and B commute with their commutator. Inserting relation (9) into (4), we get:

$$\langle\mathbf{r}_1|\,W\,(\mathbf{r},\mathbf{p})\,|\mathbf{r}_2\rangle = \frac{1}{\hbar^3}\int\frac{\mathrm{d}^3\kappa}{(2\pi)^3}\int\frac{\mathrm{d}^3x}{(2\pi)^3}\;e^{i\boldsymbol{\kappa}\cdot\mathbf{x}/2}\,\langle\mathbf{r}_1|\,e^{i(\mathbf{R}-\mathbf{r})\cdot\boldsymbol{\kappa}}e^{i(\mathbf{P}-\mathbf{p})\cdot\mathbf{x}/\hbar}\,|\mathbf{r}_2\rangle \tag{12}$$

Now relation (13) of Complement $\mathrm{E_{II}}$ tells us that the action of operator $e^{i\mathbf{P}\cdot\mathbf{x}/\hbar}$ is a simple translation of the position eigenvalue by $-\mathbf{x}$, which means that:

$$e^{i(\mathbf{P}-\mathbf{p})\cdot\mathbf{x}/\hbar}\,|\mathbf{r}_2\rangle = e^{-i\mathbf{p}\cdot\mathbf{x}/\hbar}\,|\mathbf{r}_2-\mathbf{x}\rangle \tag{13}$$

The function to be integrated in relation (12) then becomes:

$$e^{i\boldsymbol{\kappa}\cdot\mathbf{x}/2}\;\langle\mathbf{r}_1|\,e^{i(\mathbf{R}-\mathbf{r})\cdot\boldsymbol{\kappa}}e^{i(\mathbf{P}-\mathbf{p})\cdot\mathbf{x}/\hbar}\,|\mathbf{r}_2\rangle = e^{i\boldsymbol{\kappa}\cdot\mathbf{x}/2}\;e^{i(\mathbf{r}_2-\mathbf{x}-\mathbf{r})\cdot\boldsymbol{\kappa}}\;e^{-i\mathbf{p}\cdot\mathbf{x}/\hbar}\;\delta\,(\mathbf{r}_1-\mathbf{r}_2+\mathbf{x}) \tag{14}$$

The delta function, integrated over d^3x in (12), leads to replacing in the exponent $\mathbf{x}$ by $\mathbf{r}_2-\mathbf{r}_1$, and $\mathbf{r}_2-\mathbf{x}$ by $\mathbf{r}_1$. We then get:

$$\langle\mathbf{r}_1|\,W\,(\mathbf{r},\mathbf{p})\,|\mathbf{r}_2\rangle = \frac{1}{(2\pi\hbar)^3}\int\frac{\mathrm{d}^3\kappa}{(2\pi)^3}\;e^{i\left[\boldsymbol{\kappa}\cdot\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2}-\mathbf{r}\right)-\mathbf{p}\cdot\left(\frac{\mathbf{r}_2-\mathbf{r}_1}{\hbar}\right)\right]} \tag{15}$$

The integration over $\mathrm{d}^3\kappa$ then yields a delta function, and we obtain relation (8).

2-b. Expressions for the Wigner transform

Definition (5) of the Wigner transformation may lead to various expressions for the Wigner transform, depending on the representation used in the state space.

α. Position representation

Using the position representation to calculate the trace appearing in (6), we get:

$$\begin{aligned}\rho_W\,(\mathbf{r},\mathbf{p};t) &= \int\mathrm{d}^3r_2\;\int\mathrm{d}^3r_1\;\;\langle\mathbf{r}_2|\,\rho\,(t)\,|\mathbf{r}_1\rangle\,\langle\mathbf{r}_1|\,W\,(\mathbf{r},\mathbf{p})\,|\mathbf{r}_2\rangle \\ &= \int\mathrm{d}^3R\;\int\mathrm{d}^3y\;\left\langle\mathbf{R}+\frac{\mathbf{y}}{2}\right|\rho\,(t)\left|\mathbf{R}-\frac{\mathbf{y}}{2}\right\rangle\left\langle\mathbf{R}-\frac{\mathbf{y}}{2}\right|W\,(\mathbf{r},\mathbf{p})\left|\mathbf{R}+\frac{\mathbf{y}}{2}\right\rangle\end{aligned} \tag{16}$$

where, on the second line, we used the integration variables $\mathbf{R} = (\mathbf{r}_1+\mathbf{r}_2)/2$ and $\mathbf{y} = \mathbf{r}_2-\mathbf{r}_1$. Using relation (8) in this expression, we get the function $\delta\,(\mathbf{R}-\mathbf{r})$, to be integrated over d^3R, which leads to:

$$\boxed{\rho_W\,(\mathbf{r},\mathbf{p};t) = \frac{1}{(2\pi\hbar)^3}\int\mathrm{d}^3y\;e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\;\left\langle\mathbf{r}+\frac{\mathbf{y}}{2}\right|\rho\,(t)\left|\mathbf{r}-\frac{\mathbf{y}}{2}\right\rangle} \tag{17}$$

This relation is often used as a definition of the Wigner distribution. Its integration over $\mathrm{d}^3p/\left(2\pi\hbar\right)^3$ yields a function $\delta\left(\mathbf{y}\right)$, which leads to:

$$\int \mathrm{d}^3p\ \rho_W\left(\mathbf{r},\mathbf{p};t\right) = \langle\mathbf{r}|\,\rho\left(t\right)|\mathbf{r}\rangle = n\left(\mathbf{r},t\right) \tag{18}$$

We confirm a property of the classical distributions in phase space: the integral of the distribution over the momenta yields the probability density $n\left(\mathbf{r},t\right)$ of finding the particle at point $\mathbf{r}$.

In the particular case where the particle is described by a pure state $|\Psi\rangle$ – see relation (7) – the definition of the Wigner distribution becomes:

$$\boxed{\rho_W\left(\mathbf{r},\mathbf{p};t\right) = \frac{1}{\left(2\pi\hbar\right)^3}\int \mathrm{d}^3y\ e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\ \Psi\left(\mathbf{r}+\frac{\mathbf{y}}{2};t\right)\Psi^*\left(\mathbf{r}-\frac{\mathbf{y}}{2};t\right)} \tag{19}$$

where $\Psi\left(\mathbf{r};t\right)$ is the wave function in the position representation:

$$\Psi\left(\mathbf{r};t\right) = \langle\mathbf{r}\,|\Psi\left(t\right)\rangle \tag{20}$$

β. Momentum representation

As position and momentum play symmetrical roles in the argument, we expect to find a similar relation for the Wigner distribution, involving now the matrix elements of the density operator ρ in the momentum representation. This is indeed the case, and we are going to show that:

$$\boxed{\rho_W\left(\mathbf{r},\mathbf{p};t\right) = \frac{1}{\left(2\pi\hbar\right)^3}\int \mathrm{d}^3q\ e^{i\mathbf{q}\cdot\mathbf{r}/\hbar}\left\langle\mathbf{p}+\frac{\mathbf{q}}{2}\right|\rho\left(t\right)\left|\mathbf{p}-\frac{\mathbf{q}}{2}\right\rangle} \tag{21}$$

This expression is the exact analog of (17); it can be considered as an alternative definition of the Wigner distribution. As for the analogy of the property expressed by (18), it is easy to show that:

$$\int \mathrm{d}^3r\ \rho_W\left(\mathbf{r},\mathbf{p};t\right) = \langle\mathbf{p}|\,\rho\left(t\right)|\mathbf{p}\rangle \tag{22}$$

Just as for a classical distribution, the integral over the positions of the Wigner distribution yields the probability density of finding a given momentum $\mathbf{p}$.

Demonstration:

Inserting in the matrix element of (17) two closure relations on normalized momentum plane waves $\mathbf{q}'$ and $\mathbf{q}''$, yields the two scalar products:

$$\left\langle\mathbf{r}+\frac{\mathbf{y}}{2}\,\right|\left.\mathbf{q}'\right\rangle = \frac{1}{\left(2\pi\hbar\right)^{3/2}}e^{i\mathbf{q}'\cdot\left(\mathbf{r}+\frac{\mathbf{y}}{2}\right)/\hbar} \qquad \left\langle\mathbf{q}''\,\right|\left.\mathbf{r}-\frac{\mathbf{y}}{2}\right\rangle = \frac{1}{\left(2\pi\hbar\right)^{3/2}}e^{-i\mathbf{q}''\cdot\left(\mathbf{r}-\frac{\mathbf{y}}{2}\right)/\hbar} \tag{23}$$

and we can write:

$$\rho_W\left(\mathbf{r},\mathbf{p};t\right) = \frac{1}{\left(2\pi\hbar\right)^6}\int \mathrm{d}^3y\ e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\int \mathrm{d}^3q'\int \mathrm{d}^3q''\ e^{i\mathbf{q}'\cdot\left(\mathbf{r}+\frac{\mathbf{y}}{2}\right)/\hbar}e^{-i\mathbf{q}''\cdot\left(\mathbf{r}-\frac{\mathbf{y}}{2}\right)/\hbar}\left\langle\mathbf{q}'\right|\rho\left(t\right)\left|\mathbf{q}''\right\rangle \tag{24}$$

The summation over d^3y yields a delta function:

$$\frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3y\; e^{i\left(\frac{\mathbf{q}'+\mathbf{q}''}{2}-\mathbf{p}\right)\cdot\mathbf{y}/\hbar}=\delta\left(\frac{\mathbf{q}'+\mathbf{q}''}{2}-\mathbf{p}\right) \tag{25}$$

so that if we take $\mathbf{Q}=\left(\mathbf{q}'+\mathbf{q}''\right)/2$ and $\mathbf{q}=\mathbf{q}'-\mathbf{q}''$ as integration variables, we obtain (21).

In the particular case where the particle is described by a pure state $|\Psi\rangle$, relation (21) becomes:

$$\boxed{\rho_W\left(\mathbf{r},\mathbf{p};t\right)=\frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3q\; e^{i\mathbf{q}\cdot\mathbf{r}}\;\overline{\Psi}\left(\mathbf{p}+\frac{\mathbf{q}}{2};t\right)\overline{\Psi}^*\left(\mathbf{p}-\frac{\mathbf{q}}{2};t\right)} \tag{26}$$

where $\overline{\Psi}\left(\mathbf{p};t\right)$ is the wave function in the momentum representation:

$$\overline{\Psi}\left(\mathbf{p};t\right)=\langle\mathbf{p}\,|\Psi\left(t\right)\rangle \tag{27}$$

If the wave function is factored[2]:

$$\overline{\Psi}\left(\mathbf{p};t\right)=\overline{\psi}_x\left(p_x;t\right)\times\overline{\psi}_y\left(p_y;t\right)\times\overline{\psi}_z\left(p_z;t\right) \tag{28}$$

relation (26) shows that the Wigner transform is also factored:

$$\rho_W\left(\mathbf{r},\mathbf{p};t\right)=\rho_W^x\left(x,p_x;t\right)\times\rho_W^y\left(y,p_y;t\right)\times\rho_W^z\left(z,p_z;t\right) \tag{29}$$

with:

$$\rho_W^x\left(x,p_x;t\right)=\frac{1}{2\pi\hbar}\int \mathrm{d}q\; e^{iqx/\hbar}\;\overline{\psi}_x\left(p_x+\frac{q}{2};t\right)\overline{\psi}_x^*\left(p_x-\frac{q}{2};t\right) \tag{30}$$

The other two components $\rho_W^y\left(y,p_y;t\right)$ and $\rho_W^z\left(z,p_z;t\right)$ are defined in a similar way. In this particular case, one can reason independently for the three dimensions.

γ. Inverting the relations

We just saw that to each density operator corresponds a unique and well defined Wigner distribution. Inversely, starting from this distribution, one can reconstruct the corresponding density operator via its matrix elements. To take the inverse Fourier transform of (17), we multiply this relation by $e^{i\mathbf{p}\cdot\mathbf{z}/\hbar}$ and integrate over d^3p:

$$\int \mathrm{d}^3p\; e^{i\mathbf{p}\cdot\mathbf{z}/\hbar}\;\rho_W\left(\mathbf{r},\mathbf{p};t\right)=\frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3p\int \mathrm{d}^3y\; e^{i\mathbf{p}\cdot(\mathbf{z}-\mathbf{y})/\hbar}\left\langle\mathbf{r}+\frac{\mathbf{y}}{2}\right|\rho\left(t\right)\left|\mathbf{r}-\frac{\mathbf{y}}{2}\right\rangle \tag{31}$$

On the right-hand side, the integral over d^3p yields a function $(2\pi\hbar)^3\,\delta\left(\mathbf{z}-\mathbf{y}\right)$, so that this equality becomes:

$$\int \mathrm{d}^3p\; e^{i\mathbf{p}\cdot\mathbf{z}/\hbar}\;\rho_W\left(\mathbf{r},\mathbf{p};t\right)=\left\langle\mathbf{r}+\frac{\mathbf{z}}{2}\right|\rho\left(t\right)\left|\mathbf{r}-\frac{\mathbf{z}}{2}\right\rangle \tag{32}$$

[2]Whether the wave function is factored in the momentum or position representation is equivalent.

Setting $\mathbf{r}_1 = \mathbf{r} + \mathbf{z}/2$ and $\mathbf{r}_2 = \mathbf{r} - \mathbf{z}/2$, we obtain the matrix elements of ρ in the position representation:

$$\boxed{\langle \mathbf{r}_1 | \, \rho \, (t) \, | \mathbf{r}_2 \rangle = \int \mathrm{d}^3 p \; e^{i \mathbf{p} \cdot (\mathbf{r}_1 - \mathbf{r}_2)/\hbar} \; \rho_W \left(\frac{\mathbf{r}_1 + \mathbf{r}_2}{2}, \mathbf{p}; t \right)} \tag{33}$$

Knowing the Wigner distribution $\rho_W (\mathbf{r}, \mathbf{p})$ thus defines the operator ρ in a unique way.

In a similar way, multiplying (21) by $e^{-i \mathbf{p}' \cdot \mathbf{r}/\hbar}$, integrating over $\mathrm{d}^3 r$, then setting $\mathbf{p}' = \mathbf{p}_1 - \mathbf{p}_2$ and $\mathbf{p} = (\mathbf{p}_1 + \mathbf{p}_2)/2$, yields the inversion relation in the momentum representation:

$$\boxed{\langle \mathbf{p}_1 | \, \rho \, (t) \, | \mathbf{p}_2 \rangle = \int \mathrm{d}^3 r \; e^{i (\mathbf{p}_2 - \mathbf{p}_1) \cdot \mathbf{r}/\hbar} \; \rho_W \left(\mathbf{r}, \frac{\mathbf{p}_1 + \mathbf{p}_2}{2}; t \right)} \tag{34}$$

2-c. Reality, normalization, operator form

Let us take the Hermitian conjugate of relation (17). As the density operator is Hermitian, the matrix element on the right-hand side becomes:

$$\left\langle \mathbf{r} + \frac{\mathbf{y}}{2} \right| \rho \, (t) \left| \mathbf{r} - \frac{\mathbf{y}}{2} \right\rangle^* = \left\langle \mathbf{r} - \frac{\mathbf{y}}{2} \right| \rho \, (t) \left| \mathbf{r} + \frac{\mathbf{y}}{2} \right\rangle \tag{35}$$

Changing the sign of the integration variable $\mathbf{y}$ then yields again relation (17); the distribution $\rho_W (\mathbf{r}, \mathbf{p})$ is therefore equal to its complex conjugate, meaning it is real.

We now compute the integral of $\rho_W (\mathbf{r}, \mathbf{p})$ over the entire phase space. Summing (18) over $\mathrm{d}^3 r$, we get:

$$\int \mathrm{d}^3 r \int \mathrm{d}^3 p \; \rho_W (\mathbf{r}, \mathbf{p}; t) = \int \mathrm{d}^3 r \;\; \langle \mathbf{r} | \, \rho \, (t) \, | \mathbf{r} \rangle = \mathrm{Tr} \, \{ \rho \, (t) \} = 1 \tag{36}$$

where the last equality comes from the fact that the density operator has a trace equal to one. The Wigner distribution of the density operator is thus a real function, normalized to one in phase space, as is the case for a classical distribution.

Comment:

The density operator must obey a stronger constraint than having its trace equal to unity. It is defined as positive definite, meaning that for any ket $|\varphi\rangle$, we must have:

$$\langle \varphi | \, \rho \, (t) \, | \varphi \rangle \geq 0 \tag{37}$$

Now this condition is not merely equivalent to a positivity condition for the Wigner transform. In fact, we shall see below that the Wigner distribution of a density operator can become negative at certain points of phase space. However, the only Wigner distributions $\rho_W (\mathbf{r}, \mathbf{p}; t)$ acceptable for describing quantum systems are those that lead to a density operator obeying constraint (37). To know if a distribution in phase space is acceptable or not, is thus more difficult to decide in quantum mechanics than in classical mechanics.

We now show that relations (33) and (34) can be written in a simple operator form:

$$\rho(t) = (2\pi\hbar)^3 \int \mathrm{d}^3 r \int \mathrm{d}^3 p \; \rho_W(\mathbf{r},\mathbf{p};t) \; W(\mathbf{r},\mathbf{p}) \tag{38}$$

In this relation, $\rho_W(\mathbf{r},\mathbf{p};t)$ is the Wigner distribution, hence a function of position and momentum, but $W(\mathbf{r},\mathbf{p})$ is the Weyl operator defined in (4). To prove relation (38), we calculate the matrix element of this equality between the bra $\langle \mathbf{r}_1|$ and the ket $|\mathbf{r}_2\rangle$ to verify that we indeed obtain relation (33). Taking (8) into account, the matrix elements of the right-hand side are:

$$\begin{aligned}
(2\pi\hbar)^3 \int \mathrm{d}^3 r \int \mathrm{d}^3 p \; \rho_W(\mathbf{r},\mathbf{p};t) \;\; & \langle \mathbf{r}_1 | \, W(\mathbf{r},\mathbf{p}) \, |\mathbf{r}_2\rangle \\
&= \int \mathrm{d}^3 r \int \mathrm{d}^3 p \; \rho_W(\mathbf{r},\mathbf{p};t) \;\; e^{-i\mathbf{p}\cdot(\mathbf{r}_2-\mathbf{r}_1)/\hbar} \; \delta\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2} - \mathbf{r}\right) \\
&= \int \mathrm{d}^3 p \; \rho_W\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2},\mathbf{p};t\right) \;\; e^{-i\mathbf{p}\cdot(\mathbf{r}_2-\mathbf{r}_1)/\hbar}
\end{aligned} \tag{39}$$

which is equivalent to the right-hand side of (33).

