Advanced Texts in Physics

This program of advanced texts covers a broad spectrum of topics which are of current and emerging interest in physics. Each book provides a comprehensive and yet accessible introduction to a field at the forefront of modern research. As such, these texts are intended for senior undergraduate and graduate students at the MS and PhD level; however, research scientists seeking an introduction to particular areas of physics will also benefit from the titles in this collection.

Springer-Verlag Berlin Heidelberg GmbH

Louis Marchildon

Quantum Mechanics

From Basic Principles to Numerical Methods and Applications

With 50 Figures,
252 Exercises and Selected Answers

Springer

Professor Louis Marchildon
University of Quebec
Department of Physics
3351, boul. des Forges
G9A 5H7 Trois-Rivieres, Qc
Canada
E-mail: `Louis_Marchildon@UQTR.CA`

Originally published in French under the title:
Louis Marchildon: Mécanique Quantique

Editions De Boeck Université
Rue des Minimes 39, 1000 Bruxelles, Belgium

ISSN 1439-2674

ISBN 978-3-642-07767-8 ISBN 978-3-662-04750-7 (eBook)
DOI 10.1007/978-3-662-04750-7

Library of Congress Cataloging-in-Publication Data applied for.

Die Deutsche Bibliothek - CIP-Einheitsaufnahme:
Marchildon, Louis: Quantum mechanics :
from basic principles to numerical methods and applications / Louis Marchildon. -
Berlin ; Heidelberg ; New York ; Barcelona ; Hongkong ; London ; Mailand ; Paris ; Tokio :
Springer, 2002
(Advanced texts in physics)

http://www.springer.de

Originally published by Springer-Verlag Berlin Heidelberg New York in 2002.
Softcover reprint of the hardcover 1st edition 2002

Typesetting by the author
Cover picture: Designed using Fig. 18.1
Cover design: *design & production* GmbH, Heidelberg
Printed on acid-free paper SPIN 10863515 56/3141/tr 5 4 3 2 1 0

Preface

Whoever begins writing a book on quantum mechanics is struck by the breadth of the subject. In its applications first: atomic and molecular physics, nuclear physics, optics, solid state physics, theory of gases and liquids, elementary particles theory, almost all fields of contemporary physics are based on quantum mechanics. In its formulation, also, which borrows from many subfields of mathematics and reaches philosophical reflection as much as modern technology. The writing therefore implies, at the outset, making choices.

I first chose to write a book for those who strive to understand quantum mechanics. These are physics students, of course, but also students and investigators in theoretical chemistry, biophysics and engineering physics wishing to comprehend more deeply the computational methods they use. I have thus tried to clarify delicate points rather than leave them aside. Conceptual problems are treated in more detail than in most general textbooks. But understanding also involves the capability to perform concrete calculations. This motivates the development of numerical methods which, most of the time, are the only ones that yield quantitative results.

I chose also to present quantum mechanics as a self-contained theory. The exposition largely develops around the central notion of state space. This explains why the discussion of the electromagnetic field goes with the particle in three dimensions, and why spin and angular momentum are treated in three different chapters. Except for a few heuristic remarks, I did not look for a priori justification of the formalism, based for instance on analogies with classical mechanics or the physics of waves. The only convincing justification of a fundamental theory lies in its logical consistency and the agreement between its predictions and experiment. In the first chapter I briefly recall the scientific context in which quantum mechanics developed. But the complex relation between quantum mechanics and classical mechanics is treated a posteriori.

I finally chose to focus on one field of application (atomic and molecular physics) that I could investigate more fully, rather than give an overview of several. Nonetheless, methods of atomic physics often apply to nuclear physics, and the path integral and the theory of symmetry open the way to particle physics.

This work developed from courses taught at Université du Québec à Trois-Rivières. In general, training in quantum mechanics spreads over several years, so that the student's scientific background is more important at the end than at the beginning. For this reason the exposition is deliberately slower in the first few chapters. It does not require knowing any specific detail in modern physics, even though a basic knowledge is certainly helpful. The theory of vector spaces is summarized in the second chapter. Properties of functions of mathematical physics are presented where needed. A few sections (11.3, 11.8, 16.2 and 16.5) require more advanced notions of complex variables. The reader who has not acquired them at that time can look at the references proposed.

The first eleven chapters gather material for a one-year undergraduate course in quantum mechanics. To this one can add the first few sections of Chap. 12, sections in Chap. 14 on the addition of angular momentum, the first two sections of Chap. 17 and perhaps the first three of Chap. 21. With the remainder one can build a one-semester graduate course in advanced quantum mechanics or atomic and molecular physics.[1]

The problems given at the end of most chapters are an integral part of the course. Learning quantum mechanics cannot be accomplished by merely assimilating theoretical concepts, but must be completed by concrete calculations. The student is urged to work out, if not all, at least a substantial fraction of exercises. Hints on solutions are given at the end of the book.

Although the exposition is largely self-contained, numerous references are given to books or papers that the reader wishing to delve further on a topic is invited to consult. I did not attempt to go back to original papers. They or references to them can be found in [124], [125], [146], [200] and [227]. References were chosen to complement the exposition and for their pedagogical value.

It is a great pleasure to thank all those who to any degree helped to carry out this project. Several colleagues and students commented on the first version of the manuscript, in particular Jean-René Cliche, Gilles Couture, Louis J. Dubé, Joseph LeBrech, Gérald M. Lefebvre, Pierre Ouellette and David Sénéchal. Pierre Mathieu added constant heartening to his shrewd remarks. I thank Lucie Bellemare, Jacqueline Moreau, Patrick Jacob and Marie Louise Héroux for typesetting text and equations; Robert McDougall for writing software calculating values shown in Tables 9.1, 18.6, 18.7 and 19.1; Leslie Ballentine who welcomed me at Simon Fraser when this work was just beginning; and four anonymous referees who carefully evaluated the French manuscript. I am grateful to Université du Québec à Trois-Rivières for a

---

[1] The equations, formatted for projection on a screen, are available in `.dvi` and `.ps` formats at the ftp server of Université du Québec à Trois-Rivières (`ftp.uqtr.ca`), in the subdirectory `/pub/dphy/marchild`. They can also be reached from my Web page (`http://www.uqtr.ca/~marchild/`) where I intend to refresh the bibliography and further develop some of the ideas presented in this work.

sabbatical leave during which a large fraction of the writing was done. More personally, I warmly thank Adel F. Antippa who first trained me in theoretical physics and offered to use in his own classes the first version of the manuscript; the late Feza Gürsey, whose clear and profound thinking never ceased to inspire me; and finally Lucie and Vincent, with whom this work makes so much more sense.

Trois-Rivières, August 1999 *L. Marchildon*

* * * * * * * * * * * * *

In preparing the English version I have followed the original quite closely, except for a few improvements in formulation and correction of misprints. English versions of references have been given where possible.

I thank Aline Simoneau and Ghislaine Ferschke for their help in preparing the manuscript.

Trois-Rivières, February 2002 *L. Marchildon*

Contents

Units and Constants

The International System of Units (SI) is used in this book. Dynamical variables introduced here can be expressed in terms of five fundamental units: the meter (m), kilogram (kg), second (s), ampere (A) and kelvin (K).

Table A lists values of the most important constants we shall use. More complete tables, together with uncertainties, can be found in [21] and [52].

Table A. Fundamental constants and conversion factors

Quantity	Symbol	Value
Speed of light in vacuum	c	$2.997\,924\,58 \times 10^{8}\,\mathrm{m\,s^{-1}}$
Electric constant	ε_0	$8.854\,187\,817 \times 10^{-12}\,\mathrm{F\,m^{-1}}$
Planck constant/2π	$\hbar$	$1.054\,572\,66 \times 10^{-34}\,\mathrm{J\,s}$
Electron charge	q_e	$-1.602\,177\,33 \times 10^{-19}\,\mathrm{C}$
Electron mass	m_e	$9.109\,389\,7 \times 10^{-31}\,\mathrm{kg}$
Proton mass	m_p	$1.672\,623\,1 \times 10^{-27}\,\mathrm{kg}$
Neutron mass	m_n	$1.674\,928\,6 \times 10^{-27}\,\mathrm{kg}$
Fine-structure constant	α	$(137.035\,989\,5)^{-1}$
Bohr radius	a_0^∞	$5.291\,772\,49 \times 10^{-11}\,\mathrm{m}$
Bohr magneton	μ_B	$9.274\,015\,4 \times 10^{-24}\,\mathrm{J\,T^{-1}}$
Electron gyromagnetic ratio	g_e	$2.002\,319\,304\,386$
Electron volt	eV	$1.602\,177\,33 \times 10^{-19}\,\mathrm{J}$
Boltzmann constant	k	$1.380\,658 \times 10^{-23}\,\mathrm{J\,K^{-1}}$
		$8.617\,385 \times 10^{-5}\,\mathrm{eV\,K^{-1}}$
Rydberg	Ry	$2.179\,874\,1 \times 10^{-18}\,\mathrm{J}$
		$13.605\,698\,1\,\mathrm{eV}$

We point out that atomic distances are often expressed in ångströms (1 Å $= 10^{-10}$ m) or in multiples of the Bohr radius (a_0^∞). Atomic energies are most commonly expressed in electron volts (eV), in rydbergs (Ry) or in hartrees (= 2 Ry).

In many books on fundamental physics, as well as in a number of scientific journals, the Gaussian system of units is used, which is well adapted to the analysis of microscopic phenomena. Fundamental mechanical units are then the gram, centimeter and second. The unit of electric charge, called the statcoulomb, represents a charge which, when 1 cm away from an identical charge, exerts on it a force of 1 dyne. The statcoulomb can be expressed in terms of mechanical units as

$$1\,\mathrm{statcoul} = 1\,\mathrm{dyne}^{1/2}\mathrm{cm} = 1\,\mathrm{g}^{1/2}\mathrm{cm}^{3/2}\mathrm{s}^{-1}.$$

The constants ε_0 and μ_0 do not appear in Maxwell's equations when they are written in Gaussian units.

The relation between Gaussian and International Units is given in [21] and in the appendix of [123]. Table B summarizes operations useful for the transformation, into Gaussian units, of an equation written in the International System.

Table B. From the International System to Gaussian units

Dynamical variable	Transformation
Charge density	$\varrho \longrightarrow \sqrt{4\pi\varepsilon_0}\,\varrho$
Current density	$\boldsymbol{j} \longrightarrow \sqrt{4\pi\varepsilon_0}\,\boldsymbol{j}$
Electric field	$\boldsymbol{E} \longrightarrow \frac{1}{\sqrt{4\pi\varepsilon_0}}\,\boldsymbol{E}$
Magnetic induction	$\boldsymbol{B} \longrightarrow \sqrt{\frac{\mu_0}{4\pi}}\,\boldsymbol{B}$
Vector potential	$\boldsymbol{A} \longrightarrow \sqrt{\frac{\mu_0}{4\pi}}\,\boldsymbol{A}$
Scalar potential	$\phi \longrightarrow \frac{1}{\sqrt{4\pi\varepsilon_0}}\,\phi$
Magnetic moment	$\boldsymbol{\mu} \longrightarrow \sqrt{\frac{4\pi}{\mu_0}}\,\boldsymbol{\mu}$

1 A Crisis in Classical Physics

The formalism of quantum mechanics was proposed in 1925 and 1926, chiefly by W. Heisenberg, E. Schrödinger and P. A. M. Dirac. The key to its interpretation was given by M. Born in 1926. At the outset, creators of quantum mechanics presented it as a fundamental theory of atoms and molecules.

Although classical mechanics had met, ever since I. Newton, with incomparable success, at the turn of the twentieth century it appeared less and less capable of describing microscopic phenomena. In this opening chapter, we recall some of the most significant episodes in this story.[1]

1.1 The Reality of Atoms

The idea that every substance is made of extremely small, indivisible and unchangeable particles goes back to antiquity. *Atomism*, in its different versions, was not however the leading conception of Greek thinkers, much less of the scholastics. The idea nonetheless strongly resurfaced in the seventeenth century, within the period's mechanistic and anti-Aristotelian context. Let us see how Newton ([168], pp. 400–401) formulates it in the ultimate query of his celebrated Opticks.

> [It] seems probable to me, that God in the beginning formed matter in solid, massy, hard, impenetrable, movable particles. [...] [Particles] even so very hard, as never to wear or break in pieces; [...] [The] changes of corporeal things are to be placed only in the various separations and new associations and motions of these permanent particles. [...] It seems to me farther, that these particles [...] are moved by certain active principles, such as is that of gravity, and that which causes fermentation, and the cohesion of bodies.

The proposal that active principles – today's forces – move bodies according to precise equations was particularly fruitful in the realm of celestial phenomena, where Newtonian dynamics explained so well the motion of pla-

---

[1] Expectably the story is rich and complex, and we can only give here a brief overview. The reader is referred to general expositions by Segrè [207, 206], Rosmorduc [197] and Mathieu [155], to more specialized studies by Jammer [125], Pais [174, 175] and Rechenberg [193] as well as to the first chapters of Eisberg [76].

L. Marchildon, *Quantum Mechanics*

nets, their satellites and comets. Due to the complex structure of material bodies and the lack of adequate tools for investigation of their properties, the atomic hypothesis could not at once be applied to them with such success.

The atomic hypothesis had two important backings in the nineteenth century. The first came from chemists. At the beginning of the century, it became clearer and clearer that substances tend to combine in fixed proportions to form compounds. In the case where two different weights of A can combine to a given weight of B, J. Dalton noted that the ratio of the weights of A is a simple fraction. Dalton proposed that compounds can be resolved in different atoms that can only combine in integral numbers.

At the same period, J. L. Gay-Lussac observed that when gases combine, the ratio of their volumes is a simple fraction. In 1811 A. Avogadro proposed that at a given pressure and temperature, a given volume of any gas whatsoever always contains the same number of molecules. Avogadro's hypothesis was only accepted half a century later. His contemporaries, unlike himself, did not understand that a molecule of a simple gas can contain more than one atom.

The first half of the nineteenth century also witnessed a remarkable growth of thermodynamics. Several physicists, in particular R. Clausius, J. C. Maxwell and L. Boltzmann, developed the idea that thermodynamic properties can be explained through the atomic hypothesis. In that kinetic theory of gases, pressure comes from particle collisions on the container's walls and temperature is related to the particles' average energy. One can, in principle, compute the relation between the pressure, temperature and volume of a gas from simple statistical hypotheses and from the knowledge of forces that particles exert on one another.

Inasmuch as one believes in the genuine existence of atoms, the question of their size immediately comes up. On this Dalton's hypothesis is of no help since reduction of all atomic weights by the same factor leaves it with the same explanatory power. Similarly, a given pressure can be produced by small atoms making numerous collisions with the walls as easily as by big atoms that collide less often. But in 1860 Maxwell, using the concept of mean free path introduced earlier by Clausius, realized that gas diffusion, thermal conductivity and viscosity depend on atomic sizes. The quantitative investigation of these phenomena therefore gives a hold on atomic parameters or, equivalently, on Avogadro's number. First estimates left many physicists and chemists, and not the least, skeptical. For various scientific or philosophical reasons many would see in atoms nothing but a model which, although useful, did not have true physical reality. At the turn of the twentieth century, however, new methods emerged for the determination of atomic sizes. A. Einstein's famous analysis of Brownian motion is one of half a dozen such methods he himself developed. Values of Avogadro's number deduced from those independent methods all fell between 6 and 9×10^{23}. By 1910, the scientific community was convinced of the existence of atoms.

1.2 The Reality of Field

In his treatise on optics from which we quoted, Newton developed a corpuscular theory of light. This hypothesis naturally accounts for rectilinear propagation of light and for the laws of reflection. It also accounts for refraction, provided the particles of light undergo an acceleration when they enter a more refractive medium. The corpuscular theory does not so easily explain, on the other hand, the concentric rings observed when a low-curvature spherical lens is placed on a glass plate. Newton had to introduce an additional hypothesis according to which light successively undergoes fits of easy reflection and fits of easy transmission.