2-d. Gaussian wave packet

A particular case where the computation can be completed is the one-dimensional Gaussian wave packet, studied in Complement $\mathrm{G_I}$. Relation (1) of this complement yields the normalized wave function $\psi_x(x)$ of such a wave packet in the position representation. We slightly modify it to center the wave packet at an arbitrary non-zero position x_0, and replace the variable k by $p = \hbar k$ (setting in particular $p_0 = \hbar k_0$). We obtain:

$$\begin{aligned}
\psi_x(x) &= \frac{\sqrt{a}}{(2\pi)^{3/4}} \int \frac{\mathrm{d}p}{\hbar} \; e^{-a^2(p-p_0)^2/4\hbar^2} \;\; e^{ip(x-x_0)/\hbar} \\
&= \left(\frac{2}{\pi a^2}\right)^{1/4} e^{ip_0(x-x_0)/\hbar} \; e^{-(x-x_0)^2/a^2}
\end{aligned} \tag{40}$$

where the second equality corresponds to relation (9) of Complement $\mathrm{G_I}$ (within an x_0 translation along Ox). The wave functions $(2\pi\hbar)^{-1/2} e^{ipx/\hbar}$ correspond to plane waves with momentum p (normalized with respect to that momentum); looking at the first equality in (40) we recognize the wave function in the momentum representation:

$$\overline{\psi}_x(p) = \frac{1}{(2\pi)^{1/4}} \sqrt{\frac{a}{\hbar}} \; e^{-a^2(p-p_0)^2/4\hbar^2} \;\; e^{-ipx_0/\hbar} \tag{41}$$

The Wigner distribution (30) can then be written (to simplify, we temporarily ignore the time dependence):

$$\begin{aligned}
\rho_W^x(x,p_x) &= \frac{a}{(2\pi)^{3/2}\hbar^2} \int \mathrm{d}q \; e^{-\frac{a^2}{4\hbar^2}\left[\left(p_x-p_0+\frac{q}{2}\right)^2+\left(p_x-p_0-\frac{q}{2}\right)^2\right]} \;\; e^{i\frac{q(x-x_0)}{\hbar}} \\
&= \frac{a}{(2\pi)^{3/2}\hbar^2} e^{-\frac{a^2}{2\hbar^2}(p_x-p_0)^2} \int \mathrm{d}q \; e^{-\frac{a^2q^2}{8\hbar^2}} \;\; e^{i\frac{q(x-x_0)}{\hbar}}
\end{aligned} \tag{42}$$

The integral over dq is a Fourier transform whose value can be obtained by replacing a by $a/\sqrt{2}$ in relation (50) de l'Appendice I. We get:

$$\rho_W^x(x,p_x) = \frac{a}{(2\pi)^{3/2}\hbar^2} e^{-\frac{a^2}{2\hbar^2}(p_x-p_0)^2} \times \sqrt{2\pi\hbar} \times \frac{\sqrt{2\hbar}}{\left(a/\sqrt{2}\right)} e^{-2(x-x_0)^2/a^2}$$
$$= \frac{1}{\pi\hbar} e^{-\frac{a^2}{2\hbar^2}(p_x-p_0)^2} e^{-2\frac{(x-x_0)^2}{a^2}} \tag{43}$$

Taking (40) and (41) into account, we see that the Wigner distribution is simply the product of the probability densities in the position and momentum spaces:

$$\rho_W^x(x,p_x) = |\psi_x(x)|^2 \times |\overline{\psi}_x(p)|^2 \tag{44}$$

This result is particularly simple and shows that the Wigner transform of a Gaussian wave packet (40) contains no correlations between the variables x and p. It can be factored into two Gaussian functions, one concerning the momentum, the other the position. The first one is centered on the average momentum p_0, and has a width of the order of $\hbar/a$; the second, on the average position x_0, with a width of the order of a. These two widths are within the boundaries imposed by Heisenberg relations. Note that, in this case, the Wigner transform remains positive for all the values of its variables, as will also be the case for the semiclassical situations we consider in § 2-e.

Comment:

In the preceding paragraph, we ignored the time dependence of the wave packet. To take it into account and assuming we are dealing with a free particle, one can multiply, in (41), $\overline{\psi}_x(p)$ by $e^{-i\omega_p t}$, with:

$$\omega_p = \frac{p^2}{2m\hbar} \tag{45}$$

where m is the mass of the particle. This introduces, in the integral on the second line of (42), an additional exponential $e^{-ip_x qt/(m\hbar)}$ whose effect on the Fourier transform is to make the following substitution:

$$x \Rightarrow x - \frac{p_x}{m}\, t \tag{46}$$

This simply corresponds to the motion of the particle with a velocity p/m. Making this substitution in (43), we find that the Wigner distribution is still a product of two Gaussian functions, but no longer the product of a function of momentum by a function of position: correlations have appeared between the momentum and position variables.

2-e. Semiclassical situations

To what extent is it possible to consider the Wigner transform to be a true probability distribution? Relations (18) and (22) seem to be in favor of it, as they show that integrating that distribution over the momenta (or over the positions) actually yields a probability distribution of finding the particle at a given point (or with a given momentum). These two "marginal distributions" thus obtained by integration are both probability distributions. But this is not sufficient to ensure that the function $\rho_W(\mathbf{r},\mathbf{p})$

itself (before integration) has the same property. Actually, we already mentioned in the introduction that it is not possible, in general, to interpret the product $\rho_W(\mathbf{r},\mathbf{p})\,\mathrm{d}^3r\mathrm{d}^3p$ as yielding directly the probability for a particle to occupy an infinitesimal cell $\mathrm{d}^3r\mathrm{d}^3p$ of phase space, centered at $(\mathbf{r},\mathbf{p})$. Such a probability distribution is meaningless in quantum mechanics, as Heisenberg's relations forbid the existence of a quantum state defined with an arbitrary precision both in position and momentum spaces.

There are, however, some simple cases, that we shall call "semiclassical", where the Wigner transform is very similar to a classical probability distribution. They correspond to situations we will now study, where the physical quantities vary sufficiently slowly in space compared to a scale we shall define explicitly. In the following section, we will consider more general situations, where the properties of the Wigner transform are radically different. In particular, the Wigner transform can become negative, which immediately excludes any interpretation in terms of probability density.

α. Wave packet with slow spatial variations

Consider the wave function:

$$\psi(\mathbf{r}) = C(\mathbf{r})\, e^{iS(\mathbf{r})} \tag{47}$$

where $C(\mathbf{r})$ is the modulus of the wave function and $S(\mathbf{r})$ its phase. The probability density of presence is then $[C(\mathbf{r})]^2$, while relation (D-17) of Chapter III yields the probability current $\mathbf{J}(\mathbf{r})$:

$$\mathbf{J}(\mathbf{r}) = \frac{\hbar}{m}\,[C(\mathbf{r})]^2\ \boldsymbol{\nabla} S(\mathbf{r}) \tag{48}$$

The matrix elements of the corresponding density operator are written:

$$\langle \mathbf{r}'|\,\rho(t)\,|\mathbf{r}''\rangle = C(\mathbf{r}')\,C(\mathbf{r}'')\ e^{i\left[S(\mathbf{r}')-S(\mathbf{r}'')\right]} \tag{49}$$

We assume that, in the vicinity of each point $\mathbf{r}$, the wave function behaves locally as a plane wave:

$$\psi(\mathbf{r}) \simeq C(\mathbf{r})\, e^{i[\mathbf{K}(\mathbf{r})\cdot\mathbf{r}+\alpha(\mathbf{r})]} \qquad \text{in the vicinity of each point } \mathbf{r} \tag{50}$$

and that the two functions $C(\mathbf{r})$ and $\mathbf{K}(\mathbf{r})$, as well as the phase $\alpha(\mathbf{r})$, vary slowly in space: their variations are negligible over distances of the order of the de Broglie wavelength $2\pi/k(\mathbf{r})$. When $\mathbf{r}'$ and $\mathbf{r}''$ are close enough, one can expand the argument of the exponential in (49); the matrix elements $\langle \mathbf{r}'|\,\rho(t)\,|\mathbf{r}''\rangle$ of the density operator are then written:

$$\langle \mathbf{r}'|\,\rho(t)\,|\mathbf{r}''\rangle \simeq C(\mathbf{r}')\,C(\mathbf{r}'')\ e^{i\mathbf{K}\cdot\left(\mathbf{r}'-\mathbf{r}''\right)} \tag{51a}$$

where $\mathbf{K}$ is defined as:

$$\mathbf{K} = \boldsymbol{\nabla} S(\mathbf{r} = \frac{\mathbf{r}'+\mathbf{r}''}{2}) \tag{51b}$$

β. Density operator; link with the probability fluid

To characterize a semiclassical situation in a more general way, we will now reason in terms of a density operator, without restricting our study to a pure state as we did earlier. To start, we assume there is no long-range non-diagonal order[3]:

$$|\langle \mathbf{r}'| \rho(t) |\mathbf{r}''\rangle| \to 0 \qquad \text{if} \qquad |\mathbf{r}' - \mathbf{r}''| \gg l \tag{52}$$

where l is a macroscopic coherence length. For a pure state, l would be determined by the size of the domain where the wave function has a non-zero modulus $C(r)$. For a statistical mixture of states, we have a different situation: the phases of the various wave functions may interfere destructively at shorter distances, so that l can be much smaller. Nevertheless, we shall assume that l remains larger than a few de Broglie wavelengths $2\pi/k(\mathbf{r}')$ and that when $|\mathbf{r}' - \mathbf{r}''| \lesssim l$, the non-diagonal matrix elements of the density operator vary in a similar way as (51a):

$$\langle \mathbf{r}'| \rho(t) |\mathbf{r}''\rangle \simeq F(\mathbf{r}', \mathbf{r}'')\ e^{i\mathbf{K}\cdot(\mathbf{r}'-\mathbf{r}'')} \qquad \text{if } |\mathbf{r}' - \mathbf{r}''| \lesssim l \tag{53}$$

This expression is simply the generalization of (51a), only valid for a pure state; the real function $F(\mathbf{r}', \mathbf{r}'')$ replaces the modulus product $C(\mathbf{r}')\,C(\mathbf{r}'')$. Both functions $F(\mathbf{r}', \mathbf{r}'')$ and $\mathbf{K}$ are supposed to remain practically constant as the variables $\mathbf{r}'$ and $\mathbf{r}''$ vary by a quantity of the order of l.

With these assumptions, the values of the integration variable giving a significant contribution to the integral in relation (17) correspond to $|\mathbf{y}| \lesssim l$, so that we can write:

$$\rho_W(\mathbf{r}, \mathbf{p}) \simeq \frac{1}{(2\pi\hbar)^3} \int_D \mathrm{d}^3 y\ e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\ F\left(\mathbf{r} + \frac{\mathbf{y}}{2}, \mathbf{r} - \frac{\mathbf{y}}{2}\right)\ e^{i\mathbf{K}(\mathbf{r})\cdot\mathbf{y}} \tag{54}$$

where the integration domain D is centered at $\mathbf{y} = 0$ and extends over a few coherence lengths l. As the function F is practically constant in this domain, and since $F(\mathbf{r}, \mathbf{r}) = \langle \mathbf{r}| \rho(t) |\mathbf{r}\rangle$, we get:

$$\rho_W(\mathbf{r}, \mathbf{p}) \simeq \frac{1}{(2\pi\hbar)^3} \langle \mathbf{r}| \rho(t) |\mathbf{r}\rangle \int_D \mathrm{d}^3 y\ e^{i[\mathbf{K}(\mathbf{r})\cdot\mathbf{y} - \mathbf{p}\cdot\mathbf{y}/\hbar]} \tag{55}$$

or else:

$$\rho_W(\mathbf{r}, \mathbf{p}) \simeq \langle \mathbf{r}| \rho(t) |\mathbf{r}\rangle \times \overline{g}[\mathbf{p} - \mathbf{p}_0(\mathbf{r})] \tag{56}$$

To write this expression, we have used the following definitions:

$$\overline{g}(\mathbf{p}) = \frac{1}{(2\pi\hbar)^3} \int_D \mathrm{d}^3 y\ e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar} \tag{57}$$

and:

$$\mathbf{p}_0(\mathbf{r}) = \hbar\, \mathbf{K}(\mathbf{r}) \tag{58}$$

[3]The concept of long-range non-diagonal order is introduced in Complement $\mathrm{A_{XVI}}$, §§ 2-a and 3-c, where, in particular, its relation with Bose-Einstein condensation is established. The present hypothesis concerning the absence of long-range order prevents $\rho(t)$ from being the one-body density operator of a system of condensed bosons.

The function $\overline{g}(\mathbf{p})$ is a momentum distribution centered at $\mathbf{p} = 0$, with a width $\Delta p \sim \hbar/l$. It is normalized to unity as the integral of $\overline{g}(\mathbf{p})$ over the momenta yields a function $\delta(\mathbf{y})$, which integrated over $\mathrm{d}^3 y$ is equal to one. Note that in (56) this function takes on its value for a momentum equal to $\mathbf{p} - \mathbf{p}_0(\mathbf{r})$, which means the momentum distribution is centered at the value $\mathbf{p}_0(\mathbf{r})$. As this momentum value depends on $\mathbf{r}$, correlations between position and momenta are now introduced in $\rho_W(\mathbf{r}, \mathbf{p})$.

Expression (56) for the distribution $\rho_W(\mathbf{r}, \mathbf{p})$ can be interpreted as a classical distribution in the probability fluid phase space: it is the product of the local probability density $\langle \mathbf{r} | \rho(t) | \mathbf{r} \rangle$ by a function of momentum $\overline{g}[\mathbf{p} - \mathbf{p}_0(\mathbf{r})]$ centered around the value $\mathbf{p}_0(\mathbf{r})$ defined in (58). Now this $\mathbf{p}_0(\mathbf{r})$ value is precisely the momentum value that, divided by m (to go from momentum to velocity) and multiplied by the probability density, yields the fluid probability current $\mathbf{J}(\mathbf{r})$. Note that the distribution keeps a certain width around $\mathbf{p}_0(\mathbf{r})$, of the order of $\hbar/l$, as required by Heisenberg's uncertainty relation. To sum up, in such semiclassical situations, the Wigner distribution directly reflects the spatial variation of the probability, and of its associated local current. It simply describes the flow of a "probability fluid" (III, § D-1-c-β), as does the distribution in phase space of an ensemble of classical particles forming a moving fluid.

2-f. Quantum situations where the Wigner distribution is not a probability distribution

In the previous examples, the properties of the Wigner distribution are very similar to those of a classical distribution. This is, however, not always the case: as surprising as it may seem, the Wigner transform can, in general, become negative.

α. *Odd wave function*

A very simple case offers such an example. In a one-dimension problem, imagine that the system has an odd wave function, as is the case for example for the first excited state of the harmonic oscillator. We then have, according to relation (19):

$$\begin{aligned} \rho_W^x(x = 0, p_x = 0) &= \frac{1}{2\pi\hbar} \int \mathrm{d}y \ \psi_x\left(\frac{y}{2}\right) \psi_x^*\left(-\frac{y}{2}\right) \\ &= -\frac{1}{2\pi\hbar} \int \mathrm{d}y \ \left|\psi_x\left(\frac{y}{2}\right)\right|^2 \end{aligned} \tag{59}$$

which is obviously negative. As odd wave functions often occur in quantum mechanics, we see that there exist numerous situations where the Wigner distribution has some properties unexpected for a distribution. Strictly speaking, the term "quasi-distribution" should always be used.

β. *Two-peak wave function*

Imagine now the particle wave function is the sum of two wave packets, one localized around $x = +b$, the other around $x = -b$:

$$\psi_x(x) = \frac{1}{\sqrt{2}} \left[\varphi(x - b) + e^{i\chi} \varphi(x + b)\right] \tag{60}$$

where the wave function $\varphi(x)$ is normalized; the relative phase factor χ is arbitrary. For the sake of simplicity, we assume that $\varphi(x)$ is zero when $|x| > a$ and that it is even. We

also suppose that in our case $b \gg a$, meaning that the two wave packets forming the total wave function are well separated.

Let us compute the Wigner distribution at point $x = 0$, therefore at a point where the wave function $\psi_x(x)$ is zero. In one dimension, relation (19) is written as:

$$\rho_W^x(x=0,p_x) = \frac{1}{4\pi\hbar}\int \mathrm{d}y\; e^{-ip_x y/\hbar}\left[\varphi\left(\frac{y}{2}-b\right)+e^{i\chi}\varphi\left(\frac{y}{2}+b\right)\right] \times\left[\varphi^*\left(-\frac{y}{2}-b\right)+e^{-i\chi}\varphi^*\left(-\frac{y}{2}+b\right)\right] \tag{61}$$

In this expression, the functions φ are zero if their argument's modulus is larger than a. As an example, $\varphi\left(\frac{y}{2}-b\right)$ is different from zero only if $y \simeq 2b$, whereas $\varphi^*\left(-\frac{y}{2}-b\right)$ is different from zero only if $y \simeq -2b$; consequently their product is always zero. Actually, in the product of the two brackets, only the "crossed" terms are non-zero, and we obtain (with our assumption that the function φ is even):

$$\rho_W^x(x=0,p_x) = \frac{1}{4\pi\hbar}\int \mathrm{d}y\; e^{-ip_x y/\hbar}\left[e^{-i\chi}\left|\varphi\left(b-\frac{y}{2}\right)\right|^2+e^{i\chi}\left|\varphi\left(b+\frac{y}{2}\right)\right|^2\right] \tag{62}$$

Changing the sign of the integration variable y for the second term in the bracket, we can write:

$$\rho_W^x(x=0,p_x) = \frac{1}{2\pi\hbar}\int \mathrm{d}y\;\cos\left(\frac{p_x y}{\hbar}+\chi\right)\left|\varphi\left(b-\frac{y}{2}\right)\right|^2 \tag{63}$$

In the limit where the width a becomes very narrow, the squared modulus of the wave function φ in the integral behaves as a delta function $\delta(b-y/2)$, and we get:

$$\rho_W^x(x=0,p_x) \simeq \frac{1}{\pi\hbar}\cos\left(\frac{2bp_x}{\hbar}+\chi\right) \tag{64}$$

This result illustrates two properties of the Wigner distributions that both seem quite surprising. The first is that the distribution is non-zero at point $x = 0$, whereas the probability of finding the particle at this position is strictly zero. The second is that the distribution is an oscillating function of momentum, taking successively positive and negative values, whereas a classical distribution always remain positive or zero. These two properties are actually related: integrating the distribution over all possible momenta p_x yields zero, which is in agreement with relation (18) stating that the integral of the Wigner distribution over the momenta yields the probability of the particle's presence at each point. More details on the properties of a two-peak wave function will be given in § 5-a.

3. Wigner transform of an operator

Consider now any operator A acting in the particle state space. We define its Wigner transform $A_W(\mathbf{r},\mathbf{p})$ in the same way as for a density operator, but without the prefactor $1/(2\pi\hbar)^3$ that appears in front of the integrals in (17) and (21):

$$\begin{aligned} A_W(\mathbf{r},\mathbf{p}) &= \int \mathrm{d}^3y\; e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\left\langle \mathbf{r}+\frac{\mathbf{y}}{2}\right| A \left|\mathbf{r}-\frac{\mathbf{y}}{2}\right\rangle \\ &= \int \mathrm{d}^3q\; e^{i\mathbf{q}\cdot\mathbf{r}/\hbar}\left\langle \mathbf{p}+\frac{\mathbf{q}}{2}\right| A \left|\mathbf{p}-\frac{\mathbf{q}}{2}\right\rangle \end{aligned} \tag{65}$$

To simplify the notation, this definition does not include a time dependence; one can, however, directly replace A by $A(t)$ and $A_W(\mathbf{r},\mathbf{p})$ by $A_W(\mathbf{r},\mathbf{p};t)$, without any other change. The inversion relations (33) and (34) now become:

$$\begin{aligned}\langle \mathbf{r}_1 | A | \mathbf{r}_2 \rangle &= \frac{1}{(2\pi\hbar)^3} \int \mathrm{d}^3 p \; e^{i\mathbf{p}\cdot(\mathbf{r}_1-\mathbf{r}_2)/\hbar} \; A_W\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2},\mathbf{p}\right) \\ \langle \mathbf{p}_1 | A | \mathbf{p}_2 \rangle &= \frac{1}{(2\pi\hbar)^3} \int \mathrm{d}^3 r \; e^{i(\mathbf{p}_2-\mathbf{p}_1)\cdot\mathbf{r}/\hbar} \; A_W\left(\mathbf{r},\frac{\mathbf{p}_1+\mathbf{p}_2}{2}\right)\end{aligned} \tag{66}$$

Taking the complex conjugate of relation (65) shows that the Wigner transform of a Hermitian operator A is necessarily a real function. Similarly, the fact that the complex conjugate of (66) is real indicates that it is a sufficient condition for hermiticity.

As the prefactor $1/(2\pi\hbar)^3$ is no longer included in the definition (65), the equivalent of relation (38) is now:

$$A = \int \mathrm{d}^3 r \int \mathrm{d}^3 p \; A_W(\mathbf{r},\mathbf{p}) \; W(\mathbf{r},\mathbf{p}) \tag{67}$$

We saw previously that the operator $W(\mathbf{r},\mathbf{p})$ is Hermitian. The above relation then allows building a Hermitian operator A from any real function $A_W(\mathbf{r},\mathbf{p})$ of position and momentum. In other words, we found a quantization procedure for any classical function, often called "Weyl quantization" or "phase space quantization" [100, 101, 102]. Starting from two functions $A_W(\mathbf{r},\mathbf{p})$ and $B_W(\mathbf{r},\mathbf{p})$, whose product obviously commutes, this procedure yields two operators A and B that, in general, do not commute. Such an operation, which introduces in phase space a non-commutative structure, is sometimes referred to as "geometric quantization".

3-a. Average value of a Hermitian operator (observable)

We can now compute the average value of operator A in the quantum state defined by the density operator $\rho(t)$:

$$\langle A \rangle = \mathrm{Tr}\{\rho(t)\,A\} = \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2 \; \langle \mathbf{r}_1 | \rho(t) | \mathbf{r}_2 \rangle \langle \mathbf{r}_2 | A | \mathbf{r}_1 \rangle \tag{68}$$

We are going to show that:

$$\boxed{\langle A \rangle = \int \mathrm{d}^3 r \int \mathrm{d}^3 p \; \rho_W(\mathbf{r},\mathbf{p};t) \; A_W(\mathbf{r},\mathbf{p})} \tag{69}$$

This relation is the exact analog of the relation one would obtain with a classical distribution. It is the reason the Wigner transform of the density operator is referred to as a "quasi-classical distribution", or more simply as a "distribution".

Demonstration:

Inserting in (68) the equalities (33) and (66) leads to:

$$\begin{aligned}\langle A \rangle = \frac{1}{(2\pi\hbar)^3} \int \mathrm{d}^3 r_1 \int \mathrm{d}^3 r_2 \int \mathrm{d}^3 p \; e^{i\mathbf{p}\cdot(\mathbf{r}_1-\mathbf{r}_2)/\hbar} \int \mathrm{d}^3 p' \; e^{-i\mathbf{p}'\cdot(\mathbf{r}_1-\mathbf{r}_2)/\hbar} \\ \times \; \rho_W\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2},\mathbf{p};t\right) A_W\left(\frac{\mathbf{r}_1+\mathbf{r}_2}{2},\mathbf{p}'\right)\end{aligned} \tag{70}$$

We now replace the integration variables $\mathbf{r}_1$ and $\mathbf{r}_2$ by the following variables:

$$\mathbf{r} = \frac{\mathbf{r}_1 + \mathbf{r}_2}{2} \qquad ; \qquad \mathbf{r}' = \mathbf{r}_1 - \mathbf{r}_2 \tag{71}$$

The summation over $\mathrm{d}^3 r'$ introduces a delta function:

$$(2\pi)^3 \ \delta\left(\frac{\mathbf{p} - \mathbf{p}'}{\hbar}\right) = (2\pi\hbar)^3 \ \delta\left(\mathbf{p} - \mathbf{p}'\right) \tag{72}$$

which takes care of the integration over $\mathrm{d}^3 p'$. We then finally obtain (69).

3-b. Special cases

In the special case in which the operator A depends only on the position operator:

$$A = F(\mathbf{R}) \qquad \text{and hence:} \qquad \langle \mathbf{r}_1 | A | \mathbf{r}_2 \rangle = F(\mathbf{r}_1) \ \delta(\mathbf{r}_1 - \mathbf{r}_2) \tag{73}$$

the first line of (65) leads to:

$$A_W(\mathbf{r}, \mathbf{p}) = F(\mathbf{r}) \tag{74}$$

The Wigner transform of the operator is then simply the function $F(\mathbf{r})$, which does not depend on the momentum $\mathbf{p}$.

In a similar way, if A depends only on the momentum operator:

$$A = G(\mathbf{P}) \qquad \text{and hence:} \qquad \langle \mathbf{p}_1 | A | \mathbf{p}_2 \rangle = G(\mathbf{p}_1) \ \delta(\mathbf{p}_1 - \mathbf{p}_2) \tag{75}$$

the second line of (65) leads to:

$$A_W(\mathbf{r}, \mathbf{p}) = G(\mathbf{p}) \tag{76}$$

As a further illustration, let us find an operator A whose Wigner transform involves both position and momentum, for example:

$$A_W(\mathbf{r}, \mathbf{p}) = \mathbf{r} \cdot \mathbf{p} \tag{77}$$

Relation (66) yields its matrix elements:

$$\begin{aligned} \langle \mathbf{r}_1 | A | \mathbf{r}_2 \rangle &= \frac{1}{(2\pi\hbar)^3} \int \mathrm{d}^3 p \ e^{i\mathbf{p}\cdot(\mathbf{r}_1 - \mathbf{r}_2)/\hbar} \ \frac{\mathbf{r}_1 + \mathbf{r}_2}{2} \cdot \mathbf{p} \\ &= \frac{\hbar}{i} \frac{\mathbf{r}_1 + \mathbf{r}_2}{2} \cdot \boldsymbol{\nabla}_{\mathbf{r}_1} \delta(\mathbf{r}_1 - \mathbf{r}_2) \end{aligned} \tag{78}$$

We recognize in this expression the matrix elements of the operator $\mathbf{P}$, equal to the gradient of a delta function of the positions, multiplied by $\hbar/i$. Note, in addition, that $\mathbf{r}_1$ is the result of the action of the position operator on the bra, whereas $\mathbf{r}_2$ is the result of the action of the position operator on the ket. This means that:

$$A = \frac{1}{2} [\mathbf{R} \cdot \mathbf{P} + \mathbf{P} \cdot \mathbf{R}] \tag{79}$$

We thus get a Hermitian operator, as expected since its Wigner transform is real. It is however remarkable that building a quantum operator via the Wigner transforms spontaneously introduces an arrangement of the operators' order leading to the necessary symmetry. This property is quite general: starting from real classical functions, the Wigner transform allows building operators symmetrized with respect to position and momentum. This method is a real quantization procedure.

3-c. Wigner transform of an operator product

We are going to show that, in general, the Wigner transform associated with the product of two operators A and B is not simply the product of the Wigner transforms of each operator.

α. General expression

Let us apply relation (65) to obtain the Wigner transform of a product of two operators A and B. Inserting a closure relation on the kets $|\mathbf{z}\rangle$ leads to:

$$[AB]_W(\mathbf{r},\mathbf{p}) = \int \mathrm{d}^3y\; e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar} \int \mathrm{d}^3z\; \left\langle \mathbf{r}+\frac{\mathbf{y}}{2}\right| A \left|\mathbf{z}\right\rangle \left\langle \mathbf{z}\right| B \left|\mathbf{r}-\frac{\mathbf{y}}{2}\right\rangle \tag{80}$$

We can then replace the matrix elements of A and B by their expressions (66), which leads to:

$$\begin{aligned}&[AB]_W(\mathbf{r},\mathbf{p})\\ &= \frac{1}{(2\pi\hbar)^6}\int \mathrm{d}^3y \int \mathrm{d}^3z\; e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar} \int \mathrm{d}^3p_1 \int \mathrm{d}^3p_2\; e^{i\mathbf{p}_1\cdot\left(\mathbf{r}+\frac{\mathbf{y}}{2}-\mathbf{z}\right)/\hbar} e^{i\mathbf{p}_2\cdot\left(-\mathbf{r}+\frac{\mathbf{y}}{2}+\mathbf{z}\right)/\hbar}\\ &\qquad\qquad \times A_W\left(\frac{\mathbf{r}+\mathbf{z}}{2}+\frac{\mathbf{y}}{4},\mathbf{p}_1\right) B_W\left(\frac{\mathbf{r}+\mathbf{z}}{2}-\frac{\mathbf{y}}{4},\mathbf{p}_2\right)\end{aligned} \tag{81}$$

Instead of using the position representation, one can use the momentum representation; we then must use the relations on the second lines of (65) and (66). A reasoning similar to that used before leads to:

$$\begin{aligned}&[AB]_W(\mathbf{r},\mathbf{p})\\ &= \frac{1}{(2\pi\hbar)^6}\int \mathrm{d}^3q\; e^{i\mathbf{q}\cdot\mathbf{r}/\hbar} \int \mathrm{d}^3q' \int \mathrm{d}^3x \int \mathrm{d}^3y\; e^{i\left(\mathbf{q}'-\mathbf{p}-\frac{\mathbf{q}}{2}\right)\cdot\mathbf{x}/\hbar} e^{i\left(\mathbf{p}-\frac{\mathbf{q}}{2}-\mathbf{q}'\right)\cdot\mathbf{y}/\hbar}\\ &\qquad\qquad \times A_W\left(\mathbf{x},\frac{\mathbf{q}'+\mathbf{p}}{2}+\frac{\mathbf{q}}{4}\right) B_W\left(\mathbf{y},\frac{\mathbf{q}'+\mathbf{p}}{2}-\frac{\mathbf{q}}{4}\right)\end{aligned} \tag{82}$$

Depending on the case, it will be easier to use either (81) or (82). These two expressions are exact, but fairly complicated. They can be simplified, however, in a certain number of cases.

β. A few simple cases

As a first example, imagine that operator A is simply the position operator $\mathbf{R}$ while B can be any operator. As A_W is no longer dependent on $\mathbf{p}_1$, the integration over $\mathrm{d}^3p_1/(2\pi\hbar)^3$ in (81) yields a delta function $\delta\left(\mathbf{r}+\frac{\mathbf{y}}{2}-\mathbf{z}\right)$; this allows integrating over d^3z to obtain:

$$[\mathbf{R}B]_W(\mathbf{r},\mathbf{p}) = \frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3y\; e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar} \int \mathrm{d}^3p_2\; e^{i\mathbf{p}_2\cdot\mathbf{y}/\hbar} \times \left[\mathbf{r}+\frac{\mathbf{y}}{2}\right] B_W(\mathbf{r},\mathbf{p}_2) \tag{83}$$

For the term in $\mathbf{r}$, the integral over d^3y of exponential $e^{i(\mathbf{p}_2-\mathbf{p})\cdot\mathbf{y}/\hbar}$ introduces a function $\mathbf{r}\delta(\mathbf{p}_2-\mathbf{p})$ with the coefficient $(2\pi\hbar)^3$. As for the term in $\mathbf{y}/2$, it yields $\boldsymbol{\nabla}_{\mathbf{p}_2}\delta(\mathbf{p}_2-\mathbf{p})$,

with the coefficient $(\hbar/2i)\,(2\pi\hbar)^3$. After integrating over d^3p_2, we get:

$$[\mathbf{R}B]_W(\mathbf{r},\mathbf{p}) = \mathbf{r}\ B_W(\mathbf{r},\mathbf{p}) - \frac{\hbar}{2i}\boldsymbol{\nabla}_{\mathbf{p}}B_W(\mathbf{r},\mathbf{p}) \tag{84}$$

If we now reverse the order of the operators $\mathbf{R}$ and B, the roles of $\mathbf{p}_1$ and $\mathbf{p}_2$ are interchanged in (81); the integration over $\mathrm{d}^3p_2/\,(2\pi\hbar)^3$ yields a function $\delta\left(-\mathbf{r}+\frac{\mathbf{y}}{2}+\mathbf{z}\right)$ and the integration over d^3z leads to:

$$[B\mathbf{R}]_W(\mathbf{r},\mathbf{p}) = \frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3y\ e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\int \mathrm{d}^3p_1\ e^{i\mathbf{p}_1\cdot\mathbf{y}/\hbar}\times B_W(\mathbf{r},\mathbf{p}_1)\left[\mathbf{r}-\frac{\mathbf{y}}{2}\right] \tag{85}$$

Compared to (83), the only change is the sign of $\mathbf{y}$ in the final bracket, so that we simply obtain the final result by changing the sign of the gradient on the right-hand side of (84). This means that the Wigner transform of the commutator is:

$$[\mathbf{R},B]_W(\mathbf{r},\mathbf{p}) = -\frac{\hbar}{i}\boldsymbol{\nabla}_{\mathbf{p}}B_W(\mathbf{r},\mathbf{p}) \tag{86}$$

Starting from (82), the same reasoning leads to:

$$[\mathbf{P}B]_W(\mathbf{r},\mathbf{p}) = \mathbf{p}\ B_W(\mathbf{r},\mathbf{p}) + \frac{\hbar}{i}\boldsymbol{\nabla}_{\mathbf{r}}B_W(\mathbf{r},\mathbf{p}) \tag{87}$$

This relation can now be iterated to obtain:

$$\left[\mathbf{P}^2B\right]_W(\mathbf{r},\mathbf{p}) = \mathbf{p}^2\ B_W(\mathbf{r},\mathbf{p}) + 2\frac{\hbar}{i}\mathbf{p}\cdot\boldsymbol{\nabla}_{\mathbf{r}}B_W(\mathbf{r},\mathbf{p}) - \hbar^2\Delta_{\mathbf{r}}B_W(\mathbf{r},\mathbf{p}) \tag{88}$$

We then get the expression for the Wigner transform of the commutator of the momentum squared and any operator B:

$$\left[\mathbf{P}^2,B\right]_W(\mathbf{r},\mathbf{p}) = \frac{2\hbar}{i}\ \mathbf{p}\cdot\boldsymbol{\nabla}_{\mathbf{r}}B_W(\mathbf{r},\mathbf{p}) \tag{89}$$

This relation will be useful for what follows.

γ. *Gradient expansions*

We now show that relation (81) can be expressed as a series expansion of higher order derivatives of the two functions A_W and B_W, of the form:

$$[AB]_W(\mathbf{r},\mathbf{p}) = A_W(\mathbf{r},\mathbf{p})\times B_W(\mathbf{r},\mathbf{p}) + \frac{\hbar}{2i}\left\{B_W(\mathbf{r},\mathbf{p}),A_W(\mathbf{r},\mathbf{p})\right\} + ... \tag{90}$$

where we have used the classical definition of the "*Poisson bracket*" [103, 104] of classical Lagrangian mechanics:

$$\begin{aligned}&\left\{B_W(\mathbf{r},\mathbf{p}),A_W(\mathbf{r},\mathbf{p})\right\}\\ &\quad= \boldsymbol{\nabla}_{\mathbf{r}}B_W(\mathbf{r},\mathbf{p})\cdot\boldsymbol{\nabla}_{\mathbf{p}}A_W(\mathbf{r},\mathbf{p}) - \boldsymbol{\nabla}_{\mathbf{r}}A_W(\mathbf{r},\mathbf{p})\cdot\boldsymbol{\nabla}_{\mathbf{p}}B_W(\mathbf{r},\mathbf{p})\end{aligned} \tag{91}$$

This shows that, to lowest order in $\hbar$, the Wigner function of an operator product is simply the product of the Wigner transforms of these operators. To first order, a correction must be added, which contains the Poisson bracket of the two Wigner transforms. It

is remarkable that purely quantum considerations bring in this classical Poisson bracket definition; this explains why these results are well suited for the study of the classical limit of quantum mechanics.

In (90), the expansion is limited to the contribution of the first order derivatives of the two functions. The following terms involve higher order derivatives and, consequently, higher powers of $\hbar$ (the corresponding result is called the "Groenewold's formula"; see for example [99]).

Demonstration:

Let us make in (81) the following change of momentum integration variables:

$$\begin{aligned} \mathbf{P} &= \frac{\mathbf{p}_1 + \mathbf{p}_2}{2} \\ \mathbf{q} &= \mathbf{p}_1 - \mathbf{p}_2 \end{aligned} \tag{92}$$

(despite the notation with a capital letter, $\mathbf{P}$ is a classical variable, not an operator). This leads to the new expression:

$$\begin{aligned} [AB]_W(\mathbf{r},\mathbf{p}) = \frac{1}{(2\pi\hbar)^6} \int \mathrm{d}^3y \int \mathrm{d}^3z \int \mathrm{d}^3P \int \mathrm{d}^3q \; e^{i(\mathbf{P}-\mathbf{p})\cdot\mathbf{y}/\hbar} e^{i\mathbf{q}\cdot(\mathbf{r}-\mathbf{z})/\hbar} \\ \times A_W\left(\frac{\mathbf{r}+\mathbf{z}}{2}+\frac{\mathbf{y}}{4},\mathbf{P}+\frac{\mathbf{q}}{2}\right) B_W\left(\frac{\mathbf{r}+\mathbf{z}}{2}-\frac{\mathbf{y}}{4},\mathbf{P}-\frac{\mathbf{q}}{2}\right) \end{aligned} \tag{93}$$

If the two Wigner transforms A_W and B_W vary slowly with position and momentum, we can use the expansions:

$$\begin{aligned} A_W\left(\frac{\mathbf{r}+\mathbf{z}}{2}+\frac{\mathbf{y}}{4},\mathbf{P}+\frac{\mathbf{q}}{2}\right) &= A_W\left(\frac{\mathbf{r}+\mathbf{z}}{2},\mathbf{P}\right) + \frac{\mathbf{y}}{4}\cdot\boldsymbol{\nabla}_{\mathbf{r}} A_W + \frac{\mathbf{q}}{2}\cdot\boldsymbol{\nabla}_{\mathbf{p}} A_W + ... \\ B_W\left(\frac{\mathbf{r}+\mathbf{z}}{2}-\frac{\mathbf{y}}{4},\mathbf{P}-\frac{\mathbf{q}}{2}\right) &= B_W\left(\frac{\mathbf{r}+\mathbf{z}}{2},\mathbf{P}\right) - \frac{\mathbf{y}}{4}\cdot\boldsymbol{\nabla}_{\mathbf{r}} B_W - \frac{\mathbf{q}}{2}\cdot\boldsymbol{\nabla}_{\mathbf{p}} B_W + ... \end{aligned} \tag{94}$$

Keeping only the first term in each of these two expansions (zero-order term in the gradient expansion), the integrals over d^3y and d^3q introduce the delta functions $\delta(\mathbf{P}-\mathbf{p})$ and $\delta(\mathbf{r}-\mathbf{z})$ respectively, each with a coefficient $(2\pi\hbar)^3$. We then get:

$$[AB]_W(\mathbf{r},\mathbf{p}) = A_W(\mathbf{r},\mathbf{p}) \times B_W(\mathbf{r},\mathbf{p}) + ... \tag{95}$$

In this approximation, the Wigner transform of the product of two operators is thus simply the product of the Wigner transforms.