Due to its author's immense fame, the corpuscular theory of light eclipsed in the eighteenth century the wave theory of C. Huygens, a contemporary of Newton. But early in the nineteenth century, T. Young suggested that light, like sound, is a longitudinal wave. Young had observed the interference pattern of a ray of light going through two narrow slits. He noted that light is more intense where the two trajectories differ by an integral multiple of a certain length. He associated with red and blue light wavelengths of 0.7 and 0.4 μm respectively.

Independently of Young, work in the same vein was carried out in France by A. J. Fresnel about 1815. Fresnel went on to study the phenomenon of double refraction. Many had unsuccessfully attempted to produce interference fringes with the two rays. Fresnel proposed that vibrations of light are transverse rather than longitudinal and that rays coming out of double refraction vibrate in perpendicular directions. This hypothesis also explained E. L. Malus's recent observation that light reflected in a certain angle does not produce double refraction (what we call polarization by reflection). In spite of this it first met with strong opposition, due to difficulties in accounting for transverse vibrations on the basis of mechanical properties of the medium in which waves were assumed to propagate.

Along with optics, electricity and magnetism made important progress at the beginning of the nineteenth century. The pile just invented by A. Volta allowed for the first time the production of significant currents. H. C. Oersted observed in 1820 that a current in a wire deflects a nearby magnetized needle. This was the first time an electric effect on magnetism was observed. The converse effect, suspected by many, was observed eleven years later by M. Faraday. To him we owe the notion of field. Looking at lines made by iron filings near a magnet, Faraday was convinced that the magnetic effect is everywhere in space rather than acting at a distance in Newton's way.

In the hands of Maxwell this idea would become extraordinarily fruitful. In his 1873 treatise, Maxwell gave a general theory of electromagnetic phenomena. He showed that his equations admit solutions where the electric and magnetic fields propagate as transverse waves at a speed numerically equal to the speed of light. He thus proposed to identify light with an electromagnetic wave of appropriate frequency. In his exposition, he kept the ether as the medium in which electromagnetic waves propagate and the notion of real polarization to explain the displacement current. But he did not so much rely on such mechanical models as he, Faraday and W. Thomson (later Lord Kelvin) had elaborated in the previous decades.

The reality of electromagnetic waves was revealed in 1886 by H. Hertz. By means of a very-high-frequency electric oscillator Hertz produced waves which he detected at a distance. He showed these waves obey the laws of reflection and propagate at the speed of light. A few years later, H. A. Lorentz would considerably clarify Maxwell's theory. Lorentz proposed that the source of electromagnetic field lies in microscopic particles that he called *electrons*, and which move in the ether. The electrodynamics of moving bodies would reduce, in principle, to a theory of electrons.

At the end of the nineteenth century, the ether seemed to play an important role in electromagnetism. Several experiments were carried out to detect it. They all failed. In 1905 Einstein proposed the theory of special relativity. The ether became useless and the electromagnetic field no longer needed a medium for propagation.

1.3 The Discreteness of Energy

Energy is one of these concepts that have evolved over a long period. Only in the middle of the nineteenth century was the law of conservation of energy recognized. Definition and conservation of energy are connected, as energy is conserved only if all its different forms are reckoned with.

In classical physics, a system's energy is a real variable and its allowed values fill a continuum. If a projectile can be thrown with energies of 1 or 2 joules, nothing prevents ejecting it with any intermediate value. The same applies to waves. The energy of an electromagnetic wave, for instance, is proportional to the square of the electric field amplitude and the latter, in principle, is not in any way restricted.

About 1815 J. Fraunofer observed a number of narrow dark lines in sunlight decomposed by a prism. He noted they were stable and wrote down the position of the most important ones. Analogous bright lines were also observed in decomposed light originating from some flames or sparks. Nevertheless several decades went by before R. W. Bunsen and G. R. Kirchhoff realized that such lines characterize the emitting substance. They took the spectra of many elements. Spectral analysis received spectacular confirmation when they announced they had in this way discovered two new elements, which they called cesium and rubidium. Several other elements were found similarly in the following years.

As spectroscopy was developing, people realized that elements have in general extremely complex spectra. Numerical values of all these wavelengths did not seem to obey simple laws. There was, however, one exception. In 1885

J. J. Balmer showed that wavelengths of the twelve known lines of hydrogen were all given by the following formula, where n is an integer larger than 2:

$$\lambda = \left\{ \frac{n^2}{n^2 - 2^2} \right\} \times 3\,645.6\,\text{Å}\,. \tag{1.1}$$

More complicated, and less successful, relations were afterward proposed for certain elements like the alkali atoms. At any rate, atomic spectra would get more and more complicated, with the observation of fine structure and of the splitting of rays in a magnetic field (the Zeeman effect) or in an electric field (the Stark effect).

Complexity hints at structure and fortunately, structure was beginning to show up inside atoms. In 1897 J. J. Thomson established the corpuscular nature of cathode rays and measured their charge to mass ratio. He showed this ratio does not depend on the cathode material. The rays were thus negatively-charged particles common to all elements. Since atoms are electrically neutral, there must also have been positive charges, which were first supposed to be spread more or less uniformly inside the atom. But in 1910 E. Rutherford, observing the scattering of α particles through thin metal foils, inferred that the positive charge was concentrated in a much smaller volume than the atom's total volume. The structure of the atom was thus getting clearer: a nucleus with positive charge and small size, and much lighter electrons revolving about the nucleus under the electrostatic force, much like planets revolve about the sun under the gravitational force.

As with planets nothing, in Rutherford's model, fixes the radius of electronic orbits. It was suspected that the hydrogen atom, the simplest of all, contains only one electron. In 1913 N. Bohr solved the problem of the radius and accounted for the hydrogen spectrum with simple yet bold hypotheses. One of them amounts to assume that the electron's angular momentum, in its motion about the nucleus, can only take values that are integral multiples of a constant $\hbar$, to which we will soon come back. A simple argument then shows that the electron's total energy is given by

$$E_n = -\left(\frac{q_e^2}{4\pi\varepsilon_0} \right)^2 \frac{m}{2\hbar^2 n^2}\,, \tag{1.2}$$

where q_e and m are the electron's charge and mass and n is a positive integer.[2] According to Bohr, the electron's energy can thus take discrete values only.

Bohr also assumed that radiation is emitted when the electron goes from an energy level n to a level $n' < n$. Radiation energy E is then equal to the difference $E_n - E_{n'}$. Borrowing Einstein's hypothesis according to which radiation energy and wavelength are related through the equation

---

[2] More precisely, m is the reduced mass of the electron, which we will define in Chap. 9. We use modern instead of original notations.

$$E = \frac{hc}{\lambda}, \tag{1.3}$$

where $h = 2\pi\hbar$, Bohr easily reproduced Balmer's formula by setting $n' = 2$. Bohr's model also explained other line series in hydrogen, as well as the spectrum of ionized helium.

In spite of its remarkable quantitative success, Bohr's model confronted classical theory on two fronts: discrete values of energy, and also stability of orbits. Indeed Maxwell's electrodynamics predicts that an accelerated charge necessarily radiates electromagnetic energy. But the electron in the hydrogen atom, in its circular motion about the nucleus, undergoes a substantial centripetal acceleration. Quantitative evaluation shows that due to energy loss, the electron should collapse on the nucleus in a fraction of a second. The stability of orbits, assumed by Bohr, was therefore incompatible with electromagnetic theory.

The problem of stability shows up, of course, in heavier atoms also. Wavelengths in line spectra of these atoms are again related to differences of discrete energy levels. In contrast with hydrogen though, values of these energies remained unexplained.

In Bohr's model, radius and angular momentum, in addition to energy, have discrete values only. Detailed study of atomic spectra, the Zeeman effect in particular, revealed that discreteness of angular momentum plays an essential role. It was directly demonstrated in 1922, by an experiment we will examine in detail in Chap. 4.

The discreteness of energy, unexpected in classical mechanics, appeared spectacularly in atomic spectra. But it also showed up in several other phenomena.

About 1860 Kirchhoff showed, by thermodynamic arguments, that radiation from a perfect emitter (which is also a perfect absorber, that is, a *blackbody*) does not depend on details of its structure. Specifically, the energy emitted per unit frequency is a universal function of frequency and temperature. It thus became important to measure that function, and to be able to compute it.

Although the concept of blackbody is an idealization, it can be realized approximately by means of a small hole in a cavity. External radiation falling on the hole is absorbed by the cavity's internal walls, after a few reflections perhaps. Brought to a temperature T, the cavity essentially becomes a perfect emitter.

Several theoretical computations of the cavity's emission spectrum were proposed at the end of the nineteenth century. The one by Lord Rayleigh is probably the most signifiant. Rayleigh first evaluated the distribution of the electromagnetic field's resonant modes inside a cavity. He then assumed, following the principle of equipartition of energy developed by Maxwell and Boltzmann, that on the average each mode stores an energy equal to kT, where k is Boltzmann's constant. In this way he found that $I(\omega)$, the radiation intensity in the cavity per unit angular frequency, is proportional to ω^2.. The result is simple but, independently of experimental values, unacceptable. It implies that total radiation has infinite power.

Between 1895 and 1900, the function $I(\omega)$ was studied experimentally by several investigators, in various frequency and temperature intervals. Looking for a curve that would fit experimental results adequately, M. Planck in October 1900 proposed the following formula:

$$I(\omega) = \frac{\hbar\omega^3}{\pi^2 c^2}\left[\exp\left(\frac{\hbar\omega}{kT}\right) - 1\right]^{-1}. \tag{1.4}$$

There remained to justify it, which Planck did two months later. Planck assumed that the energy of a mode of vibration in the cavity is restricted to an integral multiple of $\hbar\omega$. He next showed that the average energy of a mode is different from kT, and the result he found allowed him to recover (1.4) exactly. Planck's paper is in fact the first that brought forth the discreteness of energy. The constant $\hbar$, or more precisely h, bears his name.

We pointed out that Hertz was the first to generate high-frequency electromagnetic waves. He also observed that ultraviolet light stimulated spark production in his experimental setup. The phenomenon is a particular case of the *photolectric effect*, which consists in the production of an electric current by the illumination of a conductor. About 1900 J. J. Thomson and P. Lenard showed that charges emitted by the conductor are electrons. Lenard also observed that current is produced only if the light frequency is above a threshold that depends on the material used, and that the electrons' energy is independent of light intensity. For frequency above threshold, the current appears to be produced instantaneously.

Such results were surprising. On the basis of classical electrodynamics, one rather expected the current not to show up immediately (i.e. it should take time for electrons to absorb enough energy), but to show up for whatever frequency of the incoming light.

The correct explanation was given by Einstein, and once more it involved discreteness of energy. On the basis of a penetrating analysis of the entropy associated with radiation, Einstein proposed that the energy of light is carried by quanta with energy equal to $\hbar\omega$, i.e. the value given in (1.3). As long as this is not enough to eject an electron, there is no current produced, as the electron rapidly looses the energy it just acquired. If, on the other hand, $\hbar\omega$ is large enough, the effect is instantaneous. Because the electron normally absorbes one photon only, its energy does not depend on the intensity of light.

The discreteness of energy showed up once more in the study of specific heats. In 1819 P. L. Dulong and A. T. Petit observed that for a number of metals, as well as sulfur, the specific heat (per unit atomic weight) is the same. This could be explained by the principle of equipartition of energy. But what about the atoms' internal degrees of freedom? They appeared not to contribute to specific heats either of solids or of gases. Monoatomic gases, for instance, have the specific heat of point particles. In addition, H. F. Weber observed about 1870 that the specific heat decreases with temperature. The behavior of low-temperature solids was explained by Einstein, and later by P. Debye, through the assumption that atoms vibrate like Planck oscillators which, for a given frequency, have only discrete values of energy.

1.4 The Nature of Atomic Objects

We saw that in the nineteenth century, the corpuscular theory of light was replaced by a wave theory. The phenomena of interference and diffraction, Maxwell's synthesis and its experimental verification by Hertz all pointed toward a wave theory. In 1849 A. H. Fizeau directly measured the speed of light in a refractive medium. He found a numerical value lower that the speed of light in empty space, which ran against the prediction of Newton's theory. At the end of the nineteenth century, no one doubted that light is a wave.

The hypothesis of light quanta, formulated by Einstein in 1905, thus came out as a revolution. It seemed to restore light's corpuscular character. At first few investigators believed it, even though corroboration would slowly build up. But in 1922 A. H. Compton gave the hypothesis a spectacular confirmation. Analyzing the interaction of x-rays with electrons, Compton showed that this process obeys the relativistic laws of particle collisions.

The situation thus challenged all rational analysis. Interference, diffraction and Hertz's experiments showed that light is a wave, whereas the photoelectric and Compton effects showed that it is a corpuscle. The ambiguity could be seen even in the formula $E = \hbar\omega$, where E denotes the corpuscle's energy while ω stands for the wave's frequency.

In 1924 L. de Broglie pushed the idea further. No one had hitherto doubted that electrons are corpuscles. Nevertheless their quantization rules, as prescribed by Bohr, involved integral numbers. In de Broglie's mind, integral numbers characterized phenomena such as interference and normal modes of vibration. Using relativistic reasoning, he proposed that the electron, like the photon, has wave properties and he associated with the electron a wavelength related to its momentum by the equation

$$\lambda = \frac{h}{p}\,. \tag{1.5}$$

De Broglie did not know that interference of electron waves had been observed (without however being appreciated) soon before by C. Davisson and C. Kunsman. The correct interpretation was given by W. Elsasser in 1925. The nature of electrons, like that of light, appeared incomprehensible.

1.5 At the Threshold of the Quantum Revolution

The end of the nineteenth and the first quarter of the twentieth century witnessed numerous signs of major problems in classical theory. One could not account for the line spectra of elements, and atomic stability itself was a mystery. Blackbody radiation and the photoelectric effect were explained by assumptions that could not be reconciled with electrodynamics. Properties of light and of electrons seemed inconsistent. Classical theory, to borrow the terminology of T. S. Kuhn [135], was in a deep crisis.

Such a crisis is not, in general, resolved by cosmetic changes in the theory. It foretells radical transformations. These came in 1925–26. Quantum mechanics, as the new theory was called, proposed an explanation of atomic phenomena completely different from classical theory.

In the coming chapters we will develop the formalism of quantum mechanics, its interpretation and a number of applications.

Two kinds of exposition were discarded: the historical and the inductive. The second one leads nowhere. Even through experiment is an indispensable guide in the elaboration of theories – we will see an example in Chap. 4 –, a scientific theory is not extracted directly from empirical laws. The historical way, on the other hand, has much interest. Unfortunately, it is not always consonant with pedagogical objectives. The best way to assimilate a complex theory is not to retrace the long and tortuous path that outstanding creators have trod.

Our exposition will in a sense be axiomatic. Not as in a mathematical treatise which attempts to deduce, with complete rigor, each proposition from a small number of independent statements. Rather, quantum theory will be proposed as a hypothesis, as a set of postulates whose justification is not a priori obvious but follows from the agreement between consequences drawn from them and experimental results.

Quantum concepts, very different from common sense, are best assimilated if presented in a self-contained way. Different aspects of the connection with classical concepts will be investigated in Chaps. 6, 16 and 21. Readers interested in the historical development of quantum mechanics may, once familiar with the main formalism, go back to the excellent book by Jammer [125].

2 Finite-Dimensional Vector Spaces

The theory of vector spaces and of operators defined in them is the fundamental mathematical tool of quantum mechanics. This chapter summarizes, usually without proofs, the properties of finite-dimensional vector spaces.[1] Readers familiar with these results can skip to Chap. 3, after a glance at the notations we introduce.

2.1 Vector Spaces and Scalar Product

Let $\mathcal{V}$ be a set and $\mathcal{K}$ a field (in general, the field of real numbers $\mathcal{R}$ or of complex numbers $\mathcal{C}$); let $+$ be a binary operation which, to every pair of elements in $\mathcal{V}$, associates an element in $\mathcal{V}$; let $\cdot$ be a binary operation which, to an element in $\mathcal{K}$ and an element in $\mathcal{V}$, associates an element in $\mathcal{V}$. Suppose that for all $|\phi\rangle$, $|\psi\rangle$ and $|\chi\rangle$ in $\mathcal{V}$ and α and β in $\mathcal{K}$,

i) $|\phi\rangle + |\psi\rangle = |\psi\rangle + |\phi\rangle$;
ii) $(|\phi\rangle + |\psi\rangle) + |\chi\rangle = |\phi\rangle + (|\psi\rangle + |\chi\rangle)$;
iii) there exists a $|0\rangle$ in $\mathcal{V}$ such that $|0\rangle + |\phi\rangle = |\phi\rangle$;
iv) $0 \cdot |\phi\rangle = |0\rangle$;
v) $1 \cdot |\phi\rangle = |\phi\rangle$;
vi) $\alpha \cdot (|\phi\rangle + |\psi\rangle) = \alpha \cdot |\phi\rangle + \alpha \cdot |\psi\rangle$;
vii) $(\alpha + \beta) \cdot |\phi\rangle = \alpha \cdot |\phi\rangle + \beta \cdot |\phi\rangle$;
viii) $(\alpha\beta) \cdot |\phi\rangle = \alpha \cdot (\beta \cdot |\phi\rangle)$.

Then the quadruplet $(\mathcal{V}, \mathcal{K}, +, \cdot)$ is called a *vector space.* Elements of $\mathcal{V}$ are called *vectors.* Binary operations $+$ and $\cdot$ are called *vector addition* and *scalar multiplication*, respectively. The vector space is *real* if $\mathcal{K} = \mathcal{R}$ and *complex* if $\mathcal{K} = \mathcal{C}$.

Actually, the symbol $\mathcal{V}$ is often used to denote not only the set of vectors but also the vector space, that is, the quadruplet. The symbol $\cdot$ is usually omitted.

---

[1] More on concepts introduced here, as well as proofs of theorems, can be found in [81] and [194].

L. Marchildon, *Quantum Mechanics*

Examples The set of all columns of n real numbers, together with the usual definition of addition and multiplication by a real scalar, makes up a vector space denoted by $\mathcal{R}^n$. A vector space $\mathcal{C}^n$ is defined similarly, with complex instead of real numbers. ♣[2]

Unless otherwise indicated, we will always have in mind complex vector spaces.

Let $(\mathcal{V}, \mathcal{K}, +, \cdot)$ be a vector space and let $\mathcal{V}'$ be a subset of $\mathcal{V}$. The quadruplet $(\mathcal{V}', \mathcal{K}, +, \cdot)$ is called a *subspace* of $(\mathcal{V}, \mathcal{K}, +, \cdot)$ if $(\mathcal{V}', \mathcal{K}, +, \cdot)$ is itself a vector space. $(\mathcal{V}, \mathcal{K}, +, \cdot)$ and $(|0\rangle, \mathcal{K}, +, \cdot)$ are trivial subspaces of $(\mathcal{V}, \mathcal{K}, +, \cdot)$.

We say that n nonzero vectors $|\phi_1\rangle$, $|\phi_2\rangle$, …, $|\phi_n\rangle$ are *linearly independent* if for all α_1, α_2, …, α_n in $\mathcal{K}$, the equation

$$\alpha_1|\phi_1\rangle + \alpha_2|\phi_2\rangle + \cdots + \alpha_n|\phi_n\rangle = |0\rangle \tag{2.1}$$

is true only if $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$. The *dimension* of a vector space $\mathcal{V}$ is the maximum number of linearly independent vectors that one can find in $\mathcal{V}$. Let N be the dimension of $\mathcal{V}$. Any set of N linearly independent vectors is called a *basis* of $\mathcal{V}$. In general, N can be finite or infinite. In this chapter, however, N will always be finite.

Let $\mathcal{V}$ be an N-dimensional vector space, and let $\{|u_i\rangle,\ i = 1, 2, \ldots, N\}$ be a basis of $\mathcal{V}$. Let $|\psi\rangle$ be an arbitrary nonzero vector in $\mathcal{V}$. Necessarily, $|\psi\rangle$ and the $|u_i\rangle$ are linearly dependent. Thus we can write

$$|\psi\rangle = \sum_{i=1}^{N} c_i |u_i\rangle\,, \tag{2.2}$$

where the c_i are complex constants. We say that the vector $|\psi\rangle$ is *expanded* in a linear combination of basis vectors. The constants c_i are called *components* of the vector $|\psi\rangle$ in the basis of the $|u_i\rangle$. It is very important to realize that the components depend on the basis in which the vector is expanded.

Consider a binary operation which, to any two elements $|\phi\rangle$ and $|\psi\rangle$ in $\mathcal{V}$, associates a number $(|\phi\rangle, |\psi\rangle)$ (real or complex according as $\mathcal{K} = \mathcal{R}$ or $\mathcal{C}$). The operation is called a *scalar product* if, for all $|\phi\rangle$, $|\psi\rangle$, $|\psi_1\rangle$ and $|\psi_2\rangle$ in $\mathcal{V}$ and α_1 and α_2 in $\mathcal{K}$,

i) $(|\phi\rangle, |\psi\rangle) = (|\psi\rangle, |\phi\rangle)^*$;
ii) $(|\phi\rangle, \alpha_1|\psi_1\rangle + \alpha_2|\psi_2\rangle) = \alpha_1 (|\phi\rangle, |\psi_1\rangle) + \alpha_2 (|\phi\rangle, |\psi_2\rangle)$;
iii) $(|\phi\rangle, |\phi\rangle) \geq 0$;
iv) $(|\phi\rangle, |\phi\rangle) = 0$ if and only if $|\phi\rangle = |0\rangle$.

The star denotes complex conjugation, which is trivial if $\mathcal{K} = \mathcal{R}$. Property (ii) expresses the *linearity* of the scalar product in terms of the second vector. Note that in terms of the first vector the scalar product is *antilinear*. Indeed

---

[2] The symbol ♣ indicates the end of an example or of a proof.

$$\begin{aligned}(\alpha_1|\phi_1\rangle + \alpha_2|\phi_2\rangle, |\psi\rangle) &= (|\psi\rangle, \alpha_1|\phi_1\rangle + \alpha_2|\phi_2\rangle)^* \\ &= \{\alpha_1 (|\psi\rangle, |\phi_1\rangle) + \alpha_2 (|\psi\rangle, |\phi_2\rangle)\}^* \\ &= \alpha_1^* (|\psi\rangle, |\phi_1\rangle)^* + \alpha_2^* (|\psi\rangle, |\phi_2\rangle)^* \\ &= \alpha_1^* (|\phi_1\rangle, |\psi\rangle) + \alpha_2^* (|\phi_2\rangle, |\psi\rangle) .\end{aligned}$$

From the scalar product, the *norm* of $|\phi\rangle$ is defined as $(|\phi\rangle, |\phi\rangle)^{1/2}$. Two vectors $|\phi\rangle$ and $|\psi\rangle$ are *orthogonal* if their scalar product vanishes, that is, if $(|\phi\rangle, |\psi\rangle) = 0$. Let N be the dimension of $\mathcal{V}$. The N vectors $|u_1\rangle$, $|u_2\rangle$, ..., $|u_N\rangle$ make up a *basis of orthogonal vectors* if they form a basis and if all vectors are pairwise orthogonal. Moreover, if each $|u_i\rangle$ is *normalized*, which means that $(|u_i\rangle, |u_i\rangle) = 1$ for every i, the basis is *orthonormal*. For the rest of this chapter, we will assume that $\mathcal{V}$ has dimension N, that a scalar product has been defined and that $|u_1\rangle, |u_2\rangle, \ldots, |u_N\rangle$ make up an orthonormal basis.

A *linear functional* is a function which associates to any element $|\psi\rangle$ in $\mathcal{V}$ an element $f(|\psi\rangle)$ in $\mathcal{K}$ so that, for all $|\psi_1\rangle$ and $|\psi_2\rangle$ in $\mathcal{V}$ and α_1 and α_2 in $\mathcal{K}$,

$$f\{\alpha_1|\psi_1\rangle + \alpha_2|\psi_2\rangle\} = \alpha_1 f(|\psi_1\rangle) + \alpha_2 f(|\psi_2\rangle). \tag{2.3}$$

The set of all linear functionals make up a vector space if addition and scalar multiplication are defined so that

$$[\beta_1 f_1 + \beta_2 f_2](|\psi\rangle) = \beta_1 f_1(|\psi\rangle) + \beta_2 f_2(|\psi\rangle). \tag{2.4}$$

It can be checked that all defining properties of a vector space are indeed satisfied.

With the help of the scalar product one can establish a one-to-one correspondence between vectors in $\mathcal{V}$ and linear functionals on $\mathcal{V}$. On the one hand, for any $|\phi\rangle$ in $\mathcal{V}$ it is clear that $f_{|\phi\rangle}(|\psi\rangle) \equiv (|\phi\rangle, |\psi\rangle)$ is a linear functional. On the other hand, let f be an arbitrary linear functional and let $\{|u_i\rangle\}$ be an orthonormal basis. Let us denote $f(|u_i\rangle)$ by c_i and let us form the vector

$$|\phi\rangle = \sum_{i=1}^{N} c_i^* |u_i\rangle . \tag{2.5}$$

It is easy to check that f coincides with $f_{|\phi\rangle}$.

For simplicity, we will denote by $\langle\phi|$ the linear functional $f_{|\phi\rangle}$. The vector space of linear functionals on $\mathcal{V}$ is called the *dual space* of $\mathcal{V}$ and it is denoted by $\mathcal{V}^*$. Vectors in $\mathcal{V}$ are called *kets*, those in $\mathcal{V}^*$ *bras*. Spaces $\mathcal{V}$ and $\mathcal{V}^*$ have the same dimension. The action of $f_{|\phi\rangle}$ on $|\psi\rangle$ is conveniently written in the notation $\langle\phi|\psi\rangle$ which, therefore, also represents the scalar product $(|\phi\rangle, |\psi\rangle)$. Suppose that bras $\langle\phi_1|$ and $\langle\phi_2|$ correspond to kets $|\phi_1\rangle$ and $|\phi_2\rangle$. The properties of the scalar product then imply that the bra $\alpha_1^*\langle\phi_1| + \alpha_2^*\langle\phi_2|$ corresponds to the ket $\alpha_1|\phi_1\rangle + \alpha_2|\phi_2\rangle$.

We now derive an important property of the scalar product, called the *Schwartz inequality*.

Lemma Let $|\phi\rangle$ and $|\psi\rangle$ be two arbitrary vectors. Then

$$|\langle\phi|\psi\rangle|^2 \le \langle\phi|\phi\rangle\langle\psi|\psi\rangle . \tag{2.6}$$

Proof If one vector vanishes, equality is obvious. Assume now that $|\phi\rangle \neq |0\rangle \neq |\psi\rangle$. Let us write

$$|\chi\rangle \equiv |\phi\rangle - \alpha|\psi\rangle \equiv |\phi\rangle - \frac{\langle\psi|\phi\rangle}{\langle\psi|\psi\rangle}|\psi\rangle .$$

Clearly

$$\begin{aligned} 0 &\le \langle\chi|\chi\rangle \\ &\le \{\langle\phi| - \alpha^*\langle\psi|\}\{|\phi\rangle - \alpha|\psi\rangle\} \\ &\le \langle\phi|\phi\rangle - \alpha\langle\phi|\psi\rangle - \alpha^*\langle\psi|\phi\rangle + \alpha\alpha^*\langle\psi|\psi\rangle \\ &\le \langle\phi|\phi\rangle - \frac{\langle\psi|\phi\rangle}{\langle\psi|\psi\rangle}\langle\phi|\psi\rangle . \end{aligned}$$

Therefore

$$|\langle\phi|\psi\rangle|^2 = \langle\psi|\phi\rangle\langle\phi|\psi\rangle \le \langle\phi|\phi\rangle\langle\psi|\psi\rangle .$$ ♣

2.2 Operators

An *operator* O is an object which, to an arbitrary element of $\mathcal{V}$ (or of a subset of $\mathcal{V}$) associates an element of $\mathcal{V}$. We write

$$|\psi'\rangle = O|\psi\rangle \tag{2.7}$$

and we say that O *acts in the vector space* $\mathcal{V}$, or that O is *defined* in $\mathcal{V}$.

An operator O is *linear* if it is defined on all of $\mathcal{V}$ and if, for all $|\psi_1\rangle$ and $|\psi_2\rangle$ in $\mathcal{V}$ and α_1 and α_2 in $\mathcal{K}$,

$$O\{\alpha_1|\psi_1\rangle + \alpha_2|\psi_2\rangle\} = \alpha_1 O|\psi_1\rangle + \alpha_2 O|\psi_2\rangle . \tag{2.8}$$

One easily sees that if two linear operators transform in the same way the vectors of a given basis, these operators coincide.

An important class of linear operators can be defined through bras and kets. Recall that $\langle\phi|\psi\rangle$ stands for a scalar product. Let us now introduce the notation $|\psi\rangle\langle\phi|$ to represent the operator that acts on an arbitrary vector $|\chi\rangle$ in the following way:

$$(|\psi\rangle\langle\phi|)\,|\chi\rangle \equiv (\langle\phi|\chi\rangle)\,|\psi\rangle = |\psi\rangle\,(\langle\phi|\chi\rangle) . \tag{2.9}$$

One checks that $|\psi\rangle\langle\phi|$ is a linear operator. That definition ensures the associativity of an arbitrary alternating string of bras and kets.

An operator O transforming $|\psi\rangle$ into $|\psi'\rangle$ can have an *inverse*. It does have an inverse if and only if (i) to every $|\psi\rangle$ there corresponds one and only one $|\psi'\rangle$ and (ii) every $|\psi'\rangle$ corresponds to a $|\psi\rangle$. The inverse O^{-1} is then defined as

$$|\psi\rangle = O^{-1}|\psi'\rangle\,. \tag{2.10}$$

The *identity* operator, denoted by I, is defined so that $I|\phi\rangle = |\phi\rangle$ for all $|\phi\rangle$. We have

$$OO^{-1} = I = O^{-1}O\,. \tag{2.11}$$

If O_1 and O_2 both have inverses, then $(O_1O_2)^{-1} = O_2^{-1}O_1^{-1}$.

Let O be a linear operator. The *adjoint* of O, denoted by $O^\dagger$, is defined so that for all $|\phi\rangle$ and $|\psi\rangle$ in $\mathcal{V}$

$$(|\phi\rangle, O|\psi\rangle) = \left(O^\dagger|\phi\rangle, |\psi\rangle\right). \tag{2.12}$$

The adjoint always exists. It is easy to show (Exercise 2.6) that $O^\dagger$ is linear, that $(O^\dagger)^\dagger = O$ and that $(O_1O_2)^\dagger = O_2^\dagger O_1^\dagger$. Note that we can write $(|\phi\rangle, O|\psi\rangle) = \langle\phi|O|\psi\rangle$. In the latter expression, O acts to the right on $|\psi\rangle$. But the expression can also be interpreted so that O acts to the left on $\langle\phi|$, provided that the vector $\langle\phi|O$ in $\mathcal{V}^*$ is the dual of $O^\dagger|\phi\rangle$. Note that

$$\begin{aligned}\langle\phi|O|\psi\rangle &= (|\phi\rangle, O|\psi\rangle) = (O|\psi\rangle, |\phi\rangle)^* \\ &= \left((O^\dagger)^\dagger|\psi\rangle, |\phi\rangle\right)^* = \left(|\psi\rangle, O^\dagger|\phi\rangle\right)^* = \langle\psi|O^\dagger|\phi\rangle^*\,.\end{aligned} \tag{2.13}$$

This relation is true for all vectors $|\phi\rangle$ and $|\psi\rangle$ and every linear operator O.