We now take into account the first order terms in the gradient expansion (94). The $\boldsymbol{\nabla}_{\mathbf{r}} A_W$ term on the first line contains a summation over d^3y modified by the presence of $\mathbf{y}$ in the integral:

$$\frac{1}{(2\pi\hbar)^3} \int \mathrm{d}^3y \; e^{i(\mathbf{P}-\mathbf{p})\cdot\mathbf{y}/\hbar} \; \mathbf{y} = \frac{\hbar}{i}\boldsymbol{\nabla}_{\mathbf{P}}\delta(\mathbf{P}-\mathbf{p}) \tag{96}$$

The integral over d^3P in (93) is now modified and leads to a derivation with respect to $\mathbf{P}$ of the function to be integrated, a multiplication by the coefficient $-\hbar/i$, and finally the replacement of $\mathbf{P}$ by $\mathbf{p}$. On the other hand, the integral over $\mathrm{d}^3q/(2\pi\hbar)^3$ is unchanged and leads to the replacement of $\mathbf{z}$ by $\mathbf{r}$. The corresponding term is therefore written:

$$-\frac{\hbar}{4i}\boldsymbol{\nabla}_{\mathbf{p}} \cdot [B_W(\mathbf{r},\mathbf{p})\,\boldsymbol{\nabla}_{\mathbf{r}} A_W(\mathbf{r},\mathbf{p})] \tag{97}$$

As for the $\boldsymbol{\nabla}_{\mathbf{p}} A_W$ term on the first line of (94), it can be handled in the same way. The presence of the variable $\mathbf{q}$ transforms $\delta(\mathbf{r}-\mathbf{z})$ into $\boldsymbol{\nabla}_{\mathbf{z}}\delta(\mathbf{r}-\mathbf{z})$, with a coefficient $-\hbar/i$, where the sign change of this coefficient comes from the $-\mathbf{z}$ in the exponent $\mathbf{q}\cdot(\mathbf{r}-\mathbf{z})$; the integral over d^3P is unchanged. This yields the term:

$$\frac{\hbar}{4i}\boldsymbol{\nabla}_{\mathbf{r}}\cdot\left[B_W(\mathbf{r},\mathbf{p})\,\boldsymbol{\nabla}_{\mathbf{p}}A_W(\mathbf{r},\mathbf{p})\right] \tag{98}$$

which, added to (97), leads to the contribution (the terms involving a double derivation of A_W cancel each other):

$$\frac{\hbar}{4i}\left[\boldsymbol{\nabla}_{\mathbf{r}}B_W(\mathbf{r},\mathbf{p})\cdot\boldsymbol{\nabla}_{\mathbf{p}}A_W(\mathbf{r},\mathbf{p})-\boldsymbol{\nabla}_{\mathbf{r}}A_W(\mathbf{r},\mathbf{p})\cdot\boldsymbol{\nabla}_{\mathbf{p}}B_W(\mathbf{r},\mathbf{p})\right] \tag{99}$$

Finally, the terms coming from the second line of (94) are obtained by exchanging the roles of A_W and B_W, and changing the signs because of the opposite values of $\mathbf{y}$ and $\mathbf{q}$ in the of relation (94). We thus double the result (99), and finally obtain expression (90) to first order in the gradients.

3-d. Evolution of the density operator

The Schrödinger evolution of the density operator obeys the von Neumann equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\rho(t)=[H(t),\ \rho(t)] \tag{100}$$

Taking its Wigner transformation, this equation becomes:

$$i\hbar\frac{\partial}{\partial t}\rho_W(\mathbf{r},\mathbf{p};t)=\frac{1}{(2\pi\hbar)^3}\quad[H,\rho]_W(\mathbf{r},\mathbf{p};t) \tag{101}$$

where, on the right-hand side, is written the Wigner transform associated with the commutator of $H(t)$ and $\rho(t)$; the factor $1/(2\pi\hbar)^3$ comes from the definition of the Wigner distribution of the density operator, remembering that no such coefficient appears in the transform of an arbitrary operator. We already saw that the general expression of the Wigner transform of an operator product is somewhat complex, and the same is of course true for their commutator.

α. Classical limit

If we only keep, as in (90), the first order terms in the gradients, we see that the zero-order terms disappear, and that the terms in $H(t)\,\rho(t)$ and $\rho(t)\,H(t)$ double up; in addition, factors $\hbar$ on each side of the equations cancel out. Using this approximation, we get:

$$\frac{\partial}{\partial t}\rho_W(\mathbf{r},\mathbf{p};t)=\{H_W(\mathbf{r},\mathbf{p};t),\rho_W(\mathbf{r},\mathbf{p};t)\}+\hbar... \tag{102}$$

where the Poisson bracket of $H_W(\mathbf{r},\mathbf{p}_1;t)$ and $\rho_W(\mathbf{r},\mathbf{p}_1;t)$ is defined in (91). As noticed earlier in § 3-c-γ, the neglected terms are proportional to $\hbar$, and vanish in the classical limit $\hbar\to 0$. We find in this limit, where the gradients of the Wigner transforms with respect to position and momentum are small, the usual equations of classical dynamics.

β. Particle in an external potential

An exact calculation can be made if the particle's Hamiltonian is simply the sum of a kinetic energy and an external potential energy:

$$H = \frac{\mathbf{P}^2}{2m} + V(\mathbf{R};t) \tag{103}$$

where m is the mass of the particle. The contribution of the kinetic energy to the right-hand side of (101) comes directly from relation (89):

$$\left.\frac{\partial}{\partial t}\right|_{\text{kinetic}} \rho_W(\mathbf{r},\mathbf{p};t) = -\frac{\mathbf{p}}{m}\cdot\boldsymbol{\nabla}_{\mathbf{r}}\rho_W(\mathbf{r},\mathbf{p};t) \tag{104}$$

The evolution of the Wigner distribution induced by the kinetic energy operator is thus given by a "drift term" just as in classical physics.

As for the contribution of the potential energy, the computation is very similar to the one conducted at the beginning of § 3-c-β, except that instead of dealing with the operator $\mathbf{R}$ itself, we are now dealing with a function $V(\mathbf{R})$ of that operator. Taking $B = \rho$ in relations (83) and (85), they become:

$$[V(\mathbf{R})\,\rho]_W(\mathbf{r},\mathbf{p}) = \frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3y\; e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\int \mathrm{d}^3p_2\; e^{i\mathbf{p}_2\cdot\mathbf{y}}\times V\left(\mathbf{r}+\frac{\mathbf{y}}{2};t\right)\rho_W(\mathbf{r},\mathbf{p}_2;t) \tag{105}$$

and:

$$[\rho V(\mathbf{R})]_W(\mathbf{r},\mathbf{p}) = \frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3y\; e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\int \mathrm{d}^3p_1\; e^{i\mathbf{p}_1\cdot\mathbf{y}}\times \rho_W(\mathbf{r},\mathbf{p}_1;t)\,V\left(\mathbf{r}-\frac{\mathbf{y}}{2};t\right) \tag{106}$$

Finally, the evolution of the Wigner distribution $\rho_W(\mathbf{r},\mathbf{p};t)$ obeys the following equation:

$$\frac{\partial}{\partial t}\rho_W(\mathbf{r},\mathbf{p};t) + \frac{\mathbf{p}}{m}\cdot\boldsymbol{\nabla}_{\mathbf{r}}\rho_W(\mathbf{r},\mathbf{p};t) = \frac{1}{i\hbar}\frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3y\int \mathrm{d}^3p'\; e^{i(\mathbf{p}'-\mathbf{p})\cdot\mathbf{y}/\hbar} \times\left[V\left(\mathbf{r}+\frac{\mathbf{y}}{2};t\right) - V\left(\mathbf{r}-\frac{\mathbf{y}}{2};t\right)\right]\rho_W(\mathbf{r},\mathbf{p}';t) \tag{107}$$

This is an exact equation. It contains all the quantum effects that play a role in the particle's evolution. It obeys a local conservation law for the probability:

$$\frac{\partial}{\partial t}n(\mathbf{r},t) + \boldsymbol{\nabla}_{\mathbf{r}}\cdot\mathbf{J}(\mathbf{r},t) = 0 \tag{108}$$

where the local probability density $n(\mathbf{r},t)$ is defined in (18), and its associated current $\mathbf{J}(\mathbf{r},t)$ is defined as:

$$\mathbf{J}(\mathbf{r},t) = \int \mathrm{d}^3p\;\frac{\mathbf{p}}{m}\;\rho_W(\mathbf{r},\mathbf{p};t) \tag{109}$$

This can be shown by integrating (107) over d^3p, as the left-hand side then becomes identical to the left-hand side of (108), just as in classical mechanics; as for the right-hand side, the integration over d^3p introduces a function $\delta(\mathbf{y})$ that cancels the bracket in the remaining integral.

When the external potential varies slowly enough, one can use in (107) the following approximation:

$$V\left(\mathbf{r}+\frac{\mathbf{y}}{2};t\right)-V\left(\mathbf{r}-\frac{\mathbf{y}}{2};t\right)=\mathbf{y}\cdot\boldsymbol{\nabla}_{\mathbf{r}}V(\mathbf{r};t)+... \tag{110}$$

The integration over $\mathrm{d}^3y/(2\pi\hbar)^3$ then leads to a function $(\hbar/i)\nabla_{\boldsymbol{p}'}\delta(\mathbf{p}'-\mathbf{p})$ and we get:

$$\frac{\partial}{\partial t}\rho_W(\mathbf{r},\mathbf{p};t)+\frac{\mathbf{p}}{m}\cdot\boldsymbol{\nabla}_{\mathbf{r}}\rho_W(\mathbf{r},\mathbf{p};t)=\boldsymbol{\nabla}_{\mathbf{r}}V(\mathbf{r};t)\cdot\boldsymbol{\nabla}_{\mathbf{p}}\rho_W(\mathbf{r},\mathbf{p};t)+... \tag{111}$$

One recognizes here the Liouville equation of classical mechanics. The dots at the end of the equation symbolize the possible contributions of terms containing higher order spatial derivatives of the potential $V(\mathbf{r};t)$. They come with a power of $\hbar$ increasing with the order of the derivative. This means that they correspond to quantum corrections: the faster the potential varies in space, the more terms need to be taken into account. On the other hand, when the potential varies slowly, only keeping the classical evolution term is a good approximation.

4. Generalizations

The above considerations can be directly generalized to particles with spin, or to an N-particle system.

4-a. Particle with spin

For a particle with spin, a basis in state space is formed by the kets $|\mathbf{r},\nu\rangle$, where $\mathbf{r}$ is the eigenvalue of the position operator, and ν the eigenvalue of the spin component on the quantization axis. The matrix elements of the density operator are then written:

$$\langle\mathbf{r},\nu|\,\rho(t)\,|\mathbf{r}',\nu'\rangle \tag{112}$$

For each value of ν and ν' we can perform a Wigner transformation and define, as in (17), the functions:

$$\rho_W^{\nu\nu'}(\mathbf{r},\mathbf{p};t)=\frac{1}{(2\pi\hbar)^3}\int\mathrm{d}^3y\;e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\;\left\langle\mathbf{r}+\frac{\mathbf{y}}{2},\nu\right|\rho(t)\left|\mathbf{r}-\frac{\mathbf{y}}{2},\nu'\right\rangle \tag{113}$$

As an example, for a spin 1/2 the two indices ν and ν' can take on two different values, noted $\pm$. We thus define four Wigner functions, which can be arranged in a 2×2 spin matrix:

$$\begin{pmatrix}\rho_W^{++}(\mathbf{r},\mathbf{p};t) & \rho_W^{+-}(\mathbf{r},\mathbf{p};t)\\ \rho_W^{-+}(\mathbf{r},\mathbf{p};t) & \rho_W^{--}(\mathbf{r},\mathbf{p};t)\end{pmatrix} \tag{114}$$

It is easy to show that this matrix is Hermitian:

$$\left[\rho_W^{+-}(\mathbf{r},\mathbf{p};t)\right]^*=\rho_W^{-+}(\mathbf{r},\mathbf{p};t) \tag{115}$$

Such a matrix is frequently used when studying the quantum properties of spin polarization transport in fluids (spin waves for example).

4-b. Several particles

For two spinless particles, relation (17) is easily generalized to:

$$\rho_W(\mathbf{r}_1,\mathbf{p}_1;\mathbf{r}_2,\mathbf{p}_2;t) = \frac{1}{(2\pi\hbar)^6}\int \mathrm{d}^3y_1 \int \mathrm{d}^3y_2\; e^{-i\mathbf{p}_1\cdot\mathbf{y}_1/\hbar}e^{-i\mathbf{p}_2\cdot\mathbf{y}_2/\hbar} \times \left\langle \mathbf{r}_1+\frac{\mathbf{y}_1}{2},\mathbf{r}_2+\frac{\mathbf{y}_2}{2}\middle|\,\rho(t)\,\middle|\mathbf{r}_1-\frac{\mathbf{y}_1}{2},\mathbf{r}_2-\frac{\mathbf{y}_2}{2}\right\rangle \tag{116}$$

Actually, any number of particles can be treated this way. Including the spin can be done as in the previous section, but it rapidly leads to a great number of Wigner functions (4^N for N particles each having a spin 1/2).

The Wigner distribution for a system including a large particle number N therefore depends on $6N$ variables when the particles have no spin; when the particles have a spin 1/2, it is no longer a single distribution that must be studied, but rather 4^N distributions which are the matrix elements of a spin operator. In practice, one usually uses the Wigner distribution of the one-particle density operator, resulting from the partial trace over the $N-1$ other particles, or sometimes the Wigner distribution of the two-particle density operator.

5. Discussion: Wigner distribution and quantum effects

Knowledge of the Wigner distribution allows computing the average values of observables, as seen from relation (69). It can be used to obtain the probability of any measurement result, since this probability is simply the average value of the projector onto the eigensubspace associated with this result. We simply have to compute the Wigner transform of this projector, multiply it by $\rho_W(\mathbf{r},\mathbf{p};t)$, and integrate the result over the two variables. From a practical point of view, all the information is contained in $\rho_W(\mathbf{r},\mathbf{p};t)$. However, and as already underlined with the examples given in § 2-f, that does not mean we should attribute too much physical content to the Wigner distribution itself. Strictly speaking, the Wigner distribution is rather a useful and powerful computation tool than a direct representation of the physical properties of the system.

To highlight the behavior of the Wigner transform in a situations where quantum effects are predominant, we now study an interference experiment.

5-a. An interference experiment

When the wave function of a particle goes through a screen pierced with two holes, it is split into two coherent wave packets propagating in space, and interfering when they overlap. Figure 1 represents these two wave packets after the screen, as they both propagate towards the region R where they will interfere. As they propagate in free space, the Wigner distribution associated with the particle simply obeys relation (104), which is just a classical equation of motion. What causes the interference effects in region R? To answer this question, we shall use relation (19), or its equivalent (26), which allow computing the Wigner transform associated with the particle's wave function.

This wave function is now the sum of two components, $\Psi_1(\mathbf{r},t)$ for the wave packet emerging from the first hole in the screen, and $\Psi_2(\mathbf{r},t)$ for the wave packet emerging from the second hole:

$$\Psi(\mathbf{r},t) = \Psi_1(\mathbf{r},t) + \Psi_2(\mathbf{r},t) \tag{117}$$

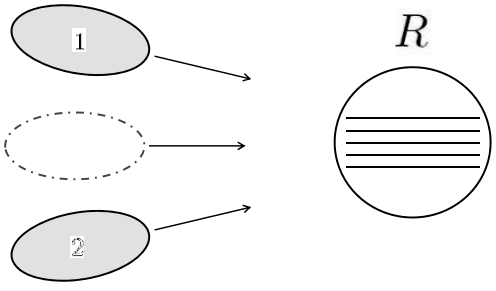


Figure 1: The wave function of a quantum particle can be split into two coherent components 1 *and* 2*, after passing, for example, through a screen pieced with two holes, or through an interferometer. As long as the two wave packets do not overlap, the Wigner distribution is the sum of three components, schematically drawn in ordinary space in the figure: a first one localized with wave packet* 1*, a second with wave packet* 2*, and finally a third one (circled with dashed lines) remaining at mid-distance from the two wave packets. This third component is called the "ghost component": when measuring its position, the particle can never be found in this component. The value of the ghost Wigner distribution oscillates rapidly as a function of the momentum* $\boldsymbol{p}$*.*
Later on, as the two wave packets 1 *and* 2 *overlap, the three components are different from zero in the same region of space; in addition, the momentum oscillations of the ghost component slow down and even vanish. This component now plays an essential role: as it is added to the terms* 1 *and* 2*, it is responsible for introducing the density oscillations producing the fringe pattern (schematized as horizontal lines in region R). It plays a virtual role as long as the wave packets are well separated, but an essential one when they overlap, as it leads to quantum interference effects.*

A similar situation has already been studied in § 2-f. Inserting (117) into relation (19), which is quadratic in Ψ, four contributions will come into play:

$$\rho_W(\mathbf{r},\mathbf{p};t) = \rho_W^1(\mathbf{r},\mathbf{p};t) + \rho_W^2(\mathbf{r},\mathbf{p};t) + \rho_W^{1,2}(\mathbf{r},\mathbf{p};t) + \rho_W^{2,1}(\mathbf{r},\mathbf{p};t) \tag{118}$$

In this equality, $\rho_W^1(\mathbf{r},\mathbf{p};t)$ is obtained when we replace in (19) the functions $\Psi(\mathbf{r},t)$ and $\Psi^*(\mathbf{r},t)$ by $\Psi_1(\mathbf{r},t)$ and $\Psi_1^*(\mathbf{r},t)$ respectively. The contribution $\rho_W^2(\mathbf{r},\mathbf{p};t)$ is obtained by replacing them by $\Psi_2(\mathbf{r},t)$ and $\Psi_2^*(\mathbf{r},t)$ respectively. Finally, the "crossed" contributions $\rho_W^{1,2}(\mathbf{r},\mathbf{p};t)$ and $\rho_W^{2,1}(\mathbf{r},\mathbf{p};t)$ come from replacing $\Psi(\mathbf{r},t)$ by $\Psi_1(\mathbf{r},t)$ and $\Psi^*(\mathbf{r},t)$ by $\Psi_2^*(\mathbf{r},t)$, and conversely. For example, relation (19) leads to:

$$\rho_W^{1,2}(\mathbf{r},\mathbf{p};t) = \frac{1}{(2\pi\hbar)^3}\int \mathrm{d}^3y\; e^{-i\mathbf{p}\cdot\mathbf{y}/\hbar}\;\Psi_1\left(\mathbf{r}+\frac{\mathbf{y}}{2};t\right)\Psi_2^*\left(\mathbf{r}-\frac{\mathbf{y}}{2};t\right) \tag{119}$$

whereas the equivalent relation (26) yields another expression as a function of the Fourier transforms $\overline{\Psi}_1$ and $\overline{\Psi}_2^*$. It can easily be shown that the two distributions $\rho_W^{1,2}(\mathbf{r},\mathbf{p};t)$ and $\rho_W^{2,1}(\mathbf{r},\mathbf{p};t)$ are complex conjugates of each other. Their sum is real, as is, consequently, $\rho_W(\mathbf{r},\mathbf{p};t)$.