A linear operator H is *Hermitian* if $H^\dagger = H$. A linear operator U is *unitary* if U^{-1} exists and if $U^{-1} = U^\dagger$. Note that if all vectors in $\mathcal{V}$ are transformed by a unitary operator U, the scalar products are invariant. Indeed

$$(U|\phi\rangle, U|\psi\rangle) = \left(U^\dagger U|\phi\rangle, |\psi\rangle\right) = \left(U^{-1}U|\phi\rangle, |\psi\rangle\right) = (|\phi\rangle, |\psi\rangle)\,. \tag{2.14}$$

Conversely, if a linear operator preserves scalar products for all vectors in $\mathcal{V}$ (or even for all vectors in a basis of $\mathcal{V}$), that operator is unitary.

Let P be a Hermitian operator. P is called a *projector* if $PP = P$. It is easy to check that if $|\psi\rangle$ is normalized, then $|\psi\rangle\langle\psi| \equiv P_\psi$ is a projector.

Often one deals with vectors and operators that depend on parameters. These are functions $|\psi(t)\rangle$, $\langle\phi(t)|$ and $A(t)$ that, respectively, take values in the vector space $\mathcal{V}$, in the space $\mathcal{V}^*$ and in the set of all operators defined in $\mathcal{V}$. It is not difficult to show that products of such objects obey the usual rules of differentiation. For example,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left\{\langle\phi(t)|A(t)|\psi(t)\rangle\right\} &= \left\{\frac{\mathrm{d}}{\mathrm{d}t}\langle\phi(t)|\right\}A(t)|\psi(t)\rangle \\ &+ \langle\phi(t)|\left\{\frac{\mathrm{d}}{\mathrm{d}t}A(t)\right\}|\psi(t)\rangle + \langle\phi(t)|A(t)|\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle\,.\end{aligned} \tag{2.15}$$

2.3 Eigenvalues and Eigenvectors

Let L be a linear operator. A nonzero vector $|\phi\rangle$ is called an *eigenvector* of L, corresponding to the *eigenvalue* λ, if the following equation is satisfied:

$$L|\phi\rangle = \lambda|\phi\rangle\,. \tag{2.16}$$

The set of all eigenvalues of an operator is its *spectrum*.

We now prove two important theorems about eigenvectors and eigenvalues of Hermitian or unitary operators.

Theorem 1 Eigenvalues of a Hermitian operator are always real. Eigenvectors of a Hermitian operator corresponding to distinct eigenvalues are always orthogonal.

Proof Let H be Hermitian and let $|\phi_1\rangle$ and $|\phi_2\rangle$ be two eigenvectors of H corrresponding to eigenvalues λ_1 and λ_2. Let us first prove the first part of the theorem. On the one hand

$$\langle\phi_1|H|\phi_1\rangle = \langle\phi_1|\lambda_1|\phi_1\rangle = \lambda_1\langle\phi_1|\phi_1\rangle\,.$$

On the other hand

$$\begin{aligned}\langle\phi_1|H|\phi_1\rangle &= \langle\phi_1|H^\dagger|\phi_1\rangle^* = \langle\phi_1|H|\phi_1\rangle^* \\ &= \lambda_1^*\langle\phi_1|\phi_1\rangle^* = \lambda_1^*\langle\phi_1|\phi_1\rangle\,.\end{aligned}$$

By comparison, we obtain $\lambda_1 = \lambda_1^*$. To prove the second part we write

$$\begin{aligned}0 &= \langle\phi_1|H|\phi_2\rangle - \langle\phi_2|H^\dagger|\phi_1\rangle^* = \langle\phi_1|H|\phi_2\rangle - \langle\phi_2|H|\phi_1\rangle^* \\ &= \langle\phi_1|\lambda_2|\phi_2\rangle - \langle\phi_2|\lambda_1|\phi_1\rangle^* = \lambda_2\langle\phi_1|\phi_2\rangle - \lambda_1^*\langle\phi_2|\phi_1\rangle^* \\ &= (\lambda_2 - \lambda_1)\langle\phi_1|\phi_2\rangle\,.\end{aligned}$$

Therefore $\lambda_1 = \lambda_2$ or $\langle\phi_1|\phi_2\rangle = 0$. ♣

Theorem 2 Eigenvalues of a unitary operator are always complex numbers of unit norm. Eigenvectors of a unitary operator corresponding to distinct eigenvalues are always orthogonal.

Proof Let U be unitary and let $|\phi_1\rangle$ and $|\phi_2\rangle$ be two eigenvectors of U corresponding to eigenvalues λ_1 and λ_2. Let us first show that $U^\dagger|\phi_1\rangle = \lambda_1^{-1}|\phi_1\rangle$. We have

$$\lambda_1^{-1}|\phi_1\rangle = \lambda_1^{-1}U^\dagger U|\phi_1\rangle = \lambda_1^{-1}U^\dagger\lambda_1|\phi_1\rangle = \lambda_1^{-1}\lambda_1 U^\dagger|\phi_1\rangle = U^\dagger|\phi_1\rangle.$$

Thus, on the one hand,

$$\langle\phi_1|U^\dagger|\phi_1\rangle = \langle\phi_1|\lambda_1^{-1}|\phi_1\rangle = \lambda_1^{-1}\langle\phi_1|\phi_1\rangle\,.$$

On the other hand

$$\langle\phi_1|U^\dagger|\phi_1\rangle = \langle\phi_1|U|\phi_1\rangle^* = \langle\phi_1|\lambda_1|\phi_1\rangle^* = \lambda_1^*\langle\phi_1|\phi_1\rangle .$$

By comparison, we obtain $\lambda_1^* = \lambda_1^{-1}$, that is, $\lambda_1\lambda_1^* = 1$. To prove the second part of the theorem we write

$$\begin{aligned} 0 &= \langle\phi_1|U|\phi_2\rangle - \langle\phi_2|U^\dagger|\phi_1\rangle^* = \langle\phi_1|\lambda_2|\phi_2\rangle - \langle\phi_2|\lambda_1^{-1}|\phi_1\rangle^* \\ &= \lambda_2\langle\phi_1|\phi_2\rangle - (\lambda_1^{-1})^*\langle\phi_2|\phi_1\rangle^* = (\lambda_2 - \lambda_1)\langle\phi_1|\phi_2\rangle . \end{aligned}$$

Therefore $\lambda_1 = \lambda_2$ or $\langle\phi_1|\phi_2\rangle = 0$. ♣

It is easy to show that eigenvalues of a projector are always equal to 1 or 0. Indeed let P be a projector, let $|\phi\rangle$ be an eigenvector of P and let λ be the corresponding eigenvalue. We have

$$\begin{aligned} PP|\phi\rangle &= P\lambda|\phi\rangle = \lambda P|\phi\rangle = \lambda^2|\phi\rangle \\ &= P|\phi\rangle = \lambda|\phi\rangle , \end{aligned}$$

from which it follows that $\lambda^2 = \lambda$.

2.4 Matrix Representation of Linear Operators

Let L be a linear operator that transforms a vector $|\psi\rangle$ into a vector $|\bar{\psi}\rangle$, that is,

$$|\bar{\psi}\rangle = L|\psi\rangle . \tag{2.17}$$

Let $\{|u_i\rangle\}$ be an orthonormal basis. One can expand $|\psi\rangle$ and $|\bar{\psi}\rangle$ in linear combinations of basis vectors as:

$$|\psi\rangle = \sum_{i=1}^{N} c_i|u_i\rangle , \qquad |\bar{\psi}\rangle = \sum_{i=1}^{N} \bar{c}_i|u_i\rangle . \tag{2.18}$$

We are going to write coefficients $\bar{c}_i$ in terms of the c_i.

We first project (2.17) on the vector $\langle u_j|$. Here it is very useful to introduce a symbol δ_{ij}, called the *Kronecker delta*, which vanishes if $i \neq j$ and is equal to 1 if $i = j$. Thus we find, on the one hand, that

$$\langle u_j|\bar{\psi}\rangle = \langle u_j| \left[\sum_{i=1}^{N} \bar{c}_i|u_i\rangle\right] = \sum_{i=1}^{N} \bar{c}_i\langle u_j|u_i\rangle = \sum_{i=1}^{N} \bar{c}_i\delta_{ij} = \bar{c}_j .$$

On the other hand

$$\begin{aligned}\langle u_j|\bar{\psi}\rangle &= \langle u_j|L|\psi\rangle = \langle u_j|L\left[\sum_{i=1}^{N} c_i|u_i\rangle\right] \\ &= \langle u_j|\left[\sum_{i=1}^{N} c_i L|u_i\rangle\right] = \sum_{i=1}^{N} c_i\langle u_j|L|u_i\rangle = \sum_{i=1}^{N} c_i L_{ji}\,,\end{aligned}$$

where $L_{ji} = \langle u_j|L|u_i\rangle$. Therefore we obtain

$$\bar{c}_j = \sum_{i=1}^{N} L_{ji} c_i\,. \tag{2.19}$$

This is the equation we were looking for. The coefficients $\bar{c}_j$ of the expansion of $|\bar{\psi}\rangle$ in the $|u_i\rangle$ basis are written in terms of the coefficients c_i of the expansion of $|\psi\rangle$ in that same basis, and of N^2 complex numbers L_{ij}. These are called *matrix elements* of L in the $|u_i\rangle$ basis. Thus knowledge of matrix elements is enough to determine the image of every vector under the action of the linear operator L. In other words, the correspondence between the set of linear operators acting in an N-dimensional vector space and the set of $N \times N$ matrices is one-to-one. Note that if we view the N numbers $\bar{c}_j$ and the N numbers c_i as column matrices $[\bar{c}]$ and $[c]$, (2.19) can be written as

$$[\bar{c}] = [L][c]\,, \tag{2.20}$$

where $[L]$ denotes the matrix of the coefficients L_{ij} and where the right-hand side represents the matrix product of $[L]$ and $[c]$.

In the $|u_i\rangle$ basis, let us denote by c_i and d_i the components of $|\psi\rangle$ and $|\phi\rangle$, respectively. It is clear that

$$\langle\phi|\psi\rangle = \sum_{i=1}^{N} d_i^* c_i = ([d]^*)^T\,[c]\,, \tag{2.21}$$

where the exponent T denotes matrix transposition. To the bra $\langle\phi|$ one can thus associate a row matrix whose elements are the complex conjugates of the d_i. Summarizing, bras, kets and linear operators in a given basis have a one-to-one correspondence with rows, columns and square matrices, respectively.

Matrix elements of the adjoint of L and of the inverse of L (if it exists) are easily obtained from matrix elements of L. For the adjoint we have[3]

$$(L^\dagger)_{ij} = \langle u_i|L^\dagger|u_j\rangle = \langle u_j|L|u_i\rangle^* = L_{ji}^*\,. \tag{2.22}$$

For the inverse, we get from $|\psi\rangle = L^{-1}|\bar{\psi}\rangle$ that

$$c_i = \sum_{i=1}^{N} (L^{-1})_{ij}\bar{c}_j\,, \tag{2.23}$$

---

[3] Sometimes the adjoint of a matrix $[M]$ is defined as $[M]^{-1}\det([M])$. That definition is completely different from the one we use.

wherefrom we immediately conclude upon comparison with (2.19) that $(L^{-1})_{ij}$ is the (ij) element of the inverse matrix of $[L]$.

Thus the matrix of L^{-1} is the inverse of the matrix of L and the matrix of $L^\dagger$ is the transpose of the complex conjugate of the matrix of L. Hence matrix elements of a Hermitian operator H or of a unitary operator U have the following properties:

$$H_{ij} = H^*_{ji}\,, \tag{2.24}$$

$$\sum_{k=1}^{N} U_{ik}U^*_{jk} = \sum_{k=1}^{N} U_{ik}(U^\dagger)_{kj} = \sum_{k=1}^{N} U_{ik}(U^{-1})_{kj} = [I]_{ij} = \delta_{ij}\,. \tag{2.25}$$

Note also that the matrix corresponding to a product of linear operators is the matrix product of the matrices that correspond to each operator.

The symbol L_{ij} stands for the matrix element $\langle u_i|L|u_j\rangle$. Conversely, the operator L can be written in terms of the L_{ij} as (Exercise 2.10)

$$L = \sum_{i,j=1}^{N} L_{ij}|u_i\rangle\langle u_j|\,. \tag{2.26}$$

In the particular case where L is the identity operator this equation becomes

$$I = \sum_{i,j=1}^{N} \delta_{ij}|u_i\rangle\langle u_j| = \sum_{i=1}^{N} |u_i\rangle\langle u_i|\,. \tag{2.27}$$

Equation (2.27) is called a *closure relation.* It means that the $|u_i\rangle$ make up an orthonormal basis and it expresses the *resolution of the identity* in terms of the projectors $|u_i\rangle\langle u_i|$.

It is interesting to see how the matrix elements of a linear operator transform when we go from one orthonormal basis to another. Let $\{|u_i\rangle\}$ and $\{|v_i\rangle\}$ be two orthonormal bases and let U be the unitary operator transforming one into the other:

$$|v_i\rangle = U|u_i\rangle\,. \tag{2.28}$$

If U_{ij} denotes the matrix elements of U in the $|u_i\rangle$ basis, we see that[4]

$$|v_i\rangle = \left[\sum_{k=1}^{N} |u_k\rangle\langle u_k|\right] U|u_i\rangle = \sum_{k=1}^{N} U_{ki}|u_k\rangle\,. \tag{2.29}$$

The operator L can be written as

---

[4] It is important to observe the position of indices in (2.19) and (2.29). The transformation rule for vector components involves a summation on the second index of the matrix elements, whereas the transformation rule for basis vectors involves a summation on the first index.

$$L = \sum_{k,l=1}^{N} \tilde{L}_{kl} |v_k\rangle\langle v_l| = \sum_{i,j=1}^{N} L_{ij} |u_i\rangle\langle u_j|, \tag{2.30}$$

from which it follows at once that

$$[\tilde{L}] = [U^\dagger][L][U]. \tag{2.31}$$

Equation (2.31) relates matrix elements of operator L in the $|v_i\rangle$ basis to matrix elements of L in the $|u_i\rangle$ basis.

A transformation that substitutes a basis to another but leaves as they are vectors in $\mathcal{V}$ and operators is called *passive*. On the other hand, a transformation that changes all vectors in $\mathcal{V}$ through the action of a unitary operator U is called *active*. Then, for instance,

$$|\psi'\rangle = U|\psi\rangle, \qquad |\bar{\psi}'\rangle = U|\bar{\psi}\rangle. \tag{2.32}$$

If $|\psi\rangle$ and $|\bar{\psi}\rangle$ are related through the operator L and if $|\psi'\rangle$ and $|\bar{\psi}'\rangle$ are related through L', we get

$$L' = ULU^\dagger. \tag{2.33}$$

Equation (2.33) represents the effect of an active transformation on L.

The *trace* of a linear operator L is defined as

$$\mathrm{Tr}(L) = \sum_{i=1}^{N} \langle u_i|L|u_i\rangle. \tag{2.34}$$

One easily checks that the value of the trace does not depend on the orthonormal basis in which it is evaluated. Furthermore it is clear that

$$\mathrm{Tr}(L_1 L_2 L_3) = \mathrm{Tr}(L_3 L_1 L_2). \tag{2.35}$$

For all vectors $|\phi\rangle$ and $|\psi\rangle$ and every linear operator L one also finds that

$$\mathrm{Tr}\{|\phi\rangle\langle\psi|\} = \langle\psi|\phi\rangle, \tag{2.36}$$

$$\mathrm{Tr}\{|\phi\rangle\langle\psi|L\} = \langle\psi|L|\phi\rangle. \tag{2.37}$$

The *determinant* of a linear operator L is defined as the determinant of the matrix associated with L. Here also, the determinant is independent of the orthonormal basis. One can show that the determinant of a product of matrices is equal to the product of determinants. Moreover, the determinant of the transpose of M is equal to the determinant of M, and the determinant of the adjoint of M is equal to the complex conjugate of the determinant of M.

2.5 Spectral Decomposition

Eigenvectors of a Hermitian operator that correspond to distinct eigenvalues are orthogonal. But several distinct and linearly independent eigenvectors often correspond to the same eigenvalue λ. In that case one easily shows that the set of eigenvectors with the same eigenvalue λ makes up a subspace $\mathcal{V}_\lambda$ of the vector space $\mathcal{V}$.

In the subspace $\mathcal{V}_\lambda$, one can find an orthonormal basis. It is possible to show that if we do so with each eigenvalue, we finally obtain an orthonormal basis of $\mathcal{V}$.

Theorem 1 Among the eigenvectors of a Hermitian operator H, one can always find a set of N vectors $|\phi_1\rangle, |\phi_2\rangle, \ldots, |\phi_N\rangle$ that make up an orthonormal basis of $\mathcal{V}$.

This theorem leads to an important representation of Hermitian operators called the *spectral decomposition.*

Theorem 2 Let H be a Hermitian operator, $\{|\phi_i\rangle\}$ a basis of orthonormal eigenvectors of H and $\{\lambda_i\}$ the corresponding eigenvalues. Then

$$H = \sum_{i=1}^{N} \lambda_i |\phi_i\rangle\langle\phi_i| \,. \tag{2.38}$$

Proof To prove that two linear operators are equal, it is enough to show that they act identically on all vectors in some basis. But

$$H|\phi_j\rangle = \lambda_j|\phi_j\rangle \,;$$

$$\left[\sum_{i=1}^{N} \lambda_i|\phi_i\rangle\langle\phi_i|\right] |\phi_j\rangle = \sum_{i=1}^{N} \lambda_i \left(\langle\phi_i|\phi_j\rangle\right)|\phi_i\rangle = \sum_{i=1}^{N} \lambda_i\delta_{ij}|\phi_i\rangle = \lambda_j|\phi_j\rangle \,. \qquad \clubsuit$$

There exists a systematic procedure to find a basis of orthonormal eigenvectors of a Hermitian operator H. Let $|\phi\rangle$ be an eigenvector of H, λ the corresponding eigenvalue and $\{|u_i\rangle\}$ an arbitrary orthonormal basis. We have

$$H|\phi\rangle = \lambda|\phi\rangle \,,$$

and we can write

$$|\phi\rangle = \sum_{j=1}^{N} c_j|u_j\rangle \,.$$

Projecting the eigenvalue equation on the vector $\langle u_i|$ we get

$$\langle u_i|H\left[\sum_{j=1}^{N} c_j|u_j\rangle\right] = \langle u_i|\lambda\left[\sum_{j=1}^{N} c_j|u_j\rangle\right] \,,$$

from which we find

$$\sum_{j=1}^{N} c_j \langle u_i | H | u_j \rangle = \lambda \sum_{j=1}^{N} c_j \langle u_i | u_j \rangle .$$

Therefore

$$\sum_{j=1}^{N} (H_{ij} - \lambda \delta_{ij}) c_j = 0 , \tag{2.39}$$

where, of course, $H_{ij} = \langle u_i | H | u_j \rangle$. Equation (2.39) can be interpreted as a system of N linear homogeneous algebraic equations for the N unknowns c_j. Such a system has a nontrivial solution if and only if the determinant of the matrix of coefficients vanishes, that is,

$$\det \left[H_{ij} - \lambda \delta_{ij} \right] = 0 . \tag{2.40}$$

Equation (2.40), called the *secular equation*, is an N-th degree equation for λ. Its solutions are the N (possibly repeated) eigenvalues of H.

The search for eigenvalues and orthornormal eigenvectors of H can therefore be effected in the following way:[5]

i) Write down H in a given basis, write the secular equation and solve it. One thus obtains the N (sometimes repeated) eigenvalues λ_k of H.
ii) For each eigenvalue λ_k, write (2.39) for coefficients c_j. In the case where λ_k is a simple root of the secular equation (λ_k is then *nondegenerate*), (2.39) specifies all c_j within a multiplicative factor. The normalization of $|\phi\rangle$ brings this factor to a complex number of unit norm.
iii) In the case where λ_k is a root of order n of the secular equation (we say that λ_k is *n-fold degenerate*), (2.39) specifies $(N - n)$ coefficients c_j in terms of n others. This is in fact the specification of the subspace $\mathcal{V}_{\lambda_k}$ of $\mathcal{V}$ made up of eigenvectors with eigenvalue λ_k. That subspace has dimension equal to n. It is easy to find n sets of N coefficients c_j that specify n orthonormal eigenvectors.
iv) Once the procedure is applied to all distinct eigenvalues, one obtains N orthonormal eigenvectors of H which can be written as

$$|\phi_k\rangle = \sum_{j=1}^{N} c_j^{(k)} |u_j\rangle , \qquad k = 1, 2, \ldots, N . \tag{2.41}$$

---

[5] Symbolic softwares evaluate eigenvalues and eigenvectors of a matrix. For example Maple V (release 5) uses the commands `eigenvals` and `eigenvects` of the package `linalg`, whereas Mathematica (release 3) uses the commands `Eigenvalues` and `Eigenvectors`. These softwares perform most of the matrix operations introduced in this chapter.

Since the two bases $\{|u_j\rangle\}$ and $\{|\phi_k\rangle\}$ are orthonormal, their vectors are related by a unitary transformation. Thus one can write

$$|\phi_j\rangle = U|u_j\rangle\,, \qquad j = 1, 2, \ldots, N\,. \tag{2.42}$$

Projecting this equation on the vector $\langle u_i|$, one easily obtains the relation $c_i^{(j)} = U_{ij}$. The matrix of U (in the $|u_i\rangle$ basis) is therefore obtained by putting the columns $[c^{(j)}]$ side by side.

The representation of H in the $|\phi_k\rangle$ basis is diagonal. In fact

$$\lambda_i \delta_{ij} = \langle \phi_i|H|\phi_j\rangle = \langle u_i|U^\dagger H U|u_j\rangle = (U^\dagger H U)_{ij}\,. \tag{2.43}$$

Thus the product $[U^\dagger][H][U]$ is a diagonal matrix with eigenvalues of H on the diagonal. This is why the computation of eigenvalues and eigenvectors of a matrix $[H]$ is called the *diagonalization* of $[H]$.

One can show that if $[H]$ is a Hermitian matrix, the matrix $[U]$ which diagonalizes it is unitary. If, in addition, $[H]$ is real, and therefore symmetric, the matrix which diagonalizes it is real and *orthogonal* (i.e. its inverse is equal to its transpose). In another connection, if $[H]$ is *block-diagonal* (i.e. made up of square blocks along the main diagonal), its eigenvalues and eigenvectors are found by diagonalizing each block independently.

It is interesting to see that an active transformation of vectors in $\mathcal{V}$ and operators acting in $\mathcal{V}$ does not change the spectrum of an operator. Indeed let $\bar{U}$ be a unitary operator and let $H' = \bar{U} H \bar{U}^\dagger$. One easily shows that if $|\phi_i\rangle$ is an eigenvector of H corresponding to the eigenvalue λ_i, $\bar{U}|\phi_i\rangle$ is an eigenvector of H' corresponding to the same eigenvalue. Moreover, if H is Hermitian, so is H'.

We saw that eigenvalues of a projector are equal to 0 or 1. The spectral decomposition of a projector is therefore given by

$$P = \sum_{\{i\}} |\phi_i\rangle\langle\phi_i|\,, \tag{2.44}$$

where the sum is restricted to values of i that correspond to the eigenvalue 1. Thus for any vector $|\psi\rangle$

$$P|\psi\rangle = \sum_{\{i\}} \left(\langle\phi_i|\psi\rangle\right)|\phi_i\rangle\,. \tag{2.45}$$

The word "projector" comes from the fact that P projects $|\psi\rangle$ on the subspace generated by the $|\phi_i\rangle$.

2.6 Functions of an Operator

Let $f(z)$ be a function that can be expanded in a Taylor series about the origin, that is,

$$f(z) = \sum_{n=0}^{\infty} c_n z^n \,. \tag{2.46}$$

Let A be an operator[6] acting in a vector space $\mathcal{V}$, and consider the following infinite series:

$$\sum_{n=0}^{\infty} c_n A^n \,.$$

In the case where this series converges to an operator acting in $\mathcal{V}$, that operator is called $f(A)$.

It is easy to show that the adjoint of $f(A)$ coincides with $f(A^\dagger)$. If A is Hermitian, so is $f(A)$. Furthermore for any invertible operator B we have (Exercise 2.16)

$$f\left(BAB^{-1}\right) = Bf(A)B^{-1} \,. \tag{2.47}$$

In particular, for a unitary operator U

$$f\left(UAU^\dagger\right) = Uf(A)U^\dagger \,. \tag{2.48}$$

Let us denote by $f'(z)$ the derivative of $f(z)$ with respect to its argument. One can show that

$$\frac{\mathrm{d}}{\mathrm{d}t} f(tA) = Af'(tA) \,. \tag{2.49}$$

The spectral decomposition of a function of an operator immediately follows from the spectral decomposition of its argument. Suppose, for instance, that H is Hermitian and that its spectral decomposition is given as in (2.38). One then checks that $|\phi_i\rangle$ is an eigenvector of $f(H)$ and that

$$f(H) = \sum_{i=1}^{N} f(\lambda_i) |\phi_i\rangle\langle\phi_i| \,. \tag{2.50}$$

The most important function of an operator is the *exponential* defined as

$$\exp(A) \equiv \mathrm{e}^A \equiv \sum_{n=0}^{\infty} \frac{1}{n!} A^n \,. \tag{2.51}$$

One can show that this series converges for all A and that it has the following properties:

$$\exp(A) = \lim_{n\to\infty} \left\{ I + \frac{1}{n} A \right\}^n , \tag{2.52}$$

$$\det\{\exp(A)\} = \exp\{\mathrm{Tr}(A)\} \,. \tag{2.53}$$

---

[6] Unless otherwise indicated, we shall henceforth assume that operators are linear.

The *commutator* $[A, B]$ of two operators A et B is defined as

$$[A, B] \equiv AB - BA . \tag{2.54}$$

The commutator has the following properties (Exercise 2.17):

$$[A + B, C] = [A, C] + [B, C] , \tag{2.55}$$

$$[AB, C] = A[B, C] + [A, C]B , \tag{2.56}$$

$$[AB, CD] = A[B, C]D + AC[B, D] + [A, C]DB + C[A, D]B . \tag{2.57}$$

We say that two operators A and B *commute* if $AB = BA$. It is clear that if A and B commute, $f(A)$ and $g(B)$ commute for any functions f and g. Moreover for every A and B, one can show (Exercise 2.18) that

$$\exp(A)B\exp(-A) = B+[A, B]+\frac{1}{2!}[A, [A, B]]+\frac{1}{3!}[A, [A, [A, B]]]+\cdots \tag{2.58}$$

If $[A, B]$ commutes with A and with B, we get from (2.47) and (2.58) that

$$\exp(A)\exp(B)\exp(-A)\exp(-B) = \exp\{[A, B]\} . \tag{2.59}$$

2.7 Commuting Operators

In quantum mechanics the notion of commuting operators is of paramount importance. We state two important properties of such operators.

Lemma Let $|b_1\rangle$ and $|b_2\rangle$ be two eigenvectors of a Hermitian operator B with eigenvalues b_1 and b_2. Let A be an operator that commutes with B. If $b_1 \neq b_2$, then $\langle b_1|A|b_2\rangle = 0$.

Proof

$$0 = \langle b_1|AB-BA|b_2\rangle = \langle b_1|Ab_2|b_2\rangle-\langle b_1|b_1A|b_2\rangle = (b_2-b_1)\langle b_1|A|b_2\rangle. \quad ♣$$

Theorem Let H and K be two Hermitian operators. There exists a basis $|\phi_1\rangle$, $|\phi_2\rangle$, $\ldots$, $|\phi_N\rangle$ of common eigenvectors of H and K, that is, such that $H|\phi_i\rangle = \lambda_i|\phi_i\rangle$ and $K|\phi_i\rangle = \eta_i|\phi_i\rangle$, if and only if H and K commute.

In other words the matrices of H and K can be diagonalized by the same operator U if and only if H and K commute. This theorem can be generalized to an arbitrary number of Hermitian operators: several Hermitian operators can be diagonalized by the same unitary operator if and only if they pairwise commute.

Let K^1, K^2, ..., K^p be a set of Hermitian operators that pairwise commute. Let $|\phi_1\rangle$, $|\phi_2\rangle$, ..., $|\phi_N\rangle$ be a basis of common eigenvectors of K^1, K^2, ..., K^p such that

$$K^a|\phi_i\rangle = k_i^a|\phi_i\rangle, \qquad a = 1, 2, \ldots, p; \quad i = 1, 2, \ldots, N. \tag{2.60}$$

We say that K^1, K^2, ..., K^p make up a *complete set of commuting operators* if all multiplets $(k_i^1, k_i^2, \ldots, k_i^p)$ are distinct. Specification of an eigenvalue for each operator then uniquely determines (within a multiplicative factor) the common eigenvector. The complete set is *minimal* if the removal of any operator destroys its completeness.

Example In the space $\mathcal{C}^5$, two operators represented by the following matrices make up a minimal complete set of commuting operators.

$$[H] = \begin{pmatrix} -2 & 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}, \qquad [K] = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}. \qquad \clubsuit$$

2.8 Direct Sum of Vector Spaces

Let $\mathcal{V}^1$ and $\mathcal{V}^2$ be two vector spaces of dimensions N_1 and N_2. Suppose that scalar products have been defined in $\mathcal{V}^1$ and $\mathcal{V}^2$. Let $\{|u_i(1)\rangle,\ i = 1,\ 2,\ \ldots,\ N_1\}$ be an orthonormal basis of $\mathcal{V}_1$ and let $\{|v_k(2)\rangle,\ k = 1,\ 2,\ \ldots,\ N_2\}$ be an orthonormal basis of $\mathcal{V}_2$. The *direct sum* of $\mathcal{V}^1$ and $\mathcal{V}^2$, denoted by $\mathcal{V}^1 \oplus \mathcal{V}^2$, is the vector space of dimension $N_1 + N_2$ defined by stipulating that $\{|u_i(1)\rangle, |v_k(2)\rangle\}$ makes up an orthonormal basis of $\mathcal{V}^1 \oplus \mathcal{V}^2$. Note that this at once defines a scalar product in $\mathcal{V}^1 \oplus \mathcal{V}^2$. To a vector $|\psi(1)\rangle$ in $\mathcal{V}^1$ given by

$$|\psi(1)\rangle = \sum_{i=1}^{N_1} c_i|u_i(1)\rangle\,,$$

there corresponds one and only one vector $|\psi\rangle$ in $\mathcal{V}^1 \oplus \mathcal{V}^2$, represented by an identical linear combination. So is it with a vector $|\phi(2)\rangle$ in $\mathcal{V}_2$. It is important to point out that such a correspondence does not depend on the choice of orthonormal bases of $\mathcal{V}^1$ and $\mathcal{V}^2$.

On the other hand, there are many vectors in $\mathcal{V}^1 \oplus \mathcal{V}^2$ that correspond neither to a vector in $\mathcal{V}^1$ nor to one in $\mathcal{V}^2$. These are linear combinations that involve at the same time vectors $|u_i(1)\rangle$ and vectors $|v_k(2)\rangle$.