As an example, imagine that the two wave packets are Gaussian, as were the wave packets studied in § 2-d. We saw in Complement G_I shows that a Gaussian wave packet, as it propagates in free space, remains Gaussian at all times; its momentum dispersion remains constant, while its spatial width changes with time. For the sake of simplicity, we shall consider a one-dimensional problem and will not explicitly write the time dependence. We assume one of the wave packet to be centered at $+x_0$, and the other at $-x_0$. Relation (41) then leads to:

$$\begin{aligned}\overline{\psi}_1(p_x) &= \frac{1}{(2\pi)^{1/4}}\sqrt{\frac{a}{2\hbar}}\; e^{-a^2(p_x-p_0)^2/4\hbar^2}\; e^{-ip_x x_0/\hbar} \\ \overline{\psi}_2(p_x) &= \frac{1}{(2\pi)^{1/4}}\sqrt{\frac{a}{2\hbar}}\; e^{-a^2(p_x-p_0)^2/4\hbar^2}\; e^{ip_x x_0/\hbar}\end{aligned} \tag{120}$$

(a factor $1/\sqrt{2}$ has been added to ensure the normalization of the total wave function; we assume $a \ll x_0$, so that the spatial overlap of the two wave packets is negligible, and the squared norm of the sum is the sum of the squared norms). The same computation as in § 2-d then yields:

$$\begin{aligned}\rho_W^1(x,p_x) &= \frac{1}{2\pi\hbar}e^{-\frac{a^2}{2\hbar^2}(p_x-p_0)^2}e^{-2\frac{(x-x_0)^2}{a^2}} \\ \rho_W^2(x,p_x) &= \frac{1}{2\pi\hbar}e^{-\frac{a^2}{2\hbar^2}(p_x-p_0)^2}e^{-2\frac{(x+x_0)^2}{a^2}}\end{aligned} \tag{121}$$

As for the crossed contributions, the computation is slightly different. Since the two lines of relation (120) have different signs in front of x_0, the product $\overline{\psi}_1(p_x+q/2)\,\overline{\psi}_2^*(p_x-q/2)$ contains the exponential $\mathrm{e}^{-2ipx_0/\hbar}$, whereas the product $\overline{\psi}_1^*(p_x+q/2)\,\overline{\psi}_2(p_x-q/2)$ contains $\mathrm{e}^{2ipx_0/\hbar}$. The computation of (42) then becomes:

$$\begin{aligned}&\rho_W^{1,2}(x,p_x)+\rho_W^{2,1}(x,p_x) \\ &= \frac{1}{2}\frac{a}{(2\pi)^{3/2}\hbar^2}\int \mathrm{d}q\; e^{-\frac{a^2}{4\hbar^2}\left[\left(p-p_0+\frac{q}{2}\right)^2+\left(p-p_0-\frac{q}{2}\right)^2\right]}\; e^{iqx/\hbar}\left[e^{-2ipx_0/\hbar}+e^{2ipx_0/\hbar}\right] \\ &= \frac{a}{(2\pi)^{3/2}\hbar^2}\cos\frac{2px_0}{\hbar}e^{-\frac{a^2}{2\hbar^2}(p-p_0)^2}\int \mathrm{d}q\; e^{-\frac{a^2q^2}{8\hbar^2}}\; e^{i\frac{qx}{\hbar}}\end{aligned} \tag{122}$$

or else:

$$\rho_W^{1,2}(x,p_x)+\rho_W^{2,1}(x,p_x) = \cos\frac{2px_0}{\hbar}e^{-\frac{a^2}{2\hbar^2}(p-p_0)^2}\; e^{-\frac{2x^2}{a^2}} \tag{123}$$

Finally, the total Wigner transform is:

$$\rho_W(x,p_x) = \frac{1}{2\pi\hbar}e^{-\frac{a^2}{2\hbar^2}(p_x-p_0)^2}\left[e^{-2\frac{(x-x_0)^2}{a^2}}+e^{-2\frac{(x+x_0)^2}{a^2}}+2\cos\frac{2p_xx_0}{\hbar}\; e^{-\frac{2x^2}{a^2}}\right] \tag{124}$$

The first two terms in the bracket are easy to understand: they are simply half the sum of the Wigner transforms associated with each of the wave packet. Each of these two terms is centered on the wave packet it corresponds to, that is at $x = \pm x_0$. The third term is the crossed term, which corresponds to an interference between the two wave packets, and is centered at $x = 0$, half way between them. In addition, this term oscillates as a function of p_x with a frequency proportional to the distance between the two wave packets.

5-b. General discussion; "ghost" component

The distribution $\rho_W^1(\mathbf{r},\mathbf{p};t)$ propagates as if it were the distribution of a free particle described by the single wave packet $\Psi_1(\mathbf{r},t)$; the distribution $\rho_W^2(\mathbf{r},\mathbf{p};t)$ corresponds to the second wave packet, here again as if it were isolated. If these were the only contributions, when the two wave packets overlap these two Wigner distributions would simply add to each other, since they follow a classical evolution; no quantum interference effects would result from this addition.

However, we saw that in (118) we must also include crossed terms (interference terms) whose properties are radically different from the first two terms. A first significant difference comes from their oscillations as a function of momentum, which necessarily involves positive and negative values of the distribution. This is definitely a quantum effect since a classical distribution must always be positive or zero. Another difference is that this crossed term in the Wigner transform propagates in a region of space where the wave function is zero, and consequently cannot correspond to any probability of the particle's presence; the integral over momentum of the last term on the right-hand side of (124) is indeed zero (in the limit $x_0 \gg a$ of well separated wave packets corresponding to the assumption made for our computation). The sum $\rho_W^{1,2}(x,p_x)+\rho_W^{2,1}(x,p_x)$ is sometimes called the "ghost component" of the Wigner distribution (or sometimes, in quantum optics, the "tamasic component"); when measuring the particle's position, it can never be found in this component[4]. Its value is always real, but not necessarily positive, because of its oscillations.

This means that, as long as the two wave packets 1 and 2 are well separated, the Wigner transform associated with the particle is the sum of three independent components: two components separately associated with each wave packet and propagating with them; one "ghost component", also propagating but remaining at mid-distance from the two wave packets. However, when the wave packets meet in region R, the three components of the Wigner transform overlap in space. The ghost component, which has a changing sign, combines with the other two components to modulate the particle's probability of presence, hence producing the interference pattern predicted by quantum mechanics. In a certain sense, one can say that the ghost component carries the quantum effects associated with the particle.

Conclusion

Quantum mechanics and classical mechanics are two very different theories. It was not obvious that, using the Wigner transforms, one could write the quantum equations in a form so akin to the classical equations of a distribution in phase space. Furthermore, we showed that any real classical function of position and momentum could be used in this formalism to generate a Hermitian operator acting in state space. In the limit where $\hbar \to 0$, the quantum equations of motion lead to the same Poisson brackets as the classical equations; quantum and classical theories then show strong similarities. Quantum effects, however, can manifest themselves in several ways:

- the evolution of the Wigner distribution can be significantly different from the classical evolution when the potentials vary rapidly on a scale on the order of $\hbar/p$ (de

[4]It is also known as the "empty component" stressing the fact that this component contains no particle.

Broglie wavelength), as higher order terms in the gradient expansion become essential.

- the Wigner transform is not always positive. We saw an example of this with the ghost component in an interference experiment, which, in a manner of speaking, carries the quantum effects to the usual components.

- whereas in classical physics any distribution in phase space, as long as it is positive and normalized, can be accepted, this is no longer the case in quantum mechanics. The only acceptable Wigner distributions are those which correspond to a density operator that is positive definite, a condition that is not expressed simply in terms of the distribution.

The Wigner transformation is frequently used in quantum physics. We already mentioned that it was introduced in 1932, while studying quantum corrections to thermal equilibrium [98]. It probably plays an even more important role in the study of transport properties where Boltzmann type equations contain simultaneous information on particles' positions and momenta. Furthermore, the Wigner transform is also useful for understanding and characterizing quantum effects, as its negativity in certain regions of phase space is a sensitive indicator of the existence of such effects. One can even use the Wigner transforms to introduce a "phase space formulation of quantum mechanics" [100, 101], totally equivalent to the usual formalism in terms of state space and operators, and which is a real quantization procedure. In a general way, the Wigner transformation belongs to the class of the so-called Liouville formulations of quantum mechanics [105], which have many uses.

Finally, there are many domains of physics (such as signal processing, in particular) in which the Wigner transformation is part of a larger class of mixed time-frequency transformations. Numerous types of such transformations exist (such as sliding window or envelope transforms, wavelets, etc.) chosen to best fit the problem at hand. Even in quantum mechanics there are other quasi-classical transforms, beside the Wigner transform, as for example the Husimi or the Kirkwood transforms, or the Glauber transform expressed in terms of creation and annihilation operators of the electromagnetic field; a review on that subject can be found in [99]. The Wigner transform still remains one of the most useful transforms, allowing, in particular, analytical calculations for many interesting cases.

Bibliography of volume III

[1] C. Cohen-Tannoudji, B. Diu and F. Laloë, *Quantum mechanics, Volume I*, Wiley (1977).

[2] C. Cohen-Tannoudji, B. Diu and F. Laloë, *Quantum mechanics, Volume II*, Wiley (1977). 1591

[3] R.K. Pathria, *Statistical mechanics*, Pergamon press (1972). 2296

[4] E.J. Mueller, Tin-Lun Ho, M. Ueda and G. Baym, "Fragmentation of Bose-Einstein condensates", *Phys. Rev.* **A 74**, 033612 (2006). 1656

[5] J.P. Blaizot et G. Ripka, *Quantum theory of finite systems*, the MIT Press (1986). 1678, 1701, 1809

[6] Wikipedia, "Density functional theory", https://en.wikipedia.org/wiki/Density_functional_theory 1699

[7] L.P. Kadanoff et G. Baym, *Quantum statistical mechanics*, Benjamin (1976). 1798

[8] A.J. Leggett, *Quantum liquids*, Oxford University Press, 2006. 1820, 1890, 1926

[9] J. Bardeen, L.N. Cooper and J.R. Schrieffer, "Theory of superconductivity", *Phys. Rev.* **108**, 1175-1204 (1957). 1889

[10] W. Ketterle and M. Zwierlein, "Making, probing and understanding ultracold Fermi gases" in *Proceedings of the international school of physics Enrico Fermi, Course CLXIV, Varenna*, edited by M. Inguscio, W. Ketterle and C. Salomon, IOS Press (Amsterdam), 2008; arXiv:0801.2500v1. 1926

[11] W. Zwerger, "The BCS-BEC crossover and the unitary Fermi gas", Springer, 2012. 1926

[12] M. Tinkham, "Introduction to superconductivity", Dover books on physics, 2004. 1926

[13] R.D. Parks, "Superconductivity", Volume 1 and 2, Dekker, 1969. 1926

[14] M. Combescot and S-Y Shiau, "Excitons and Cooper pairs", Oxford Univedrsity Press, 2016. 1926

[15] L. Pitaevskii and S. Stringari, "Bose-Einstein condensation and superfluidity", Oxford University Press, 2016. 1944

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

[16] C. Cohen-Tannoudji, J. Dupont-Roc and G. Grynberg, *Photons and atoms, introduction to quantum electrodynamics*, Wiley (1997). 1960, 1963, 1968, 1980, 1990, 2007, 2008, 2014, 2053, 2063

[17] T.A. Welton, "Some observable effects of the quantum-mechanical fluctuations of the electromagnetic field", *Phys. Rev.* **74**, 1157-1167 (1948). 2008

[18] J.V. Prodan, W. D. Phillips and H. Metcalf, "Laser production of a very slow monoenergetic atomic beam", *Phys. Rev. Lett.* **49**, 1149-1153 (1982). 2025

[19] T.W. Hänsch and A. Schawlow, "Cooling of gases by laser radiation" *Opt. Comm.* **13**, 68-69 (1975). 2026

[20] D.J. Wineland and H. Dehmelt, "Proposed $10^{14}\delta\nu < \nu$ laser fluorescence spectroscopy on Tl^+ mono-ion oscillator III (sideband cooling)", *Bull. Am. Phys. Soc.* **20**, 637 (1975). 2026

[21] C. Cohen-Tannoudji, J. Dupont-Roc and G. Grynberg, *Atom-Photon Interactions. Basic Processes and Applications*, Wiley-Interscience (1992). 2028, 2107, 2117, 2130, 2136, 2145

[22] Special issue on laser cooling and trapping, *JOSA B, Optical Physics*, **6**, Number 11 (1989). 2034

[23] W. Ketterle and N.L Van Druten, "Evaporative cooling of trapped atoms", *Advances in atomic, molecular, and optical physics*, **37**, 181-236 (1996). 2034

[24] C. Cohen-Tannoudji and D. Guéry-Odelin, *Advances in atomic physics. An Overview*, World Scientific, Singapour (2011). 1779, 2034, 2036, 2039, 2065, 2127, 2152, 2153

[25] W.E. Lamb, "Capture of neutrons by atoms in a crystal", *Phys. Rev.* **55**, 190-197 (1939). 2040

[26] R.V. Pound and G. A. Rebka Jr., "Apparent weight of photons", *Phys. Rev. Lett.* **4**, 337-341 (1960). 2040

[27] L.S. Vasilenko,V. P. Chebotayev and A. V. Shishaev, "Line shape of two-photon absorption in a standing-wave field in a gas", *JETP Lett.* **12**, 113-116 (1970). 2041

[28] B. Cagnac, G. Grynberg and F. Biraben, "Spectroscopie d'absorption multiphotonique sans effet Doppler", *J. Phys. (Paris)* **34**, 845-858 (1973). 2041

[29] T.W. Hänsch, "Passion for precision", *Rev. Mod. Phys.* **78**, 1297-1309 (2006). 2041

[30] A. Kastler, "Projet d'expérience sur le moment cinétique de la lumière", Société des Sciences physiques et naturelles de Bordeaux, Jan. 28 (1932). 2052

[31] R.A. Beth, "Mechanical Detection and Measurement of the Angular Momentum of Light", *Phys. Rev.* **50**, 115-125 (1936). 2052

[32] J.W. Simmons and M.J. Guttmann, *States, waves and photons: a modern introduction to light*, Addison-Wesley (1970), Chap. 9. 2052

[33] A.M. Yao and M. J. Padgett, "Orbital angular momentum: origins, behavior and applications", *Advances in Optics and Photonics*, IOP Publishing **3**, 161-204 (2011). 2052

[34] J.D. Jackson, *Classical electrodynamics, 3rd ed.*, Wiley (1999). 2053, 2063

[35] J. Brossel et A. Kastler, "La détection de la résonance magnétique des niveaux excités ", *C. R. Acad. Sci.* **229**, 1213 (1949). 2059

[36] J. Brossel, and F. Bitter, "A new 'Double Resonance' method for investigating atomic energy levels. Application to Hg $^3P_1$", *Phys. Rev.* **86**, 308 (1952). 2059, 2061

[37] J.N. Dodd, W. N. Fox, G. W. Series and M. J. Taylor, "Light beats as indicators of structure of atomic energy levels", *Proc. Phys. Soc.* **74**, 789 (1959). 2061

[38] E. Majorana, "Atomi orientati in campo magnetico variabile", *Nuovo Cimento* **9**, 43-50 (1932). 2061

[39] A. Kastler, "Quelques suggestions concernant la production optique et la détection optique d'une inégalité de population des niveaux de quantifigation spatiale des atomes. Application à l'expérience de Stern et Gerlach et à la résonance magnétique", *J. Phys. Radium* **11**, 255-265 (1950). 2062

[40] N. F. Ramsey, *Molecular beams*, Oxford University Press (1956). 2064

[41] L. Allen, S.M. Barnett and M.J. Padgett, *Optical angular momentum*, IOP Publishing (2003). 2065

[42] M.F. Andersen, C. Ryu, P. Cladé, V. Natarajan, A. Vaziri, K. Helmerson, and W.D. Phillips, "Quantized rotation of atoms from photons with orbital angular momentum" *Phys. Rev. Lett.* **97**, 170406 (2006). 2066

[43] A. Einstein, "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt", *Annalen der Physik* **17**, 132-149 (1905). 2110

[44] R.A. Millikan, "On the elementary electric charge and the Avogadro constant", *Physical Review* **2**, 109-143 (1913). 2111

[45] W.E. Lamb and M.O. Scully, "The photoelectric effect without photons", in *Polarisation, Matière et Rayonnement*, (Presses Universitaires de France), Jubilee volume in honour of Alfred Kastler, p. 363 (1969). 2110

[46] E. Hanbury Brown and R.Q. Twiss, "A test of a new type of stellar interferometer on Sirius", *Nature*, **178**, 1046-1048 (1956). 2120

[47] P. Grangier, G. Roger and A. Aspect, "Experimental evidence for a photon anticorrelation effect on a beam splitter: a new light on single-photon interferences", *Europhys. Lett.* **1**, 173-179 (1986). 2121

[48] J.T. Höffges, H.W. Baldauf, W. Lange and H. Walther "Heterodyne measurements of the resonance fluorescence of a single ion", *J. of Mod. Optics* **44**, 1999-2010 (1997). 2121, 2122

[49] H.J. Kimble, M. Dagenais and L. Mandel, "Photon antibunching in resonance fluorescence", *Phys. Rev. Lett.* **39**, 691 (1977). 2121

[50] S. Pancharatnam, "Light shifts in semiclassical dispersion theory", *J. Opt. Soc. Am.* **56**, 1636 (1966). 2139

[51] C. Cohen-Tannoudji, "Théorie quantique du cycle de pompage optique. Vérification expérimentale des nouveaux effets prévus", *Ann. Phys.* **13**, 423-461 and 469-504(1962). 2140

[52] C. Cohen-Tannoudji and J. Dupont-Roc, "Experimental study of light shifts in weak magnetic fields", *Phys. Rev. A* **5**, 968-984 (1972). 2140

[53] C. Cohen-Tannoudji, "Observation d'un déplacement de raie de résonance magnétique causé par l'excitation optique", *C. R. Acad. Sci.* **252**, 394-396 (1961). 2140, 2141

[54] B.R. Mollow, "Power spectrum of light scattered by two-level systems", *Phys. Rev.* **188**, 1969-1975 (1969). 2144

[55] S.H. Autler and C. H. Townes, "Stark effect in rapidly varying fields", *Phys. Rev.* **100**, 703-722 (1955). 2144

[56] S. Chu, J. Bjorkholm, A. Ashkin, and A. Cable, "Experimental observation of optically trapped atoms", *Phys. Rev. Lett.* **57**, 314-317 (1986). 2152

[57] R.J. Cook and R.K. Hill, "An electromagnetic mirror for neutral atoms", *Opt. Commun.* **43**, 258-260 (1982). 2153

[58] M.Greiner, O. Mandel, T. Esslinger, T.W. Hänsch and I. Bloch, I. *Nature* **415**, "Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms", 39-44 (2002). 2154

[59] M. Ben Dahan, E. Peik, J. Reichel, Y. Castin and C. Salomon, "Bloch oscillations of atoms in an optical potential", *Phys. Rev. Lett.* **76**, 4508-4511 (1996). 2155

[60] P.D. Lett, R.N. Watts, C.I. Westbrook, W.D. Phillips, P.L. Gould and H.J. Metcalf, "Observation of atoms laser coooled below the Doppler limit", *Phys. Rev. Lett.* **61**, 169-172 (1988). 2155

[61] J. Dalibard and C. Cohen-Tannoudji, "Laser cooling below the Doppler limit by polarization gradients: simple theoretical models", *J. Opt. Soc. Am. B* **6**, 2023-2045 (1989). 2159

[62] P.J. Ungar, D.S.Weiss, E. Riis and S. Chu, "Optical molasses and multilevel atoms: theory", *J. Opt. Soc. Am. B* **6**, 2058-2071 (1989). 2159

[63] C. Salomon, J. Dalibard, W.D. Phillips, A. Clairon and S. Guellati, *Europhys. Lett.*, "Laser cooling of Cesium atoms below 3 μK", **12**, 683-688 (1990). 2159

[64] S. Gleyzes, S. Kuhr, C. Guerlin, J. Bernu, S. Deléglise, U.B. Hoff, M. Brune, J-M. Raimond and S. Haroche, "Quantum jumps of light recording the birth and death of a photon in a cavity", *Nature* **446**, 297-300 (2007). 2160

[65] G. Grynberg, A. Aspect et C. Fabre, *Introduction to quantum optics*, with contributions from F. Bretenaker and A. Browaeys, Cambridge University Press (2010). 2186

[66] D.F. Walls and G.J. Milburn, *Quantum optics*, Springer (1994). 2186

[67] E. Schrödinger, "Discussion of probability relations between separated systems", *Proc. Cambridge Phil. Soc.* **31**, 555 (1935); "Probability relations between separated systems", *Proc. Cambridge Phil. Soc.* **32**, 446 (1936). 2190

[68] F. Laloë, *Do we really understand quantum mechanics?* , Cambridge University Press (2012); second expanded edition (2019). 2202, 2213, 2330

[69] A. Einstein, B. Podolsky and N. Rosen, "Can quantum-mechanical description of physical reality be considered complete?", *Phys. Rev.* **47**, 777–780 (1935); *Quantum Theory of Measurement*, J.A. Wheeler and W.H. Zurek eds., Princeton University Press (1983), pp. 138–141. 2204, 2207

[70] A. Einstein, "Quantenmechanik und Wirklichkeit", *Dialectica* **2**, 320–324 (1948). 2207

[71] N. Bohr, "Can quantum-mechanical description of physical reality be considered complete?", *Phys. Rev.* **48**, 696–702 (1935). 2207

[72] N. Bohr, "On the notions of causality and complementarity", *Dialectica* **2**, 312–319 (1948); Science, New Series **111**, 51-54 (Jan20, 1950). 2207

[73] J.S. Bell, "On the problem of hidden variables in quantum mechanics", *Rev. Mod. Phys.* **38**, 447–452 (1966); reprinted in *Quantum Theory and Measurement*, J.A. Wheeler and W.H. Zurek editors, Princeton University Press (1983), 396–402 and in chapter 1 of [74]. 2208

[74] J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press (1987); second augmented edition (2004), which contains the complete set of J. Bell's articles on quantum mechanics. 2329

[75] A. Peres, "Unperformed experiments have no results", *Am. J. Phys.* **46**, 745–747 (1978). 2212

[76] J.A. Wheeler, "Niels Bohr in today's words" in *Quantum Theory and Measurement*, J.A. Wheeler and W.H. Zurek editors, Princeton University Press (1983), pp. 182–213. 2212

[77] A list of references describing Bell's experiments can be found in: A. Aspect, "Closing the door on Einstein and Bohr's quantum debate", *Physics* **8**, 123 (2015). 2212

[78] D. Mermin, *Quantum computer science*, Cambridge University Press (2007). 2212

[79] N. Gisin, G. Ribordy, W. Tittel and H. Zbinden, "Quantum cryptography", *Rev. Mod. Phys.* **74**, 145-195 (2002). 2213

[80] V. Coffman, J. Kundu and W.K. Wootters, "Distributed entanglement", *Phys. Rev.* **A 61**, 052306 (2000). 2223

[81] R.F. Werner, "Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden variable model", *Phys. Rev.* **A 40**, 4277-4281 (1989). 2224

[82] A. Peres, "Separability criterion for density matrices", *Phys. Rev. Lett.* **77**, 1413-1415 (1996). 2224

[83] D.M. Greenberger, M.A. Horne and A. Zeilinger, "Going beyond Bell's theorem", pp. 69-72 in *Bell's theorem, quantum theory and conceptions of the Universe*, M. Kafatos (ed.), Kluwer Academic Publishers (1989). 2227

[84] D.M. Greenberger, M.A. Horne, A. Shimony, and A. Zeilinger, "Bell's theorem without inequalities", *Am. J. Phys.* **58**, 1131-1143 (1990). 2227

[85] D. Bouwmeester, J.W. Pan, M. Daniell, H. Weinfurter, and A. Zeilinger, "Observation of three-photon Greenberger-Horne-Zeilinger entanglement", *Phys. Rev. Lett.* **82**, 1345-1349 (1999). 2231

[86] see for example § 5-1-2 of [68]. 2231

[87] J.W. Pan, D. Bouwmeester, H. Weinfurter, and A. Zeilinger, "Experimental entanglement swapping: entangling photons that never interacted", *Phys. Rev. Lett.* **80**, 3891-3894 (1998). 2235

[88] B. Hensen, H. Bernien, A.E. Dréau, A. Reiserer, N. Kalb, M.S. Blok, J. Ruitenberg, R.F. Vermeulen, R.N. Schouten, C. Abellan, W. Amaya, V. Pruneri, M.W. Mitchell, M. Markham, D.J. Twitchen, D. Elkouss, S. Wehner, T.H. Taminiau and R. Hanson, "Experimental loophole-free violation of a Bell inequality using electron spins separated by 1.3 km", *Nature* **526**, 682-686 (2015). 2235

[89] M.R. Andrews, C.G. Townsend, H.J. Miesner, D.S. Durfee, D.M. Kurn and W. Ketterle, "Observation of interference between two Bose condensates", *Science* **275**, 637-641 (1997). 2238

[90] F. Laloë, "The hidden phase of Fock states, quantum non-local effects", *Eur. Phys. J. D* **33**, 87-97 (2005); F. Laloë and W.J. Mullin, "Non-local quantum effects with Bose-Einstein condensates", *Phys. Rev. Lett.* **99**, 150401 (2007). 2262, 2263, 2265

[91] R.P. Feynman and A.R. Hibbs, *Quantum mechanics and path integrals*, McGraw Hill (1965). 2267

[92] J. Zinn-Justin, *Intégrale de chemin en mécanique quantique: Introduction*, CNRS Editions et EDP Sciences (2003). 2267, 2280

[93] M. Le Bellac, *Physique quantique*, EDP Sciences et CNRS Editions (2013), tome II, chapitre 12. 2267

[94] D.M. Ceperley, "Path integrals in the theory of condensed helium", *Rev. Modern Physics*, **67**, 279-355 (1995). 2280

[95] B. Diu, C. Guthmann, D. Lederer et B. Roulet, *Physique statistique*, Hermann (1989). 2296

[96] K. Huang, *Statistical mechanics*, Wiley (1963). 2296

[97] F. Reif, *Fundamental of statistical and thermal physics*, McGraw-Hill (1965). 2296

[98] E. Wigner, "On the quantum correction for thermodynamic equilibrium", *Phys. Rev.* **40**, 749-759 (1932). 2298, 2323

[99] M. Hillery, R.F. O'Connell, M.O. Scully and E.P. Wigner, "Distribution functions in physics; fundamentals", *Physics Reports*, **106**, 121-167 (1984). 2298, 2315, 2323

[100] A. Perelomov, "Generalized coherent states and their applications", Springer (1986); see in particular Chap. 16. 2311, 2323

[101] C.K. Zachos, D. Fairlie and T.L. Cutright, *Quantum mechanics in phase space*, World Scientific, Singapore (2005); with the same title, *Asia Pacific Newsletters*, **01**, 37-46 (2012) or ArXiv:1104.5269v2. 2311, 2323

[102] G.G. Athanasiu and E.G. Fioratos, "Coherent states in finite quantum mechanics", *Nuclear Physics* **B 425**, 343-364 (1994). 2311

[103] L. Landau and E. Lifchitz, *Mechanics, Course of theoretical physics, Vol. I*, C § 42, Pergamon Press (1960) and Elsevier Butterworth-Heinemann (1976). 2314

[104] H. Goldstein, C.P. Poole et J.L. Safko, *Classical mechanics*, Addison-Wesley (2001). 2314

[105] N.L. Balazs and B.K. Jennings, "Wigner's function and other distribution functions in mock phase spaces", *Physics Reports*, **104**, 347-391 (1984). 2323

Absorption
and emission of photons, 2073
collision with, 971
of a quantum, a photon, 1311, 1353
of field, 2149
of several photons, 1368
rates, 1334
Acceptor (electron acceptor), 1495
Acetylene (molecule), 878
Action, 341, 1539, 1980
Addition
of angular momenta, 1015, 1043
of spherical harmonics, 1059
of two spins 1/2, 1019
Adiabatic
branching of the potential, 932
Adjoint
matrix, 123
operator, 112
Algebra (commutators), 165
Allowed energy band, 381, 1481, 1491
Ammonia (molecule), 469, 873
Amplitude
scattering amplitude, 929, 953
Angle (quantum), 2258
Angular momentum
addition of momenta, 1015, 1043
and rotations, 717
classical, 1529
commutation relations, 669, 725
conservation, 668, 736, 1016
coupling, 1016
electromagnetic field, 1968, 2043
half-integral, 987
of identical particles, 1497(ex.)
of photons, 1370
orbital, 667, 669, 685
quantization, 394
quantum, 667
spin, 987, 991
standard representation, 677, 691
two coupled momenta, 1091
Anharmonic oscillator, 502, 1135
Annihilation operator, 504, 513, 514, 1597
Annihilation-creation (pair), 1831, 1878
Anomalous
average value, 1828, 1852
dispersion, 2149
Zeeman effect, 987
Anti-normal correlation function, 1782, 1789
Anti-resonant term, 1312
Anti-Stokes (Raman line), 532, 752
Antibunching (photon), 2121
Anticommutation, 1599
field operator, 1754
Anticrossing of levels, 415, 482
Antisymmetric ket, state, 1428, 1431
Antisymmetrizer, 1428, 1431
Applications of the perturbation theory, 1231
Approximation
central field approximation, 1459
secular approximation, 1374
Argument (EPR), 2205
Atom(s), *see* helium, hydrogenoid
donor, 837
dressed, 2129, 2133
many-electron atoms, 1459, 1467
mirrors for atoms, 2153
muonic atom, 541
single atom fluorescence, 2121
Atomic
beam (deceleration), 2025
orbital, 869, 1496(ex.)
parameters, 41
Attractive bosons, 1747
Autler-Townes
doublet, 2144
effect, 1410
Autoionization, 1468
Average value (anomalous), 1828
Azimuthal
quantum number, 811

Band (energy), 381
Bardeen-Cooper-Schrieffer, 1889
Barrier (potential barrier), 68, 367, 373
Basis

Quantum Mechanics, Volume III, First Edition. C. Cohen-Tannoudji, B. Diu, and F. Laloë.
© 2020 Wiley-VCH Verlag GmbH & Co. KGaA. Published 2020 by Wiley-VCH Verlag GmbH & Co. KGaA.

change of bases, 174
characteristic relations, 101, 119
continuous basis in the space of states, 99
mixed basis in the space of states, 99
BCHSH inequalities, 2209, 2210
BCS, 1889
broken pairs and excited pairs, 1920
coherent length, 1909
distribution functions, 1899
elementary excitations, 1923
excited states, 1919
gap, 1894, 1896, 1923
pairs (wave function of), 1901
phase locking, 1893, 1914, 1916
physical mechanism, 1914
two-particle distribution, 1901
Bell's
inequality, 2208
theorem, 2204, 2208
Benzene (molecule), 417, 495
Bessel
Bessel-Parseval relation, 1507
spherical Bessel function, 944
spherical equation, 961
spherical function, 966
Biorthonormal decomposition, 2194
Bitter, 2059
Blackbody radiation, 651
Bloch
equations, 463, 1358, 1361
theorem, 659
Bogolubov
excitations, 1661
Hamiltonian, 1952
operator method, 1950
phonons, spectrum, 1660
transformation, 1950
Bogolubov-Valatin transformation, 1836, 1919
Bohr, 2207
electronic magneton, 856
frequencies, 249
magneton, *see* front cover pages
model, 40, 819
nuclear magneton, 1237
radius, 820
Boltzmann
constant, *see* front cover pages
distribution, 1630
Born
approximation, 938, 977, 1320
Born-Oppenheimer approximation, 528, 1177, 1190
Born-von Karman conditions, 1490
Bose-Einstein
condensation, 1446, 1638, 1940
condensation (repulsive bosons), 1933
condensation of pairs, 1857
distribution, 652, 1630
statistics, 1446
Bosons, 1434
at non-zero temperature, 1745
attractive, 1747
attractive instability, 1745
condensed, 1638
in a Fock state, 1775
paired, 1881
Boundary conditions (periodic), 1489
Bra, 103, 104, 119
Bragg reflection, 382
Brillouin
formula, 452
zone, 614
Broadband
detector, 2165
optical excitation, 1332
Broadening (radiative), 2138
Broken pairs and excited pairs (BCS), 1920
Brossel, 2059
Bunching of bosons, 1777

C.S.C.O., 133, 137, 153, 236
Canonical
commutation relations, 142, 223, 1984
ensemble, 2289
Hamilton-Jacobi canonical equations, 214
Hamilton-Jacobi equations, 1532
Cauchy principal part, 1517
Center of mass, 812, 1528
Center of mass frame, 814
Central

field approximation, 1459
potential, 1533
Central potential, 803, 841
scattering, 941
stationary states, 804
Centrifugal potential, 809, 888, 893
Chain (von Neumann), 2201
Chain of coupled harmonic oscillators, 611
Change
of bases, 124, 174, 1601
of representation, 124
Characteristic equation, 129
Characteristic relation of an orthonormal basis, 116
Charged harmonic oscillator in an electric field, 575
Charged particle
in an electromagnetic field, 1536
Charged particle in a magnetic field, 240, 321, 771
Chemical bond, 417, 869, 1189, 1210
Chemical potential, 1486, 2287
Circular quanta, 761, 783
Classical
electrodynamics, 1957
histories, 2272
Clebsch-Gordan coefficients, 1038, 1051
Closure relation, 93, 117
Coefficients
Clebsch-Gordan, 1038
Einstein, 1334, 2083
Coherences (of the density matrix), 307
Coherent length (BCS), 1909
Coherent state (field), 2008
Coherent superposition of states, 253, 301, 307
Collision, 923
between identical particles, 1454, 1497(ex.)
between identical particles in classical mechanics, 1420
between two identical particles, 1450
cross section, 926
scattering states, 928
total scattering cross section, 926
with absorption, 971
Combination
of atomic orbitals, 1172
Commutation, 1599
canonical relations, 142, 223
field operator, 1754
of pair field operators, 1861
relations, 1984
Commutation relations
angular momentum, 669, 725
field, 1989, 1996
Commutator algebra, 165
Commutator(s), 91, 167, 171, 187
of functions of operators, 168
Compatibility of observables, 232
Complementarity, 45
Complete set of commuting observables (C.S.C.O.), 133, 137, 236
Complex variables (Lagrangian), 1982
Compton wavelength of the electron, 825, 1235
Condensates
relative phase, 2237
with spins, 2254
Condensation
BCS condensation energy, 1917
Bose-Einstein, 1446, 1857, 1933
Condensed bosons, 1638
Conduction band, 1492
Conductivity (solid), 1492
Configurations, 1467
Conjugate momentum, 214, 323, 1531, 1983, 1987, 1995
Conjugation (Hermitian), 111
Conservation
local conservation of probability, 238
of angular momentum, 668, 736, 1016
of energy, 248
of probability, 237
Conservative systems, 245, 315
Constants of the motion, 248, 317
Contact term, 1273
Contact term (Fermi), 1238, 1247
Contextuality, 2231
Continuous
spectrum, 133, 219, 264, 1316
variables (in a Lagrangian), 1984
Continuum of final states, 1316, 1378, 1380
Contractions, 1802

Convolution product of two functions, 1510
Cooling
Doppler, 2026
down atoms, 2025
evaporative, 2034
Sisyphus, 2034
sub-Doppler, 2155
subrecoil, 2034
Cooper model, 1927
Cooper pairs, 1927
Cooperative effects (BCS), 1916
Correlation functions, 1781, 1804
anti-normal, 1782, 1789
dipole and field, 2113
for one-photon processes, 2084
normal, 1782, 1787
of the field, spatial, 1758
Correlations, 2231
between two dipoles, 1157
between two physical systems, 296
classical and quantum, 2221
introduced by a collision, 1104
Coulomb
field, 1962
gauge, 1965
Coulomb potential
cross section, 979
Coupling
between angular momenta, 1016
between two angular momenta, 1091
between two states, 412
effect on the eigenvalues, 438
spin-orbit coupling, 1234, 1241
Creation and annihilation operators, 504, 513, 514, 1596, 1990
Creation operator (pair of particles), 1813, 1846
Critical velocity, 1671
Cross section
and phase shifts, 951
scattering cross section, 926, 933, 953, 972
Current
metastable current in superfluid, 1667
of particles, 1758
of probability, 240
probability current in hydrogen atom, 851
Cylindrical symmetry, 899(ex.)

Darwin term, 1235, 1279
De Broglie
relation, 10
wavelength, *see* front cover pages, 11, 35
Decay of a discrete state, 1378
Deceleration of an atomic beam, 2025
Decoherence, 2199
Decomposition (Schmidt), 2193
Decoupling (fine or hyperfine structure), 1262, 1291
Degeneracy
essential, 811, 825, 845
exchange degeneracy, 1423
exchange degeneracy removal, 1435
lifted by a perturbation, 1125
rotation invariance, 1072
systematic and accidental, 203
Degenerate eigenvalue, 127, 203, 217, 260
Degereracy
lifted by a perturbation, 1117
parity, 199
Delta Dirac function, 1515
potential well and barriers, 83–85(ex.)
use in quantum mechanics, 97, 106, 280
Density
Lagrangian, 1986
of probability, 264
of states, 389, 1316, 1484, 1488
operator, 449, 1391
operator and matrix, 299
particle density operator, 1756
Density functions
one and two-particle, 1502(ex.)
Depletion (quantum), 1940
Derivative of an operator, 169
Detection probability amplitude (photon), 2166
Detectors (photon), 2165
Determinant
Slater determinant, 1438, 1679
Deuterium, 834, 1107(ex.)
Diagonalization

of a 2×2 matrix, 429
of an operator, 128
Diagram (dressed-atom), 2133
Diamagnetism, 855
Diatomic molecules
rotation, 739
Diffusion (momentum), 2030
Dipole
-dipole interaction, 1142, 1153
-dipole magnetic interaction, 1237
electric dipole transition, 863
electric moment, 1080
Hamiltonian, 2011
magnetic dipole moment, 1084
magnetic term, 1272
trap, 2151
Dirac, *see* Fermi
delta function, 97, 106, 280, 1515
equation, 1233
notation, 102
Direct
and exchange terms, 1613, 1632, 1634, 1646, 1650
term, 1447, 1453
Discrete
bases of the state space, 91
spectrum, 132, 217
Dispersion (anomalous), 2149
Dispersion and absorption (field), 2147
Distribution
Boltzmann, 1630
Bose-Einstein, 1630
Fermi-Dirac, 1630
function (bosons), 1629
function (fermions), 1629
functions, 1625, 1733
functions (BCS), 1899
Distribution law
Bose-Einstein, 652
Divergence (energy), 2007
Donor atom, 837, 1495
Doppler
cooling, 2026
effect, 2022
effect (relativistic), 2022
free spectroscopy, 2105
temperature, 2033
Double
condensate, 2237
resonance method, 2059
spin condensate, 2254
Doublet (Autler-Townes), 2144
Down-conversion (parametric), 2181
Dressed
states and energies, 2133
Dressed-atom, 2129, 2133
diagram, 2133
strong coupling, 2141
weak coupling, 2137