Let $A(1)$ and $A(2)$ be linear operators defined in $\mathcal{V}^1$ and $\mathcal{V}^2$. The direct sum of $A(1)$ and $A(2)$, denoted by $A(1) \oplus A(2)$, is the linear operator defined so that

$$\{A(1) \oplus A(2)\} \left[\sum_{i=1}^{N_1} c_i |u_i(1)\rangle + \sum_{k=1}^{N_2} d_k |v_k(2)\rangle \right]$$
$$\equiv \sum_{i=1}^{N_1} c_i \, A(1)|u_i(1)\rangle + \sum_{k=1}^{N_2} d_k \, A(2)|v_k(2)\rangle \, . \tag{2.61}$$

Once again the direct sum of operators does not depend on the choice of orthonormal bases of $\mathcal{V}^1$ and $\mathcal{V}^2$.

In Chap. 12, we shall define another kind of vector space made up from $\mathcal{V}^1$ and $\mathcal{V}^2$.

Exercises

2.1. Let $|u_1\rangle$, $|u_2\rangle$ and $|u_3\rangle$ be an orthonormal basis of a three-dimensional complex space and let

$$|\phi_1\rangle = |u_1\rangle - 2|u_2\rangle \, ,$$
$$|\phi_2\rangle = |u_1\rangle + \mathrm{i}|u_3\rangle \, ,$$
$$|\phi_3\rangle = |u_3\rangle \, .$$

a) Evaluate the scalar products $\langle\phi_1|\phi_2\rangle$, $\langle\phi_2|\phi_3\rangle$ and $\langle\phi_3|\phi_1\rangle$.
b) Find three constants c_1, c_2 and c_3 such that vectors $c_1|\phi_1\rangle$, $c_2|\phi_2\rangle$ and $c_3|\phi_3\rangle$ are normalized.

2.2. Let $|u_1\rangle$, $|u_2\rangle$ and $|u_3\rangle$ be an orthonormal basis and let

$$|\phi_1\rangle = \alpha \left\{\mathrm{i}|u_1\rangle + |u_2\rangle - |u_3\rangle\right\} ,$$
$$|\phi_2\rangle = \beta \left\{|u_2\rangle + |u_3\rangle\right\} .$$

a) Show that $\langle\phi_1|\phi_2\rangle = 0$ for all α and β.
b) Find α and β so that $\langle\phi_1|\phi_1\rangle = 1 = \langle\phi_2|\phi_2\rangle$.
c) Find a vector $|\phi_3\rangle$ which, together with $|\phi_1\rangle$ and $|\phi_2\rangle$, makes up an orthonormal basis.

2.3. Let $|u_1\rangle$, $|u_2\rangle$ and $|u_3\rangle$ be an orthonormal basis and let

$$|\psi\rangle = 2\mathrm{i}|u_1\rangle - |u_2\rangle + 4|u_3\rangle \, ,$$
$$|\phi\rangle = |u_1\rangle + 3\mathrm{i}|u_2\rangle - |u_3\rangle \, .$$

Evaluate $\langle\phi|\phi\rangle$, $\langle\psi|\psi\rangle$ and $\langle\phi|\psi\rangle$ and check that the Schwartz inequality (2.6) is satisfied.

2.4. Show that the Schwartz inequality (2.6) becomes an equality if and only if $|\phi\rangle$ and $|\psi\rangle$ are proportional.

2.5. Let $|u_1\rangle$, $|u_2\rangle$ and $|u_3\rangle$ be an orthonormal basis. Define two operators O and P so that for all complex numbers α_1, α_2 and α_3,

$$\begin{aligned} &O\left\{\alpha_1|u_1\rangle + \alpha_2|u_2\rangle + \alpha_3|u_3\rangle\right\} = \alpha_1^2|u_1\rangle + \alpha_2^2|u_2\rangle + \alpha_3^2|u_3\rangle\,, \\ &P\left\{\alpha_1|u_1\rangle + \alpha_2|u_2\rangle + \alpha_3|u_3\rangle\right\} = \alpha_1|u_1\rangle\,. \end{aligned}$$

Are these operators linear? Do they have inverses? Justify your answers.

2.6. Let O, O_1 and O_2 be operators and $O^\dagger$, $O_1^\dagger$ and $O_2^\dagger$ their adjoints.
a) Show that $O^\dagger$ is linear.
b) Show that $(O^\dagger)^\dagger = O$.
c) Show that $(O_1O_2)^\dagger = O_2^\dagger O_1^\dagger$.

2.7. Let $|u_1\rangle$, $|u_2\rangle$ and $|u_3\rangle$ be an orthonormal basis and let

$$\begin{aligned} &|\phi\rangle = 2|u_1\rangle + \mathrm{i}|u_3\rangle\,, \\ &|\psi\rangle = |u_2\rangle - |u_3\rangle\,. \end{aligned}$$

Let O be a linear operator whose matrix representation in the $|u_i\rangle$ basis is given by

$$\begin{pmatrix} 4 & 2\mathrm{i} & 3 \\ -2\mathrm{i} & 0 & 1+\mathrm{i} \\ 3 & 1-\mathrm{i} & -2 \end{pmatrix}.$$

Is O Hermitian? Illustrate your answer by evaluating $\langle\phi|O|\psi\rangle$ and $\langle\psi|O|\phi\rangle$.

2.8. Show that for every $|\phi\rangle$, the operator $|\phi\rangle\langle\phi|$ is Hermitian.

2.9. Show that an operator U whose matrix elements in some orthonormal basis are given by

$$[U] = \begin{pmatrix} -\mathrm{i}/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/2 & -\mathrm{i}/2 & 1/\sqrt{2} \\ -1/2 & \mathrm{i}/2 & 1/\sqrt{2} \end{pmatrix}$$

is unitary.

2.10.
a) Show that matrix elements of the operator

$$L = \sum_{i,j=1}^{N} L_{ij}|u_i\rangle\langle u_j|,$$

in the orthonormal basis $\{|u_i\rangle\}$, coincide with the L_{ij}.
b) Let $\mathcal{L}_{kl}$ be defined so that

$$L^2 = \sum_{k,l=1}^{N} \mathcal{L}_{kl}|u_k\rangle\langle u_l|\,.$$

Write $\mathcal{L}_{kl}$ in terms of the L_{ij}.

2.11. Find the eigenvalues and orthonormal eigenvectors of the following matrix:

$$[H] = \begin{pmatrix} 2 & -\sqrt{2} & \mathrm{i}\sqrt{2} \\ -\sqrt{2} & 3 & \mathrm{i} \\ -\mathrm{i}\sqrt{2} & -\mathrm{i} & 3 \end{pmatrix}.$$

2.12. An operator H has the following matrix representation in the $|u_i\rangle$ basis:

$$[H] = \begin{pmatrix} 1 & 0 & -\mathrm{i} \\ 0 & 1 & 0 \\ \mathrm{i} & 0 & 1 \end{pmatrix}.$$

a) Find the eigenvalues of H and three orthonormal eigenvectors $|\phi_1\rangle$, $|\phi_2\rangle$ and $|\phi_3\rangle$.
b) Let U be the operator defined so that

$$|\phi_1\rangle = U|u_1\rangle, \qquad |\phi_2\rangle = U|u_2\rangle, \qquad |\phi_3\rangle = U|u_3\rangle.$$

Evaluate matrix elements of U in the $|u_i\rangle$ basis and check that the matrix $[U]$ is unitary.

2.13. The matrix representation of an operator H is given by

$$[H] = \begin{pmatrix} 0 & \mathrm{i}/\sqrt{2} & -\mathrm{i}/\sqrt{2} \\ -\mathrm{i}/\sqrt{2} & 1 & 1 \\ \mathrm{i}/\sqrt{2} & 1 & 1 \end{pmatrix}.$$

a) Find the eigenvalues λ_i and orthonormal eigenvectors $|\phi_i\rangle$ of H.
b) By means of the correspondence ket ↔ column matrix and bra ↔ row matrix, check that

$$I = \sum_i |\phi_i\rangle\langle\phi_i|, \qquad H = \sum_i \lambda_i|\phi_i\rangle\langle\phi_i|.$$

2.14. Find the eigenvalues and eigenvectors of the following matrix, where E is real and a is complex:

$$[H] = \begin{pmatrix} E & a \\ a^* & E \end{pmatrix}.$$

2.15. Find the eigenvalues and eigenvectors of the following matrix, where E and a are real:

$$[H] = \begin{pmatrix} E & a & 0 \\ a & E & a \\ 0 & a & E \end{pmatrix}.$$

2.16. Making use of the series expansion of a function f, show that, for any operator A and any invertible operator B

$$f\left(BAB^{-1}\right) = Bf(A)B^{-1}\,.$$

2.17. Prove (2.55)–(2.57).

2.18. Let A et B be two operators and let

$$F(t) = \exp(tA)B\exp(-tA)\,.$$

a) By induction show that

$$\frac{\mathrm{d}^n F}{\mathrm{d}t^n} = \left[A, \frac{\mathrm{d}^{n-1}F}{\mathrm{d}t^{n-1}}\right]\,.$$

b) Use this result to derive (2.58).

2.19. Is the lemma on p. 25 true if B is an arbitrary operator? if B is a unitary operator?

2.20. Find an orthonormal basis of column vectors that are common eigenvectors of the two following matrices:

$$[H] = \begin{pmatrix} 0 & \mathrm{i} & 0 & 0 \\ -\mathrm{i} & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & -3 \end{pmatrix}, \qquad [K] = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Do these matrices make up a complete set of commuting operators? Justify your answer.

2.21. Two matrices are given by

$$[H] = \begin{pmatrix} a & 0 & 0 \\ 0 & -a & 0 \\ 0 & 0 & -a \end{pmatrix}, \qquad [K] = \begin{pmatrix} b & 0 & 0 \\ 0 & 0 & b \\ 0 & b & 0 \end{pmatrix},$$

where a and b are real numbers.

a) Find a basis of common eigenvectors of $[H]$ and $[K]$.

b) Among the following four sets, which one is a complete set of commuting operators?

$$\{[H]\}, \qquad \{[K]\}, \qquad \{[H],[K]\}, \qquad \{[H]^2,[K]\}\,.$$

3 Finite-Dimensional State Spaces

A mathematical formalism, like the one introduced in Chap. 2, is not by itself a physical theory. The latter includes, in addition, interpretation rules that associate, more or less directly, empirical concepts or procedures to objects of the formalism. The purpose of this chapter is twofold: to state the fundamental interpretation rules of quantum mechanics, and to write down the equation that governs the time evolution of a quantum system. From a logical point of view, interpretation rules and the time-evolution equation are postulates. At first sight, it is not easy to understand their relevance. They are justified through the experimental confirmation of predictions of quantum mechanics, many examples of which will be given in forthcoming chapters.

3.1 The State of a System

A *quantum system* is a physical system correctly described by quantum mechanics. Atoms, molecules, atomic nuclei and elementary particles are quantum systems. But these are not limited to microscopic objects. Indeed crystals, superfluids and superconductors, to give but a few examples, are macroscopic quantum systems.

In general a quantum system interacts, one way or another, with its environment. Nevertheless, it is very useful to introduce the notion of isolation. A quantum system is *isolated* if it interacts with the environment through classical fields only. This notion will eventually get clearer. A fortiori, a system which interacts in no way with its environment is isolated. In the forthcoming discussion we will assume, unless otherwise indicated, that quantum systems are isolated. Nonisolated systems will be examined specifically in Chaps. 12 and 21.

Is it possible that all physical systems, isolated or not, be quantum systems? This question turns out to be extremely complex. Eventually, we shall have the opportunity to apprehend it a little better. At any rate, the formulation of interpretation rules that we shall give in this section and the next presupposes the existence of *measurement apparatus*. Apparatus are macroscopic objects obeying the laws of classical mechanics. We will take them as such and will not attempt, at this stage, to describe their behavior in terms of quantum mechanics.

L. Marchildon, *Quantum Mechanics*

With this in mind, let us now state the first interpretation rule of quantum mechanics.

Interpretation Rule 1 To every quantum system is associated a complex vector space called the *state space.*

Depending on the system, the state space dimension can be finite or infinite. In this chapter, we will examine quantum systems with finite-dimensional state spaces only. Nevertheless much of the formalism developed will apply, mutatis mutandis, to systems with infinite-dimensional state spaces.

State spaces in quantum mechanics have a richer structure than the one of more general vector spaces. The most important additional characteristic is that in a state space, a scalar product is necessarily defined. Infinite-dimensional state spaces satisfy further restrictions which make them akin to what mathematicians call *Hilbert spaces.* We will investigate these spaces in more detail in Chap. 5.

To a quantum system therefore, there corresponds a complex vector space supplied with a scalar product. This correspondence, however, is not one-to-one. Indeed the same complex vector space (or in any case *isomorphic*, that is, mathematically identical spaces), with the same scalar product, can correspond to two different systems. In the next sections, we will see what tells one from the other two quantum systems with identical state spaces.

The second quantum-mechanical interpretation rule has to do with vectors in the state space. Not everyone agrees on how to formulate it. Here we shall give the simplest formulation, and postpone discussion of other options to Chap. 6. Fortunately, the application of quantum mechanics to concrete situations does not depend on subtle distinctions in the interpretation of state space vectors.

Interpretation Rule 2 At any given time, the state of a quantum system is described by a nonzero vector in the state space. This vector, called the *state vector*, is a function of time. Two vectors that differ only by a scalar multiple represent the same quantum state. Otherwise, they represent physically different states.

The state vector is denoted by the symbol $|\psi(t)\rangle$. In a finite-dimensional vector space, a nonzero vector can always be written as a multiple of a normalized vector. This means that the state of a quantum system, at a given time, can always be represented by a normalized vector.

In the next sections we will see how, from the knowledge of the state vector, one can make predictions about results of physical measurements. In general, these predictions are probabilistic. This means that quantum mechanics does not predict the result of a physical measurement, but rather the probability to obtain such and such result. Nevertheless, the prevalent interpretation of quantum mechanics asserts that the state vector incorporates everything that one can possibly know about a quantum system.

3.2 Dynamical Variables

By a *dynamical variable* we mean any property of a physical system which is susceptible of being (directly or indirectly) measured. The most common dynamical variables are energy, momentum, position, angular momentum, spin, mass, electric charge, etc.

Interpretation Rule 3 With every dynamical variable of a quantum system is associated a Hermitian operator acting in the state space.

The Hermitian operator associated with a dynamical variable is called an *observable*. Different operators are associated with different dynamical variables. One cannot, however, associate a dynamical variable with every Hermitian operator.[1]

How, in each case, do we determine the Hermitian operator that corresponds to a dynamical variable? How, for instance, do we find the operator that corresponds to the momentum of an electron? From a logical point of view, these associations are all postulates. Fortunately, work with a few examples soon develops the intuition necessary to set these correspondences correctly.

A dynamical variable is a property susceptible of being measured. Suppose the Hermitian operator corresponding to a dynamical variable has been determined. The fourth quantum-mechanical interpretation rule specifies how this operator is connected with the variable's measurement results.

Interpretation Rule 4 The only possible results of the measurement of a dynamical variable are the eigenvalues of the corresponding Hermitian operator.

The measurement of a dynamical variable requires interaction of the quantum system with an apparatus. During measurement, the quantum system is not isolated. This process will be analyzed in detail in Chap. 21.

To make the connection with notations introduced in Chap. 2, let us denote by A the Hermitian operator associated with some dynamical variable. It is useful to denote the variable itself by the same symbol. Let $\{|a_i\rangle\}$ be a basis of orthonormal eigenvectors of A and let $\{a_i\}$ be the corresponding eigenvalues. From the spectral decomposition theorem (2.38),

$$A = \sum_{i=1}^{N} a_i |a_i\rangle\langle a_i| \, .$$

The only possible results of the measurement of A are the numbers a_i. They are always real. Note that dimensions of an operator's matrix elements coin-

---

[1] There are at least three reasons for that, explained in Sects. 5.9.2 and 7.7 as well as in [237]. We should point out that the physical meaning of dynamical variables associated with some Hermitian operators can be highly opaque.

cide with dimensions of its eigenvalues. Bras and kets we are presently examining are dimensionless, as suggested by normalization conditions.