E.P.R., 1225(ex.)
Eckart (Wigner-Eckart theorem), *see* Wigner
Effect
Autler-Townes, 2144
Mössbauer, 2040
photoelectric, 2110
Effective Hamiltonian, 2141
Ehrenfest theorem, 242, 319, 522
Eigenresult, 9
Eigenstate, 217, 232
Eigenvalue, 11, 25, 176, 216
degenerate, 217, 260
equation, 126, 429
of an operator, 126
Eigenvector, 176
of an operator, 126
Einstein, 2110
coefficients, 1334, 1356, 2083
EPR argument, 297, 1104
model, 534, 653
Planck-Einstein relations, 3
temperature, 659
Einstein-Podolsky-Rosen, 2204, 2261
Elastic
scattering, 925
scattering (photon), 2086
scattering, form factor, 1411(ex.)
total cross section, 972
Elastically bound electron model, 1350
Electric
conductivity of a solid, 1492
Electric dipole
Hamiltonian, 2011
interaction, 1342

matrix elements, 1344
moment, 1080
selection rules, 1345
transition and selection rules, 863
transitions, 2056
Electric field (quantized), 2000, 2005
Electric polarisability
NH_3, 484
Electric polarizability
of the 1s state in Hydrogen, 1299
Electric quadrupole
Hamiltonian, 1347
moment, 1082
transitions, 1348
Electric susceptibility
bound electron, 577
of an atom, 1351
Electrical
susceptibility, 1223(ex.)
Electrodynamics
classical, 1957
quantum, 1997
Electromagnetic field
and harmonic oscillators, 1968
and potentials, 321
angular momentum, 1968, 2043
energy, 1966
Lagrangian, 1986, 1992
momentum, 1967, 2019
polarization, 1970
quantization, 631, 637
Electromagnetic interaction of an atom with a wave, 1340
Electromagnetism
fields and potentials, 1536
Electron spin, 393, 985
Electron(s)
configurations, 1463
gas in solids, 1491
in solids, 1177, 1481
mass and charge, *see* front cover pages
Electronic
configuration, 1459
paramagnetic resonance, 1225(ex.)
shell, 827
Elements of reality, 2205
Emergence of a relative phase, 2248, 2253
Emission
of a quantum, 1311
photon, 2080
spontaneous, 2081, 2135
stimulated (or induced), 2081
Energy, *see* Conservation, Uncertainty
and momentum of the transverse electromagnetic field, 1973
band, 381
bands in solids, 1177, 1481
conservation, 248
electromagnetic field, 1966
Fermi energy, 1772
fine structure energy levels, 986
free energy, 2290
levels, 359
levels of harmonic oscillator, 509
levels of hydrogen, 823
of a paired state, 1869
recoil energy, 2023
Ensemble
canonical, 2289
grand canonical, 2291
microcanonical, 2285
statistical ensembles, 2295
Entanglement
quantum, 2187, 2193, 2203, 2242
swapping, 2232
Entropy, 2286
EPR, 2204, 2261
elements of reality, 2205
EPRB, 2205
paradox/argument, 1104
Equation of state
ideal quantum gas, 1640
repulsive bosons, 1745
Equation(s)
Bloch, 1361
Hamilton-Jacobi, 1982, 1983, 1988
Lagrange, 1982, 1993
Lorentz, 1959
Maxwell, 1959
Schrödinger, 11, 12, 306
von Neumann, 306
Essential degeneracy, 811, 825
Ethane (molecule), 1223
Ethylene (molecule), 536, 881

Evanescent wave, 29, 67, 70, 78, 285
Evaporative cooling, 2034
Even operators, 196
Evolution
  field operator, 1765
  of quantum systems, 223
  of the mean value, 241
  operator, 313, 2069
  operator (expansion), 2070
  operator (integral equation), 2069
Exchange, 1611
  degeneracy, 1423
  degeneracy removal, 1435
  energy, 1469
  hole, 1774
  integral, 1474
  term, 1447, 1451, 1453
Excitations
  BCS, 1923
  Bogolubov, 1661
  vacuum, 1623
Excited states (BCS), 1919
Exciton, 838
Exclusion principle (Pauli), 1437, 1444, 1463, 1484
Extensive (or intensive) variables, 2292

Fermi
  contact term, 1238
  energy, 1445, 1481, 1486, 1772
  gas, 1481
  golden rule, 1318
  level, 1486, 1621
  radius, 1621
  surface (modified), 1914
  , *see* Fermi-Dirac
Fermi level
  and electric conductivity, 1492
Fermi-Dirac
  distribution, 1486, 1630, 1717
  statistics, 1446
Fermions, 1434
  in a Fock state, 1771
  paired, 1874
Ferromagnetism, 1477
Feynman
  path, 2267
  postulates, 341
Fictitious spin, 435, 1359
Field
  absorption, 2149
  commutation relations, 1989, 1996
  dispersion and absorption, 2147
  intense laser, 2126
  interaction energy, 1764
  kinetic energy, 1763
  normal variables, 1971
  operator, 1752
  operator (evolution), 1763, 1765
  pair field operator, 1861
  potential energy, 1764
  quantization, 1765, 1999
  quasi-classical state, 2008
  spatial correlation functions, 1758
Final states continuum, 1378, 1380
Fine and hyperfine structure, 1231
Fine structure
  constant, *see* front cover pages, 825
  energy levels, 1478
  Hamiltonian, 1233, 1276, 1478
  Helium atom, 1478
  Hydrogen, 1238
  of spectral lines, 986
  of the states $1s$, $2s$ et $2p$, 1276
Fletcher, 2111
Fluctuations
  boson occupation number, 1633
  intensity, 2125
  vacuum, 644, 2007
Fluorescence (single atom), 2121
Fluorescence triplet, 2144
Fock
  space, 1593, 2004
  state, 1593, 1614, 1769, 2103
Forbidden, *see* Band
  energy band, 381, 390, 1481
  transition, 1345
Forces
  van der Waals, 1151
Form factor
  elastic scattering, 1411(ex.)
Forward scattering (direct and exchange), 1874
Fourier

series and transforms, 1505
Fragmentation (condensate), 1654, 1776
Free
electrons in a box, 1481
energy, 2290
particle, 14
quantum field (Fock space), 2004
spherical wave, 941, 944, 961
spherical waves and plane waves, 967
Free particle
stationary states with well-defined angular momentum, 959
stationary states with well-defined momentum, 19
wave packet, 14, 57, 347
Frequency
Bohr, 249
components of the field (positive and negative), 2072
Rabi's frequency, 1325
Friction (coefficient), 2028
Function
of operators, 166
periodic functions, 1505
step functions, 1521
Fundamental state, 41

Gap (BCS), 1894, 1896, 1923
Gauge, 1343, 1536, 1960, 1963
Coulomb, 1965
invariance, 321
Lorenz, 1965
Gaussian
wave packet, 57, 292, 2305
Generalized velocities, 214, 1530
Geometric quantization, 2311
Gerlach, *see* Stern
GHZ state, 2222, 2227
Gibbs-Duhem relation, 2296
Golden rule (Fermi), 1318
Good quantum numbers, 248
Grand canonical, 1626, 2291
Grand potential, 1627, 1721, 2292
Green's function, 337, 936, 1781, 1786, 1789
evolution, 1785
Greenberger-Horne-Zeilinger, 2227
Groenewold's formula, 2315
Gross-Pitaevskii equation, 1643, 1657
Ground state, 363
harmonic oscillator, 509, 520
Hydrogen atom, 1228(ex.)
Group velocity, 55, 60, 614
Gyromagnetic ratio, 396, 455
orbital, 860
spin, 988

H_2^+ molecular ion, 85(ex.), 417, 1189
Hadronic atoms, 840
Hall effect, 1493
Hamilton
function, 1532
function and equations, 1531
Hamilton-Jacobi canonical equations, 214, 1532, 1982, 1983, 1988
Hamiltonian, 223, 245, 1527, 1983, 1988, 1995
classical, 1531
effective, 2141
electric dipole, 1342, 2011
electric quadrupole, 1347
fine structure, 1233, 1276
hyperfine, 1237, 1267
magnetic dipolar, 1347
of a charged particle in a vector potential, 1539
of a particle in a central potential, 806, 1533
of a particle in a scalar potential, 225
of a particle in a vector potential, 225, 323, 328
Hanbury Brown and Twiss, 2120
Hanle effect, 1372(ex.)
Hard sphere
scattering, 980, 981(ex.)
Harmonic oscillator, 497
in an electric field, 575
in one dimension, 527, 1131
in three dimensions, 569
in two dimensions, 755
infinite chain of coupled oscillators, 611
quasiclassical states, 583
thermodynamic equilibrium, 647

three-dimensional, 841, 899(ex.)
two coupled oscillators, 599
Hartree-Fock
approximation, 1677, 1701
density operator (one-particle), 1691
equations, 1686, 1731
for electrons, 1695
mean field, 1677, 1693
potential, 1706
thermal equilibrium, 1711, 1733
time-dependent, 1701, 1708
Healing length, 1652
Heaviside step function, 1521
Heisenberg
picture, 317, 1763
relations, 19, 39, 41, 45, 55, 232, 290
Helicity (photon), 2051
Helium
energy levels, 1467
ion, 838
isotopes, 1480
isotopes ^{3}He and ^{4}He, 1435, 1446
solidification, 535
Hermite polynomials, 516, 547, 561
Hermitian
conjugation, 111
matrix, 124
operator, 115, 124, 130
Histories (classical), 2272
Hole
creation and annihilation, 1622
exchange, 1774
Holes, 1621
Hybridization of atomic orbitals, 869
Hydrogen, 645
atom, 803
atom in a magnetic field, 853, 855, 862
atom, relativistic energies, 1245
Bohr model, 40, 819
energy levels, 823
fine and hyperfine stucture, 1231
ionisation energy, *see* front cover pages
ionization energy, 820
maser, 1251
molecular ion, 85(ex.), 417, 1189
quantum theory, 41
radial equation, 821
Stark effect, 1298
stationary states, 851
stationary wave functions, 830
Hydrogen-like systems in solid state physics, 837
Hydrogenoid systems, 833
Hyperfine
decoupling, 1262
Hamiltonian, 1237, 1267
Hyperfine structure, *see* Hydrogen, muonium, positronium, Zeeman effect, 1231
Muonium, 1281

Ideal gas, 1625, 1787, 1791, 1804
correlations, 1769
Identical particles, 1419, 1591
Induced
emission, 1334, 1366, 2081
emission of a quantum, 1311
emission of photons, 1355
Inequality (Bell's), 2208
Infinite one-dimensional well, 271
Infinite potential well, 74
in two dimensions, 201
Infinitesimal unitary operator, 178
Insulator, 1492
Integral
exchange integral, 1474
scattering equation, 935
Intense laser fields, 2126
Intensive (or extensive) variables, 2292
Interaction
between magnetic dipoles, 1141
dipole-dipole interaction, 1141, 1153
electromagnetic interaction of an atom with a wave, 1340
field and particles, 2009
field and atom, 2010
magnetic dipole-dipole interaction, 1237
picture, 353, 1393, 2070
tensor interaction, 1141
Interference
photons, 2167
two-photon, 2170, 2183
Ion H_2^+, 1189

Ionization
photo-ionization, 2109
tunnel ionization, 2126
Isotropic radiation, 2079

Jacobi, *see* Hamilton

Kastler, 2059, 2062
Ket, *see* state, 103, 119
for identical particles, 1436
Kuhn, *see* Thomas

Lagrange
equations, 1530, 1982, 1993
fonction and equations, 214
multipliers, 2281
Lagrangian, 1530, 1980
densities, 1986
electromagnetic field, 1986, 1992
formulation of quantum mechanics, 339
of a charged particle in an electromagnetic field, 1538
particle in an electromagnetic field, 323
Laguerre-Gaussian beams, 2065
Lamb shift, 645, 1245, 1388, 2008
Landau levels, 771
Landé factor, 1072, 1107(ex.), 1256, 1292
Laplacian, 1527
of $1/r$, 1524
of $Y_l^m(\theta,\varphi)/r^{l+1}$, 1526
Larmor
angular frequency, 857
precession, 394, 396, 410, 455, 857, 1071
Laser, 1359, 1365
Raman laser, 2093
saturation, 1370
trap, 2151
Lattices (optical), 2153
Least action
principle of, 1539
Legendre
associated function, 714
polynomial, 713
Length (healing), 1652
Level
anticrossing, 415, 482
Fermi level, 1621
Lifetime, 343, 485, 645
of a discrete state, 1386
radiative, 2081
Lifting of degeneracy by a perturbation, 1125
Light
quanta, 3
shifts, 1334, 2138, 2151, 2156
Linear, *see* operator
combination of atomic orbitals, 1172
operators, 90, 108, 163
response, 1350, 1357, 1364
superposition of states, 253
susceptibility, 1365
Local conservation of probability, 238
Local realism, 2209, 2230
Longitudinal
fields, 1961
relaxation, 1400
relaxation time, 1401
Lorentz equations, 1959
Lorenz (gauge), 1965

Magnetic
dipole term, 1272
dipole-dipole interaction, 1237
effect of a magnetic field on the levels of the Hydrogen atom, 1251
hyperfine Hamiltonian, 1267
interactions, 1232, 1237
quantum number, 811
resonance, 455
susceptibility, 1224, 1487
Magnetic dipole
Hamiltonian, 1347
transitions and selection rules, 1084, 1098, 1348
Magnetic dipoles
interactions between two dipoles, 1141
Magnetic field
and vector potential, 321
charged particle in a, 240, 771
effects on hydrogen atom, 853, 855
harmonic oscillator in a, 899(ex.)
Hydrogen atom in a magnetic field, 1263, 1289

multiplets, 1074
quantized, 2000, 2005
Magnetism (spontaneous), 1737
Many-electron atoms, 1459
Maser, 477, 1359, 1365
hydrogen, 1251
Mass correction (relativistic), 1234
Master equation, 1358
Matrice(s), 119, 121
diagonalization of a 2×2 matrix, 429
Pauli matrices, 425
unitary matrix, 176
Maxwell's equations, 1959
Mean field (Hartree-Fock), 1693, 1708, 1725
Mean value of an observable, 228
evolution, 241
Measurement
general postulates, 216, 226
ideal von Neumann measurement, 2196
of a spin 1/2, 394
of observables, 216
on a part of a physical system, 293
state after measurement, 221, 227
Mendeleev's table, 1463
Metastable superfluid flow, 1671
Methane (molecule), 883
Microcanonical ensemble, 2285
Millikan, 2111
Minimal wave packet, 290, 520, 591
Mirrors for atoms, 2153
Mixing of states, 1121, 1137
Model
Cooper model, 1927
Einstein model, 534
elastically bound electron, 1350
vector model of atom, 1071
Modes
vibrational modes, 599, 611
Modes (radiation), 1974, 1975
Molecular ion, 417
Molecule(s)
chemical bond, 417, 869, 873, 878, 883, 1189
rotation, 796
vibration, 527, 1137
vibration-rotation, 885
Mollow, 2144
Moment
quadrupole electric moment, 1225(ex.)
Momentum, 1539
conjugate, 214, 323, 1983, 1987, 1995
diffusion, 2030
electromagnetic field, 1967, 2019
mechanical momentum, 328
Monogamy (quantum), 2221
Mössbauer effect, 1415, 2040
Motional narrowing, 1323
condition, 1323, 1398, 1408
Multiphoton transition, 1368, 2040, 2097
Multiplets, 1072, 1074, 1467
Multipliers (Lagrange), 2281
Multipolar waves, 2052
Multipole
moments, 1077
Multipole operators
introduction, 1077, 1083
parity, 1082
Muon, 527, 541, 1281
Muonic atom, 541, 839
Muonium, 835
hyperfine structure, 1281
Zeeman effect, 1281

Narrowing (motional), 1323, 1408
condition, 1398
Natural width, 345, 1388
Need for a quantum treatment, 2118, 2120
Neumann
spherical function, 967
Neutron mass, *see* front cover pages
Non-destructive detection of a photon, 2159
Non-diagonal order (BCS), 1912
Non-locality, 2204
Non-resonant excitation, 1350
Non-separability, 2207
Nonlinear
response, 1357, 1368
susceptibility, 1369
Norm
conservation, 238
of a state vector, 104, 237
of a wave function, 13, 90, 99

Normal
correlation function, 1782, 1787
variables, 602, 616, 631, 633
variables (field), 1971
Nuclear
multipole moments, 1088
Bohr magneton, 1237
Nucleus
spin, 1088
volume effect, 1162, 1268
Number
occupation number, 1439, 1593
photon number, 2135
total number of particles in an ideal gas, 1635

Observable(s), 130
C.S.C.O., 133, 137
commutation, 232
compatibility, 232
for identical particles, 1429, 1441
mean value, 228
measurement of, 216, 226
quantization rules, 223
symmetric observables, 1441
transformation by permutation, 1434
whose commutator is $i\hbar$, 187, 289
Occupation number, 1439, 1593
operator, 1598
Odd operators, 196
One-particle
Hartree-Fock density operator, 1691
operators, 1603, 1605, 1628, 1756
Operator(s)
adjoint operator, 112
annihilation operator, 504, 513, 514, 1597
creation and annihilation, 1990
creation operator, 504, 513, 514, 1596
derivative of an operator, 169
diagonalization, 126, 128
even and odd operators, 196
evolution operator, 313, 2069
field, 1752
function of, 166
Hermitian operators, 115
linear operators, 90, 108, 163
occupation number, 1598
one-particle operator, 1603, 1605, 1628, 1756
parity operator, 193
particle density operator, 1756
permutation operators, 1425, 1430
potential, 168
product of, 90
reduced to a single particle, 1607
representation, 121
restriction, 165
restriction of, 1125
rotation operator, 1001
symmetric, 1628, 1755
translation operator, 190
two-particle operator, 1608, 1610, 1631, 1756
unitary operators, 173
Weyl operator, 2300
Oppenheimer, *see* Born, 1177, 1190
Optical
excitation (broadband), 1332
lattices, 2153
pumping, 2062, 2140
Orbital
angular momentum (of radiation), 2052
atomic orbital, 1496(ex.)
hybridization, 869
linear combination of atomic orbitals, 1172
quantum number, 1463
state space, 988
Order parameter for pairs, 1851
Orthonormal basis, 91, 99, 101, 133
characteristic relation, 116
Orthonormalization
and closure relations, 101, 140
relation, 116
Oscillation(s)
between two discrete states, 1374
between two quantum states, 418
Rabi, 2134
Oscillator
anharmonic, 502
harmonic, 497
strength, 1352

Pair(s)

annihilation-creation of pairs, 1831, 1874, 1887
BCS, wave function, 1909
Cooper, 1927
of particles (creation operator), 1813, 1846
pair field (commutation), 1861
pair field operator, 1845
pair wave function, 1851
Paired
bosons, 1881
fermions, 1874
state energy, 1869
states, 1811
states (building), 1818
Pairing term, 1878
Paramagnetism, 855
Parametric down-conversion, 2181
Parity, 2106
degeneracy, 199
of a permutation operator, 1431
of multipole operators, 1082
operator, 193
Parseval
Parseval-Plancherel equality, 20
Parseval-Plancherel formula, 1511, 1521
Partial
reflection, 79
trace of an operator, 309
waves in the potential, 948
waves method, 941
Particle (current), 1758
Particles and holes, 1621
Partition function, 1626, 1627, 1717
Path
integral, 2267
space-time path, 339
Pauli
exclusion principle, 1437, 1444, 1463, 1481
Hamiltonian, 1009(ex.)
matrices, 425, 991
spin theory, 986
spinor, 993
Penetrating orbit, 1463
Penrose-Onsager criterion, 1776, 1860, 1947
Peres, 2212
Periodic
boundary conditions, 1489
classification of elements, 1463
functions, 1505
potential (one-dimensional), 375
Permutation operators, 1425, 1430
Perturbation
applications of the perturbation theory, 1231
lifting of a degeneracy, 1125
one-dimensional harmonic oscillator, 1131
random perturbation, 1320, 1325, 1390
sinusoidal, 1311
stationary perturbation theory, 1115
Perturbation theory
time dependent, 1303
Phase
locking (BCS), 1893, 1916
locking (bosons), 1938, 1944
relative phase between condensates, 2237, 2248
velocity, 37
Phase shift (collision), 951, 1497(ex.)
with imaginary part, 971
Phase velocity, 21
Phonons, 611, 626
Bogolubov phonons, 1660
Photodetection
double, 2172, 2184
single, 2169, 2171
Photoelectric effect, 1412(ex.), 2110
Photoionization, 2109, 2165
rate, 2115, 2124
two-photon, 2123
Photon, 3, 631, 651, 2004, 2005, 2110
absorption and emission, 2067
angular momentum, 1370
antibunching, 2121
detectors, 2165
non-destructive detection, 2159
number, 2135
scattering (elastic), 2086
scattering by an atom, 2085
vacuum, 2007
, *see* Absorption, Emission
Picture