The set of possible measurement results of a dynamical variable is completely determined by the Hermitian operator associated with it. To predict (either categorically or probabilistically) the result of a particular measurement of a dynamical variable, however, it is natural to think that knowledge of the associated operator is not enough. One must, in addition, know the state of the quantum system. The fifth quantum-mechanical interpretation rule pertains to this type of prediction. Before giving a general formulation, we will examine the case where all eigenvalues of A are different. Let $|\psi(t)\rangle$ be the normalized state vector of the system. The probability that a measurement of A at time t yields the value a_i, denoted by $P(A \to a_i; t)$, is then given by

$$P(A \to a_i; t) = |\langle a_i|\psi(t)\rangle|^2 . \tag{3.1}$$

As it must be, $|\langle a_i|\psi(t)\rangle|^2 \geq 0$, and the sum of the probabilities of all possible results is equal to 1. Indeed

$$\begin{aligned}\sum_{i=1}^{N} |\langle a_i|\psi(t)\rangle|^2 &= \sum_{i=1}^{N} \langle a_i|\psi(t)\rangle\langle a_i|\psi(t)\rangle^* = \sum_{i=1}^{N} \langle\psi(t)|a_i\rangle\langle a_i|\psi(t)\rangle \\ &= \langle\psi(t)| \left[\sum_{i=1}^{N} |a_i\rangle\langle a_i|\right] |\psi(t)\rangle = \langle\psi(t)|I|\psi(t)\rangle = 1 .\end{aligned}$$

We have made use of (2.27), true when the vectors $|a_i\rangle$ make up an orthonormal basis, and of the normalization condition on $|\psi(t)\rangle$.

We note that if $|\psi(t)\rangle$ is not normalized, $\langle\psi|\psi\rangle^{-1/2}|\psi\rangle$ is. The probability of obtaining result a_i can then be written as

$$P(A \to a_i; t) = \left|\langle a_i|\frac{1}{\langle\psi|\psi\rangle^{1/2}}|\psi(t)\rangle\right|^2 = \frac{|\langle a_i|\psi\rangle|^2}{\langle\psi|\psi\rangle} . \tag{3.2}$$

It is important to understand the operational meaning of the rule just stated. The rule specifies how to compute the probability of getting such and such result upon measurement of a dynamical variable. How are we going to compare this prediction with experiment? It is clear that no single measurement will provide an experimental determination of a probability. To empirically determine the probability that the tossup of a die yields a 3, for instance, the die must be tossed a large number of times. So is it with quantum systems. To empirically determine the probability that a measurement of a dynamical variable A pertaining to a quantum system in state $|\psi\rangle$ yields value a_i, one must perform the measurement a large number of times. One must also make sure that before each measurement the state of the system is actually described by the vector $|\psi\rangle$.

In Chap. 6, we will see that we can make sure the quantum system is in state $|\psi\rangle$ by a process of *state preparation.* In practice, the measurements mentioned are most often made not on the same, but on identical systems (one thousand hydrogen atoms, for instance). Summarizing, probabilities are evaluated by measuring a dynamical variable on a large number of identical systems which have been prepared similarly.

Quantum mechanics makes well-defined predictions on probabilities of measurement results. From these one can extract predictions on statistical moments like the mean value or the standard deviation of results. Assume that a large number of quantum systems have been prepared similarly and are all described by the normalized state vector $|\psi(t)\rangle$. At time t, a measurement of A is made on each system. One can compute the mean value of these experimental results. Let us denote it by $\langle A\rangle_\psi$. What numerical value does quantum mechanics predict for $\langle A\rangle_\psi$? From the definition of the mean, $\langle A\rangle_\psi$ is the sum of possible results multiplied by the probability of each result, that is,

$$\begin{aligned}\langle A\rangle_\psi &= \sum_{i=1}^{N} a_i\,|\langle a_i|\psi\rangle|^2 = \sum_{i=1}^{N} a_i\langle\psi|a_i\rangle\langle a_i|\psi\rangle \\ &= \langle\psi|\left[\sum_{i=1}^{N} a_i|a_i\rangle\langle a_i|\right]|\psi\rangle = \langle\psi|A|\psi\rangle\,. \qquad (3.3)\end{aligned}$$

Likewise one easily shows that the mean value of the squares of possible results is equal to

$$\sum_{i=1}^{N} a_i^2\,|\langle a_i|\psi\rangle|^2 = \langle\psi|A^2|\psi\rangle\,.$$

The standard deviation of measurement results of A is therefore given by

$$(\Delta A)_\psi = \{\mathrm{Var}(A)\}^{1/2} = \left\{\langle\psi|A^2|\psi\rangle - \langle\psi|A|\psi\rangle^2\right\}^{1/2}, \qquad (3.4)$$

where the *variance* is the difference between the mean value of squares and the square of the mean value.

In general, eigenvalues of a Hermitian operator are not all different. Formulated as follows, the fifth interpretation rule holds for any degree of degeneracy of the eigenvalue a_i.

Interpretation Rule 5 Let $|\psi(t)\rangle$ be the normalized state vector of a quantum system and let A be a dynamical variable pertaining to the system. The probability that the result of a measurement of A at time t is equal to a_i is given by

$$P(A \to a_i; t) = \sum_{\{i\}} |\langle a_i|\psi(t)\rangle|^2\,, \qquad (3.5)$$

where the sum runs on all values of the index i that correspond to the eigenvalue a_i.

Equations (3.3) and (3.4) remain true when there is degeneracy.

Example Let A be a dynamical variable associated with a Hermitian operator whose matrix elements in the orthonormal basis $\{|u_i\rangle\}$ are given by

$$[A] = \begin{pmatrix} 2 & -3\sqrt{2} & 3\sqrt{2} \\ -3\sqrt{2} & -1 & -3 \\ 3\sqrt{2} & -3 & -1 \end{pmatrix} .$$

Let

$$|\psi\rangle = \frac{1}{\sqrt{5}} \{|u_1\rangle + 2\mathrm{i}|u_3\rangle\}$$

be the state vector of the system at time t. What are the possible results of a measurement of A, what are the respective probabilities, what is the mean value of A and what is its standard deviation?

Eigenvalues and eigenvectors of A are given by (Exercise 3.7)

$$a_1 = 8 , \qquad |a_1\rangle = \frac{-1}{\sqrt{2}}|u_1\rangle + \frac{1}{2}|u_2\rangle - \frac{1}{2}|u_3\rangle ;$$

$$a_2 = -4 , \qquad |a_2\rangle = \frac{1}{\sqrt{2}}|u_2\rangle + \frac{1}{\sqrt{2}}|u_3\rangle ;$$

$$a_3 = -4 , \qquad |a_3\rangle = \frac{1}{\sqrt{2}}|u_1\rangle + \frac{1}{2}|u_2\rangle - \frac{1}{2}|u_3\rangle .$$

For measurement result probabilities we have

$$P(A \rightarrow 8; t) = |\langle a_1|\psi\rangle|^2 = 0.3 ;$$
$$P(A \rightarrow -4; t) = |\langle a_2|\psi\rangle|^2 + |\langle a_3|\psi\rangle|^2 = 0.7 .$$

The mean value of A is given by

$$\langle A\rangle_\psi = \langle\psi|A|\psi\rangle = -0.4 .$$

For the standard deviation of A, we finally obtain

$$(\Delta A)_\psi = \left\{\langle\psi|A^2|\psi\rangle - \langle\psi|A|\psi\rangle^2\right\}^{1/2} = 0.4\sqrt{189}. \qquad ♣$$

Equation (3.5) for $P(A \rightarrow a_i; t)$ can also be written in terms of P_{a_i}, the projector on the subspace of eigenvectors associated with eigenvalue a_i. Thus

$$P(A \rightarrow a_i; t) = \langle\psi(t)| \left[\sum_{\{i\}} |a_i\rangle\langle a_i|\right] |\psi(t)\rangle = \langle\psi(t)|P_{a_i}|\psi(t)\rangle. \tag{3.6}$$

Equation (3.6) means that the probability that a measurement of A yields the value a_i is equal to the mean value of P_{a_i}.[2]

By a *property* of a quantum system, we mean the specification that the value of a dynamical variable A is equal to a number a_i; or, more generally, the specification that the value of A belongs to a set of values $\mathcal{E}$. It is clear that a property uniquely corresponds to a projector, defined by the subspace associated with eigenvalues of A belonging to $\mathcal{E}$.

We know that a projector is a Hermitian operator. Is there a corresponding dynamical variable? To answer this question it is useful to consider an operator that simply projects on a one-dimensional subspace. Analysis of more general projectors is carried out similarly. Let $|\phi\rangle$ be a normalized vector and let $P_\phi = |\phi\rangle\langle\phi|$. Eigenvalues of P_ϕ are 1 and 0. Its eigenvectors are $|\phi\rangle$, on the one hand, and any vector orthogonal to $|\phi\rangle$, on the other hand. Thus one can say that the dynamical variable associated with P_ϕ is "being in the state represented by $|\phi\rangle$." For a quantum system in a state represented by the normalized vector $|\psi\rangle$, measurement of that property indeed yields value 1 if $|\psi\rangle = |\phi\rangle$ and value 0 if $\langle\phi|\psi\rangle = 0$.

If $|\psi\rangle$ is neither proportional nor orthogonal to $|\phi\rangle$, then

$$0 < |\langle\phi|\psi\rangle|^2 = \langle\psi|P_\phi|\psi\rangle < 1 .$$

The mean value of the dynamical variable "being in the state represented by $|\phi\rangle$" is then strictly included between 0 and 1. The *overlap* of $|\psi\rangle$ and $|\phi\rangle$ is the absolute square of their scalar product or, equivalently, the mean value of P_ϕ in $|\psi\rangle$.

We should point out that the five interpretation rules stated above do not uniquely determine the description of a quantum system. Indeed let $|\psi(t)\rangle$ be the state vector of the system and let A be the Hermitian operator associated with a dynamical variable. Let U be a unitary operator. Let us perform on $|\psi(t)\rangle$ and A an active transformation by means of U, in other words let us define $|\psi'(t)\rangle$ and A' so that

$$|\psi'(t)\rangle = U|\psi(t)\rangle , \qquad A' = UAU^\dagger . \tag{3.7}$$

We saw that A' is Hermitian, and one easily checks that the mean values of A', $(A')^2$, $(A')^3$, ... in $|\psi'(t)\rangle$ coincide with the mean values of A, $(A)^2$, $(A)^3$, ... in $|\psi(t)\rangle$. Eigenvalues of A' coincide with those of A and eigenvectors $|a_i'\rangle$ of A' are related to eigenvectors $|a_i\rangle$ of A through

---

[2] Note that all predictions in quantum mechanics boil down to the computation of mean values $\langle\psi|O|\psi\rangle$, where O is a Hermitian operator. Indeed eigenvalues of O (that is, possible measurement results) can be written as $\langle\chi|O|\chi\rangle$, where χ is a normalized eigenvector of O. Measurement result probabilities are given by (3.6). Generalization of this observation to nonisolated systems is immediate (Chap. 12). It is shown in [9] that quantum mechanics obeys the axioms of the theory of probability.

$$|a_i'\rangle = U|a_i\rangle . \tag{3.8}$$

All this implies that predictions of measurement results, whether made with the primed or the unprimed description, are the same. The two descriptions are called *equivalent.*

3.3 Hamiltonian

The most important dynamical variable in quantum mechanics is energy. The corresponding operator is called the *Hamiltonian* and is denoted by H. In Sect. 3.4 we will see that the Hamiltonian determines the dynamics of a quantum system. It also determines, of course, possible results of measurement of energy.

The Hamiltonian's specific form depends on the quantum system. In particular, it can be a function of time or not. Intuitively, we can think the Hamiltonian will depend on time if the system's dynamics is controlled by time-dependent agents. Thus the Hamiltonian of an atom which in no way interacts with its environment does not depend on time. The Hamiltonian of an atom in a harmonic electromagnetic field, on the other hand, is a function of time.

Eigenvalues of H are denoted by E_i. The symbol $|E_i\rangle$ is used for eigenvectors. The spectral decomposition of the Hamiltonian can therefore be written as

$$H = \sum_{i=1}^{N} E_i |E_i\rangle\langle E_i| . \tag{3.9}$$

The Hamiltonian has dimensions of energy. Thus its matrix elements, as well as its eigenvalues, can be expressed in joules (or in electron-volts).

3.4 The Schrödinger Equation

Now that the fundamental interpretation rules of quantum mechanics have been formulated, we state the general equation that governs the dynamics of an isolated system.

Postulate Let $H(t)$ be the Hamiltonian of a quantum system and let $|\psi(t)\rangle$ be the state vector. The time evolution of $|\psi(t)\rangle$ satisfies the *Schrödinger equation*:

$$\mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H(t)|\psi(t)\rangle . \tag{3.10}$$

The symbol $\hbar$ denotes Planck's constant divided by 2π. Its numerical value is equal to 1.055×10^{-34} J s.[3]

Let us investigate, from a mathematical point of view, the nature of Schrödinger's equation. For this purpose, we will write it in matrix form. Let $\{|u_i\rangle\}$ be an orthonormal basis of constant vectors. We can write

$$|\psi(t)\rangle = \sum_{i=1}^{N} c_i(t)|u_i\rangle\,, \tag{3.11}$$

where, as indicated, the coefficients $c_i(t)$ depend on time. We substitute (3.11) in (3.10) and project the resulting equation on $\langle u_j|$. Inserting the identity operator we find that

$$\mathrm{i}\hbar\langle u_j|\frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{i=1}^{N} c_i(t)|u_i\rangle\right] = \langle u_j|H(t)\left[\sum_{k=1}^{N}|u_k\rangle\langle u_k|\right]\left[\sum_{i=1}^{N} c_i(t)|u_i\rangle\right].$$

Consequently,

$$\mathrm{i}\hbar\sum_{i=1}^{N}\frac{\mathrm{d}}{\mathrm{d}t}c_i(t)\langle u_j|u_i\rangle = \sum_{k=1}^{N}\sum_{i=1}^{N}\langle u_j|H(t)|u_k\rangle c_i(t)\langle u_k|u_i\rangle\,,$$

from which we get

$$\mathrm{i}\hbar\sum_{i=1}^{N}\frac{\mathrm{d}}{\mathrm{d}t}c_i(t)\delta_{ji} = \sum_{k=1}^{N}\sum_{i=1}^{N}\langle u_j|H(t)|u_k\rangle c_i(t)\delta_{ki}\,.$$

Making use of properties of Kronecker's delta we finally obtain

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}c_j(t) = \sum_{i=1}^{N} H_{ji}(t)c_i(t)\,. \tag{3.12}$$

Note that functions $H_{ji}(t)$ are known, whereas coefficients $c_i(t)$ are not known. In its matrix formulation, Schrödinger's equation is therefore a system of N coupled first-order differential equations for the unknowns $c_i(t)$. In differential equations theory, one shows that knowledge of the N coefficients at some arbitrary time t_0 determines a unique solution. This means that, at least in principle, the Schrödinger equation yields the state vector $|\psi(t)\rangle$ at all time t if it is known at a specific instant t_0.

The foregoing observations imply that for all t_0 and t, there exists an operator $U(t, t_0)$ such that

$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle\,. \tag{3.13}$$

---

[3] The precise value of $\hbar$ is given on p. XV.