Heisenberg, 317, 1763
interaction, 1393, 2070
Pitaevskii (Gross-Pitaevskii equation), 1643, 1657
Plancherel, *see* Parseval
Planck
constant, *see* front cover pages, 3
law , 2083
Planck-Einstein relations, 3, 10
Plane wave, 14, 19, 95, 943
Podolsky (EPR argument), 297, 1104
Pointer states, 2199
Polarizability
of the 1s state in Hydrogen, 1299
Polarization
electromagnetic field, 1970
of Zeeman components, 1295
space-dependent, 2156
Polynomial method (harmonic oscillator), 555, 842
Polynomials
Hermite polynomials, 516, 547, 561
Position and momentum representations, 181
Positive and negative frequency components, 2072
Positron, 1281
Positronium, 836
hyperfine structure, 1281
Zeeman effect, 1281
Postulate (von Neumann projection), 2202
Postulates of quantum mechanics, 215
Potential
adiabatic branching, 932
barrier, 26, 68, 367, 373
centrifugal potential, 809, 888, 893
Coulomb potential, cross section, 979
cylindrically symmetric, 899(ex.)
Hartree-Fock, 1706
infinite one-dimensional well, 74
operator, 168
scalar and vector potentials, 1536, 1960, 1963
scattering by a, 923
self-consistent potential, 1461
square potential, 63
square well, 29
step, 28, 65, 75, 284
well, 71, 367
well (arbitrary shape), 359
well (infinite one-dimensional), 271
well (infinite two-dimensional, 201
Yukawa potential, 977
Precession
Larmor precession, 396, 1071
Thomas precession, 1235
Preparation of a state, 235
Pressure (ideal quantum gas), 1640
Principal part, 1517
Principal quantum number, 827
Principle
of least action, 1539, 1980
of spectral decomposition, 11, 216
of superposition, 237
Probability
amplitude, 11, 253, 259
conservation, 237
current, 240, 283, 333, 349, 932
current in hydrogen atom, 851
density, 11, 264
fluid, 932
of photon absorption, 2076
of the measurement results, 9, 11
transition probability, 439
Process (pair annihilation-creation), 1878, 1887
Product
convolution product of functions, 1510
of matrices, 122
of operators, 90
scalar product, 101, 141, 149, 161
state (tensor product), 311
tensor product, 147
tensor product, applications, 441
Projection theorem, 1070
Projector, 109, 133, 165, 218, 222, 1108(ex.)
Propagator
for the Schrödinger equation, 335
of a particle, 2267, 2272
Proper result, 9
Proton
mass, *see* front cover pages
spin and magnetic moment, 1237, 1274
Pumping, 1358

Pure (state or case), 301

Quadrupolar electric moment, 1082, 1225(ex.)
Quanta (circular), 761, 783
Quantization
  electrodynamics, 1997
  electromagnetic field, 631, 637, 1997
  of a field, 1765
  of angular momentum, 394, 677
  of energy, 3, 11, 71, 359
  of measurement results, 9, 216, 398
  of the measurement results, 405
  rules, 11, 223, 226, 2274
Quantum
  angle, 2258
  electrodynamics, 1245, 1282, 1997
  entanglement, 2187, 2193
  monogamy, 2221
  number
    orbital, 1463
    principal quantum number, 827
  numbers (good), 248
  resonance, 417
  treatment needed, 2118, 2120
Quasi-classical
  field states, 2008
  states, 765, 791, 801
  states of the harmonic oscillator, 583
Quasi-particles, 1736, 1840
  Bogolubov phonons, 1954
  Quasi-particle vacuum, 1836

Rabi
  formula, 440, 460, 1324, 1376
  formula), 419
  frequency, 1325
  oscillation, 2134
Radial
  equation, 842
  equation (Hydrogen), 821
  equation in a central potential, 808
  integral, 1277
  quantum number, 811
Radiation
  isotropic, 2079
  pressure, 2024
Radiative
  broadening, 2138
  cascade of the dressed atom, 2145
Raman
  effect, 532, 740, 1373(ex.)
  laser, 2093
  scattering, 2091
  scattering (stimulated), 2093
Random perturbation, 1320, 1325, 1390
Rank (Schmidt), 2196
Rate (photoionization), 2115, 2124
Rayleigh
  line, 752
  scattering, 532, 2089
Realism (local), 2205, 2209
Recoil
  blocking, 2036
  effect of the nucleus, 834
  energy, 1415, 2023
  free atom, 2020
  suppression, 2040
Reduced
  density operator, 1607
  mass, 813
Reduction of the wave packet, 221, 279
Reflection on a potential step, 285
Refractive index, 2149
Reiche, *see* Thomas
Relation (Gibbs-Duhem), 2296
Relative
  motion, 814
  particle, 814
  phase between condensates, 2248, 2258
  phase between spin condensates, 2253
Relativistic
  corrections, 1233, 1478
  Doppler effect, 2022
  mass correction, 1234
Relaxation, 465, 1358, 1390, 1413, 1414(ex.)
  general equations, 1397
  longitudinal, 1400
  longitudinal relaxation time, 1401
  transverse, 1403
  transverse relaxation time, 1406
Relay state, 2086, 2098, 2106
Renormalization, 2007
Representation(s)
  change of, 124
  in the state space, 116

of operators, 121
position and momentum, 139, 181
Schrödinger equation, 183–185
Repulsion between electrons, 1469
Resonance
magnetic resonance, 455
quantum resonance, 417, 1158
scattering resonance, 69, 954, 983(ex.)
two resonnaces with a sinusoidal excitation, 1365
width, 1312
with sinusoidal perturbation, 1311
Restriction of an operator, 165, 1125
Rigid rotator, 740, 1222(ex.)
Ritz theorem, 1170
Root mean square deviation
general definition, 230
Rosen (EPR argument), 297, 1104
Rotating frame, 459
Rotation(s)
and angular momentum, 717
invariance and degeneracy, 734
of diatomic molecules, 739
of molecules, 796, 885
operator(s), 720, 1001
rotation invariance, 1478
rotation invariance and degeneracy, 1072
Rotator
rigid rotator, 740, 1222(ex.)
Rules
quantization rules, 2274
selection rules, 197
Rutherford's formula, 979
Rydberg constant, *see* front cover pages

Saturation
of linear response, 1368
of the susceptibility, 1369
Scalar
and vector potentials, 321, 1536
interaction between two angular momenta, 1091
observable, operator, 732, 737
potential, 225
product, 89, 92, 101, 141, 149, 161
product of two coherent states, 593
Scattering
amplitude, 929, 953
by a central potential, 941
by a hard sphere, 980, 981(ex.)
by a potential, 923
cross section, 933, 953, 972
cross section and phase shifts, 951
inelastic, 2091
integral equation, 935
of particles with spin, 1102
of spin 1/2 particles, 1108(ex.)
photon, 2086
Raman, 2091
Rayleigh, 532, 2089
resonance, 954, 983(ex.)
resonant, 2089
stationary scattering states, 951
stationary states, 928
stimulated Raman, 2093
Schmidt
decomposition, 2193
rank, 2196
Schottky anomaly, 654
Schrödinger, 2190
equation, 11, 12, 223, 306
equation in momentum representation, 184
equation in position representation, 183
equation, physical implications, 237
equation, resolution for conservative systems, 245
picture, 317
Schwarz inequality, 161
Second
quantization, 1766
harmonic generation, 1368
Secular approximation, 1316, 1374
Selection rules, 197, 863, 2014, 2056
electric quadrupolar, 1348
magnetic dipolar, 1098, 1348
Self-consistent potential, 1461
Semiconductor, 837, 1493
Separability, 2207, 2223
Separable density operator, 2223
Shell (electronic), 827
Shift

light shift, 2138
of a discrete state, 1387
Singlet, 1024, 1474
Sinusoidal perturbation, 1311, 1374
Sisyphus
cooling, 2034
effect, 2155
Slater determinant, 1438, 1679
Slowing down atoms, 2025
Solids
electronic bands, 1177
energy bands of electrons, 1491
energy bands of electrons in solids, 381
hydrogen-like systems in solid state physics, 837
Space (Fock), 1593
Space-dependent polarization, 2156
Space-time path, 339, 1539
Spatial correlations (ideal gas), 1769
Specific heat
of an electron gas, 1484
of metals, 1487
of solids, 653
two level system, 654
Spectral
decomposition principle, 7, 11, 216
function, 1795
terms, 1469
Spectroscopy (Doppler free), 2105
Spectrum
BCS elementary excitation, 1923
continuous, 219, 264
discrete, 132, 217
of an observable, 126, 216
Spherical
Bessel equation, 961
Bessel function, 944, 966
free spherical waves, 961
free wave, 944
Neumann function, 967
wave, 941
waves and plane waves, 967
Spherical harmonics, 689, 705
addition of, 1059
expression for $l = 0, 1, 2$, 709
general expression, 707
Spin
and magnetic moment of the proton, 1237
angular momentum, 987
electron, 985, 1289
fictitious, 435
gyromagnetic ratio, 396, 455, 988
nuclear, 1088
of the electron, 393
Pauli theory, 986, 988
quantum description, 985, 991
rotation operator, 1001
scattering of particles with spin, 1102
spin 1 and radiation, 2044, 2049, 2050
system of two spins, 441
Spin 1/2
density operator, 449
ensemble of, 1358
fictitious, 1359
interaction between two spins, 1141
preparation and measurement, 401
scattering of spin 1/2 particles, 1108(ex.)
Spin-orbit coupling, 1018, 1234, 1241, 1279
Spin-statistics theorem, 1434
Spinor, 993
rotation, 1005
Spontaneous
emission, 343, 645, 1301, 2081, 2135
emission of photons, 1356
magnetism of fermions, 1737
Spreading of a wave packet, 59, 348
Square
barrier of potential, 26, 68
potential, 26, 63, 75, 283
potential well, 71, 271
spherical well, 982(ex.)
Standard representation (angular momentum), 677, 691
Stark effect in Hydrogen atom, 1298
State(s), *see* Density operator
density of, 389, 1316, 1484, 1488
Fock, 1593, 1614, 1769, 2103
ground state, 363
mixing of states by a perturbation, 1121
orbital state space, 988
paired, 1811

pointer states, 2199
quasi-classical states, 583, 765, 791, 801
relay state, 2086, 2098, 2106
stable and unstable states, 485
state after measurement, 221
state preparation, 235
stationary, 63, 359, 375
stationary state, 24, 246
stationary states in a central potential, 804
unstable, 343
vacuum state, 1595
vector, 102, 215
Stationary
perturbation theory, 1115
phase condition, 18, 54
scattering states, 928, 951
states, 24, 63, 246, 359
states in a periodic potential, 375
states with well-defined angular momentum, 944, 959
states with well-defined momentum, 943
Statistical
entropy, 2217
mechanics (review of), 2285
mixture of states, 253, 299, 304, 450
Statistics
Bose-Einstein, 1446
Fermi-Dirac, 1446
Step
function, 1521
potential, 28, 65, 75, 284
Stern-Gerlach experiment, 394
Stimulated
(or induced) emission, 1334, 1366, 2081
Raman scattering, 2093
Stokes Raman line, 532, 752
Stoner (spontaneous magnetism), 1737
Strong coupling (dressed-atom), 2141
Subrecoil cooling, 2034
Sum rule (Thomas-Reiche-Kuhn), 1352
Superfluidity, 1667, 1674
Superposition
of states, 253
principle, 7, 237
principle and physical predictions, 253
Surface (modified Fermi surface), 1914
Susceptibility, *see* Linear, nonlinear, tensor
electric susceptibility of an atom, 1351
electrical susceptibility, 577, e1223
electrical susceptibility of NH_3, 484
magnetic susceptibility, 1224
tensor, 1224, 1410(ex.)
Swapping (entanglement), 2232
Symmetric
ket, state, 1428, 1431
observables, 1429, 1441
operators, 1603, 1605, 1608, 1610, 1628, 1631, 1755
Symmetrization
of observables, 224
postulate, 1434
Symmetrizer, 1428, 1431
System
time evolution of a quantum system, 223
two-level system, 435
Systematic
and accidental degeneracies, 203
degeneracy, 845

Temperature (Doppler), 2033
Tensor
interaction, 1141
product, 147, 441
product of operators, 149
product state, 295, 311
product, applications, 201
susceptibility tensor, 1224
Term
direct and exchange terms, 1613, 1632, 1634, 1646, 1650
pairing, 1878
spectral terms, 1467, 1469
Theorem
Bell, 2204, 2208
Bloch, 659
projection, 1070
Ritz, 1170
Wick, 1799, 1804

Wigner-Eckart, 1065, 1085, 1254
Thermal wavelength, 1635
Thermodynamic equilibrium, 308
harmonic oscillator, 647
ideal quantum gas, 1625
spin 1/2, 452
Thermodynamic potential (minimization), 1715
Thomas precession, 1235
Thomas-Reiche-Kuhn sum rule, 1352
Three-dimensional harmonic oscillator, 569, 841, 899(ex.)
Three-level system, 1409(ex.)
Three-photon transition, 1370
Time evolution of quantum systems, 223
Time-correlations (fluorescent photons), 2145
Time-dependent
Gross-Pitaevskii equation, 1657
perturbation theory, 1303
Time-energy uncertainty relation, 250, 279, 345, 1312, 1389
Torsional oscillations, 536
Torus (flow in a), 1667
Total
elastic scattering cross section, 972
reflection, 67, 75
scattering cross section (collision), 926
Townes
Autler-Townes effect, 1410
Trace
of an operator, 163
partial trace of an operator, 309
Transform (Wigner), 2297
Transformation
Bogolubov, 1950
Bogolubov-Valatin, 1836, 1919
Gauge, 1960
of observables by permutation, 1434
Transition, *see* Probability, Forbidden, Electric dipole, Magnetic dipole, Quadrupole electric dipole, 2056
magnetic dipole transition, 1098
probability, 439, 1308, 1321, 1355
probability per unit time, 1319
probability, spin 1/2, 460
three-photon transition, 1370
two-photon, 2097
virtual, 2100
Translation operator, 190, 579, 791
Transpositions, 1431
Transverse
fields, 1961
relaxation, 1403
relaxation time, 1406
Trap
dipolar, 2151
laser, 2151
Triplet, 1024, 1474
fluorescence triplet, 2144
Tunnel
effect, 29, 70, 365, 476, 540, 1177
ionization, 2126
Two coupled harmonic oscillators, 599
Two-dimensional
harmonic oscillator, 755
infinite potential well, 201
wave packets, 49
Two-level system, 393, 411, 435, 1357
Two-particle operators, 1608, 1610, 1631, 1756
Two-photon
absorption, 1373(ex.)
interference, 2170, 2183
transition, 1409(ex.), 2097

Uncertainty
relation, 19, 39, 41, 45, 232, 290
time-energy uncertainty relation, 1312
Uniqueness of the measurement result, 2201
Unitary
matrix, 125, 176
operator, 173, 314
transformation of operators, 177
Unstable states, 343

Vacuum
electromagnetism, 644, 2007
excitations, 1623
fluctuations, 2007
photon vacuum, 2007
quasi-particule vacuum, 1836
state, 1595
Valence band, 1493

Van der Waals forces, 1151
Variables
intensive or extensive, 2292
normal variables, 602, 616, 631, 633
Variational method, 1169, 1190, 1228(ex.)
Vector
model, 1091
model of the atom, 1071, 1256
observable, operator, 732
operator, 1065
potential, 225
potential of a magnetic dipole, 1268
Velocity
critical, 1671
generalized velocities, 214, 1530
group velocity, 23, 614
phase velocity, 21, 37
Vibration(s)
modes, 599, 611
modes of a continuous system, 631
of molecules, 885, 1137
of nuclei in a crystal, 534, 611, 653
of the nuclei in a molecule, 527
Violations of Bell's inequalities, 2210, 2265
Virial theorem, 350, 1210
Virtual transition, 2100
Volume effect, 544, 840, 1162, 1268
Von Neumann
chain, 2201
equation, 306
ideal measurement, 2196
reduction postulate, 2202
statistical entropy, 2217
Vortex in a superfluid, 1667

Water (molecule), 873, 874
Wave (evanescent), 67
Wave function, 88, 140, 226
BCS pairs, 1901, 1909
Hydrogen, 830
norm, 90
pair wave functions, 1851
particle, 11
Wave packet(s)
Gaussian, 57, 2305
in a potential step, 75
in three dimensions, 53
minimal, 290, 520, 591
motion in a harmonic potential, 596
one-photon, 2168
particle, 13
photon, 2163
propagation, 20, 57, 242, 398
reduction, 221, 227, 265, 279
spreading, 57, 59, 347, 348(ex.)
two-dimension, 49
two-photons, 2181
Wave(s)
de Broglie wavelength, 10, 35
evanescent, 29
free spherical waves, 961
multipolar, 2052
partial waves, 948
plane, 14, 19, 943
wave function, 11, 88, 140, 226
Wave-particle duality, 3, 45
Wavelength
Compton wavelength, 1235
de Broglie, 10
Weak coupling (dressed-atom), 2137
Well
potential square well, 29
potential well, 367
Weyl
operator, 2300
quantization, 2311
Which path type of experiments, 2202
Wick's theorem, 1799, 1804
Wigner transform, 2297
Wigner-Eckart theorem, 1065, 1085, 1254

Young (double slit experiment), 4
Yukawa potential, 977

Zeeman
components, polarizations, 865
effect, 855, 862, 987, 1251, 1253, 1257, 1261, 1281
polarization of the components, 1295
slower, 2025
Zeeman effect
Hydrogen, 1289
in muonium, 1281
in positronium, 1281
Muonium, 1284

Zone (Brillouin zone), 614

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.

WILEY-VCH

This new, third volume of Cohen-Tannoudji's groundbreaking textbook covers advanced topics of quantum mechanics such as uncorrelated and correlated identical particles, the quantum theory of the electromagnetic field, absorption, emission and scattering of photons by atoms, and quantum entanglement. Written in a didactically unrivalled manner, the textbook explains the fundamental concepts in seven *chapters,* each followed by *complements* that provide more detailed discussions, examples and applications.

- Completing the success story: the third and final volume of the quantum mechanics textbook written by 1997 Nobel laureate Claude Cohen-Tannoudji and his colleagues Bernard Diu and Franck Laloë
- As easily comprehensible as possible: all steps of the physical background and its mathematical representation are spelled out explicitly
- Comprehensive: in addition to the fundamentals themselves, the books comes with a wealth of elaborately explained examples and applications

Claude Cohen-Tannoudji *is a researcher at the Kastler-Brossel laboratory of the Ecole Normale Supérieure in Paris where he also studied and received his PhD in 1962. In 1973 he became Professor of atomic and molecular physics at the Collège des France. His main research interests were optical pumping, quantum optics and atom-photon interactions. In 1997, Claude Cohen-Tannoudji, together with Steven Chu and William D. Phillips, was awarded the Nobel Prize in Physics for his research on laser cooling and trapping of neutral atoms.*

Bernard Diu *was Professor at the Denis Diderot University (Paris VII). He was engaged in research at the Laboratory of Theoretical Physics and High Energy where his focus was on strong interactions physics and statistical mechanics.*

Franck Laloë *is a researcher at the Kastler-Brossel laboratory of the Ecole Normale Supérieure in Paris. His first assignment was with the University of Paris VI before he was appointed to the CNRS, the French National Research Center. His research was focused on optical pumping, statistical mechanics of quantum gases, musical acoustics and the foundations of quantum mechanics.*

Cover Design: Wiley
Cover Image: © antishock/Getty Images

ISBN 978-3-527-34555-7

www.wiley-vch.de

Also available as an e-book