Here $|\psi(t_0)\rangle$ is an arbitrary vector and $|\psi(t)\rangle$ is the unique vector obtained by solving the Schrödinger equation in the time interval (t_0, t), with $|\psi(t_0)\rangle$ as initial condition. $U(t, t_0)$ is called the *evolution operator.*

One cannot overestimate the importance of the evolution operator in quantum mechanics. We will immediately derive its main properties. As the notation suggests, the evolution operator depends on both arguments t and t_0. We note that $U(t, t_0)$ is defined for $t < t_0$ as well as for $t > t_0$. Furthermore, we see that $U(t, t_0)$ is a linear operator. This is easily shown from the linearity of Schrödinger's equation.

It is clear that for any t

$$U(t,t) = I\,. \tag{3.14}$$

Let t' be some arbitrary time. We have

$$|\psi(t')\rangle = U(t', t_0)|\psi(t_0)\rangle$$

and

$$|\psi(t)\rangle = U(t, t')|\psi(t')\rangle\,,$$

from which we get

$$|\psi(t)\rangle = U(t, t')U(t', t_0)|\psi(t_0)\rangle\,.$$

This equation being true for all $|\psi(t_0)\rangle$, we conclude that

$$U(t, t_0) = U(t, t')U(t', t_0)\,. \tag{3.15}$$

Letting $t_0 = t$ in (3.15), we see immediately that the evolution operator has an inverse and that

$$[U(t, t')]^{-1} = U(t', t)\,. \tag{3.16}$$

We will now show that the evolution operator is unitary. For this it is enough to check that action of the operator preserves scalar products. In the present case, this amounts to showing that the scalar product $\langle\phi(t)|\psi(t)\rangle$, where $|\phi(t)\rangle$ and $|\psi(t)\rangle$ are obtained from two arbitrary vectors $|\phi(t_0)\rangle$ and $|\psi(t_0)\rangle$ through the evolution operator, does not depend on time. Indeed we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\langle\phi(t)|\psi(t)\rangle &= \left\{\frac{\mathrm{d}}{\mathrm{d}t}\langle\phi(t)|\right\}|\psi(t)\rangle + \langle\phi(t)|\left\{\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle\right\}\\ &= \left\{\frac{1}{-\mathrm{i}\hbar}\langle\phi(t)|H\right\}|\psi(t)\rangle + \langle\phi(t)|\left\{\frac{1}{\mathrm{i}\hbar}H|\psi(t)\rangle\right\}\\ &= 0\,.\end{aligned}$$

♣

This result entails that the state vector norm is preserved by time evolution. In other words, a state vector that is normalized at t_0 remains normalized at any other time t. We should also point out that all properties we just derived for the evolution operator are true for any time dependence of the Hamiltonian.

In the formalism we introduced, the time evolution of a system is represented by the fact that the state vector depends on time. Operators associated with dynamical variables are then usually time independent. There exists an equivalent description wherein the state vector is constant and operators depend on time. Indeed let us define

$$|\psi'(t)\rangle = U(t_0, t)|\psi(t)\rangle \tag{3.17}$$

and

$$A'(t) = U(t_0, t)A(t)U^\dagger(t_0, t)\,. \tag{3.18}$$

It is easy to see that $|\psi'(t)\rangle = |\psi(t_0)\rangle$ for all t. The description of a system by a constant state vector and time-dependent operators is called the *Heisenberg picture*. Here, however, we shall always work in the *Schrödinger picture*, the description we have hitherto developed.

3.5 Time-Independent Hamiltonian

Often the Hamiltonian of a quantum system does not depend on time. This happens, for instance, with atoms, molecules or isolated ions that are not acted upon by an electromagnetic field.

Let us consider an arbitrary quantum system whose Hamiltonian H is time independent. The Schrödinger equation is given by

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H|\psi(t)\rangle\,. \tag{3.19}$$

We first look for a particular solution of that equation. A vector can depend on time in different ways. In the simplest case, it can be the product of a constant vector by a time-dependent scalar function. Accordingly, let us look for a solution like $|\psi(t)\rangle = f(t)|\phi\rangle$, where $f(t)$ is a complex function of t and $|\phi\rangle$ is a constant vector. On the one hand

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left\{f(t)|\phi\rangle\right\} = \mathrm{i}\hbar\left(\frac{\mathrm{d}f}{\mathrm{d}t}\right)|\phi\rangle\,.$$

On the other hand

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left\{f(t)|\phi\rangle\right\} = H\left\{f(t)|\phi\rangle\right\} = f(t)H|\phi\rangle\,.$$

Thus

$$\mathrm{i}\hbar \frac{1}{f(t)} \frac{\mathrm{d}f}{\mathrm{d}t} |\phi\rangle = H|\phi\rangle ,$$

from which we obtain

$$\mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t} \{\ln f(t)\} |\phi\rangle = H|\phi\rangle .$$

The right-hand side of this equation is constant. So must be the left-hand side. This means that

$$\mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t} \{\ln f(t)\} = \text{const} = E .$$

The general solution of this differential equation is given by

$$f(t) = C \exp\left\{-\frac{\mathrm{i}Et}{\hbar}\right\} ,$$

where C is a constant. So we find that

$$H|\phi\rangle = E|\phi\rangle . \tag{3.20}$$

Equation (3.20) is nothing else than the eigenvalue equation for the Hamiltonian. Its solutions are

$$|\phi\rangle = |E_i\rangle , \qquad E = E_i ,$$

where $|E_i\rangle$ and E_i are the orthonormal eigenvectors and the eigenvalues of H.

To sum up, we have shown that the Schrödinger equation has solutions like

$$|\psi(t)\rangle = f(t)|\phi\rangle = C_i \exp\left\{-\frac{\mathrm{i}E_i t}{\hbar}\right\} |E_i\rangle . \tag{3.21}$$

This is true for every eigenvalue E_i and every eigenvector $|E_i\rangle$ of the Hamiltonian. States represented by (3.21) are called *stationary.* The state vector then evolves through a multiplicative factor only, so that it always represents the same quantum state.

For the stationary state (3.21) it is easy to show that at any time t

$$\langle H\rangle_\psi = E_i, \qquad (\Delta H)_\psi = 0 . \tag{3.22}$$

Such a solution therefore represents, at any t, a quantum state whose energy is well-defined and equal to E_i.

Eigenvectors of the Hamiltonian represent quantum states whose energy is well-defined. Moreover, in the case where H is time independent, their time evolution is simply given through an oscillating complex factor of unit norm. We will now show that, when the Hamiltonian is time independent, knowledge of the eigenvectors of H is sufficient to obtain the general solution of Schrödinger's equation. For all these reasons, the computation of eigenvectors of the Hamiltonian is the central problem of quantum mechanics.

3.6 Solution of the Schrödinger Equation

The Schrödinger equation is linear. This means that any linear combination of solutions is also a solution. In other words for any set of N complex constants C_i, the following expression is a solution of Schrödinger's equation:

$$|\psi(t)\rangle = \sum_{i=1}^{N} C_i \mathrm{e}^{-\mathrm{i}E_i t/\hbar}|E_i\rangle . \tag{3.23}$$

Suppose that at $t = t_0$, the state of the quantum system is represented by the vector $|\psi(t_0)\rangle$. Pick N constants C_i such that

$$C_i = \mathrm{e}^{\mathrm{i}E_i t_0/\hbar}\langle E_i|\psi(t_0)\rangle .$$

Equation (3.23) then becomes

$$|\psi(t)\rangle = \sum_{i=1}^{N} \langle E_i|\psi(t_0)\rangle \mathrm{e}^{-\mathrm{i}E_i(t-t_0)/\hbar}|E_i\rangle . \tag{3.24}$$

Here is an expression that, at any time t, solves the Schrödinger equation and that, at $t = t_0$, coincides with $|\psi(t_0)\rangle$. Since $|\psi(t_0)\rangle$ is arbitrary, (3.24) is the general solution of Schrödinger's equation.

Let us sum up what we have found. When the Hamiltonian is time independent, it is enough, for predicting the time evolution of a quantum system, to know the eigenvalues and eigenvectors of H.

Equation (3.24) can be written slightly differently as

$$|\psi(t)\rangle = \left[\sum_{i=1}^{N} \mathrm{e}^{-\mathrm{i}E_i(t-t_0)/\hbar}|E_i\rangle\langle E_i|\right] |\psi(t_0)\rangle . \tag{3.25}$$

Let us compare this expression with (3.13), which defines the evolution operator. We see immediately that

$$U(t,t_0) = \sum_{i=1}^{N} \mathrm{e}^{-\mathrm{i}E_i(t-t_0)/\hbar}|E_i\rangle\langle E_i| . \tag{3.26}$$

The spectral decomposition of the Hamiltonian, together with (2.50), allows one to rewrite (3.26) as

$$U(t,t_0) = \exp\left\{-\frac{\mathrm{i}(t-t_0)}{\hbar}H\right\} . \tag{3.27}$$

This is the general form of the evolution operator when the Hamiltonian is time independent. One easily checks that this expression has all the properties we derived in Sect. 3.4 and that the result of its action on an arbitrary ket obeys the Schrödinger equation.

When the Hamiltonian is time dependent, the general solution of the Schrödinger equation can no longer be represented by (3.24). Nonetheless, the evolution operator can be represented formally. We set $t = t_n$ and divide the interval (t_0, t_n) into n equal subintervals labeled by instants $t_l = t_0 + l\Delta t$. Here $\Delta t = (t_n - t_0)/n$. If Δt is small enough, the Hamiltonian is essentially time independent in the interval $(t_l - \Delta t, t_l)$, so that we can write

$$U(t_l, t_l - \Delta t) \approx \exp\left\{-\frac{\mathrm{i}\Delta t}{\hbar} H(t_l)\right\} \approx I - \frac{\mathrm{i}\Delta t}{\hbar} H(t_l).$$

We have neglected terms of order $(\Delta t)^2$. But we have

$$U(t_n, t_0) = \prod_{l=n}^{1} U(t_l, t_l - \Delta t) \approx \prod_{l=n}^{1} \left\{I - \frac{\mathrm{i}\Delta t}{\hbar} H(t_l)\right\}.$$

Neglected terms vanish as $\Delta t \to 0$, that is, as $n \to \infty$. So we find that

$$U(t, t_0) = \lim_{n\to\infty} \prod_{l=n}^{1} \left\{I - \frac{\mathrm{i}(t - t_0)}{n\hbar} H(t_l)\right\}. \tag{3.28}$$

This reduces to (3.27) when the Hamiltonian is time independent, as follows from (2.52).

3.7 Commuting Operators

Let A be a Hermitian operator corresponding to some dynamical variable. Suppose that at t the state vector $|\psi(t)\rangle$ is an eigenvector of A with eigenvalue a. The mean value of A in $|\psi\rangle$ is then equal to a and the standard deviation vanishes. This means that at t, a measurement of A will yield value a with probability equal to 1. We say that the system is in a state where dynamical variable A is *well-defined.*

Now let two Hermitian operators A and B correspond to two dynamical variables. We know that a basis of orthonormal eigenvectors of A and B can be found if and only if they commute. Hence if A and B commute, there exists an orthonormal basis of vectors representing states where both dynamical variables are well-defined.

The commutator appears in two important results summarized in the following theorems.

Theorem 1 Let A be a dynamical variable associated with a system whose Hamiltonian is time independent and assume that $[A, H] = 0$. Suppose also that at some time t_0, $A|\psi(t_0)\rangle = a|\psi(t_0)\rangle$. Then at any time t, $A|\psi(t)\rangle = a|\psi(t)\rangle$.

Proof Since A commutes with H, A commutes with every function of H. Therefore,

$$\begin{aligned} A|\psi(t)\rangle &= AU(t,t_0)|\psi(t_0)\rangle = A\exp\left\{-\frac{\mathrm{i}(t-t_0)}{\hbar}H\right\}|\psi(t_0)\rangle \\ &= \exp\left\{-\frac{\mathrm{i}(t-t_0)}{\hbar}H\right\}A|\psi(t_0)\rangle = U(t,t_0)a|\psi(t_0)\rangle \\ &= aU(t,t_0)|\psi(t_0)\rangle = a|\psi(t)\rangle\,. \end{aligned}$$ ♣

This result means that if an operator commutes with H, states wherein the associated dynamical variable is well-defined remain so.

Theorem 2 Let $|\psi(t)\rangle$ be the state vector of a quantum system and A a dynamical variable. The time derivative of the mean value of A in $|\psi\rangle$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle_\psi = \frac{1}{\mathrm{i}\hbar}\langle[A,H]\rangle_\psi + \left\langle\frac{\mathrm{d}A}{\mathrm{d}t}\right\rangle_\psi\,. \tag{3.29}$$

Proof

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\langle A\rangle_\psi &= \frac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)|A|\psi(t)\rangle \\ &= \left[\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)|\right]A|\psi(t)\rangle + \langle\psi(t)|\frac{\mathrm{d}A}{\mathrm{d}t}|\psi(t)\rangle + \langle\psi(t)|A\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle \\ &= \frac{1}{-\mathrm{i}\hbar}\langle\psi(t)|HA|\psi(t)\rangle + \langle\psi(t)|\frac{\mathrm{d}A}{\mathrm{d}t}|\psi(t)\rangle + \langle\psi(t)|A\frac{1}{\mathrm{i}\hbar}H|\psi(t)\rangle \\ &= \frac{1}{\mathrm{i}\hbar}\langle\psi(t)|[A,H]|\psi(t)\rangle + \langle\psi(t)|\frac{\mathrm{d}A}{\mathrm{d}t}|\psi(t)\rangle \\ &= \frac{1}{\mathrm{i}\hbar}\langle[A,H]\rangle_\psi + \left\langle\frac{\mathrm{d}A}{\mathrm{d}t}\right\rangle_\psi\,. \end{aligned}$$ ♣

This theorem has an important corollary. If A does not explicitly depend on time (that is, $\mathrm{d}A/\mathrm{d}t = 0$) and if A and H commute, then the mean value of A is constant.

3.8 The Hydrogen Molecular Ion

The positive ion of the hydrogen molecule (H_2^+) is a quantum system made of two protons and one electron. We know that the electron mass is approximately 1836 times smaller than the proton mass. This means that the electron is much more mobile than each of the two protons. Accordingly we shall first consider a configuration in which the distance r between the two protons is fixed (Fig. 3.1).

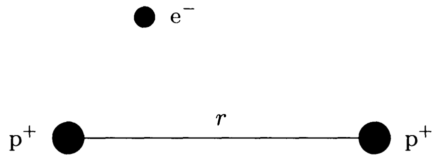

Fig. 3.1. Two protons separated by a distance r with one electron

The quantum problem of the H_2^+ ion is complicated. A quantitative analysis, like the one we will carry out in Chap. 20, requires the introduction of an infinite-dimensional state space. A qualitative treatment is possible, however, under certain plausible hypotheses and with the introduction of a two-dimensional state space. This is the road we will follow here.[4]

We assume the state space has an orthonormal basis made up of kets $|u_1\rangle$ and $|u_2\rangle$. These vectors represent the following states:

$|u_1\rangle$: state in which the electron is around proton 1 ;

$|u_2\rangle$: state in which the electron is around proton 2 .

According to the formalism of quantum mechanics, the electron state is not necessarily represented by $|u_1\rangle$ or $|u_2\rangle$. It can be represented by any linear combination of $|u_1\rangle$ and $|u_2\rangle$. In this situation, the electron has some probability to be found around proton 1, and some probability to be found around proton 2.

We point out that the electron position can be associated with the operator

$$O = \frac{r}{2}\left\{-|u_1\rangle\langle u_1| + |u_2\rangle\langle u_2|\right\} . \tag{3.30}$$

The mean value of O is equal to $-r/2$ in $|u_1\rangle$ and to $r/2$ in $|u_2\rangle$. It falls between $-r/2$ and $r/2$ if the state of the system is a linear combination of $|u_1\rangle$ and $|u_2\rangle$.

The fundamental question we must now ask is this: What is the system's Hamiltonian? In our simplified treatment, H is represented by a Hermitian 2×2 matrix. The most general Hermitian 2×2 matrix can be written as

$$\begin{pmatrix} E & a' \\ (a')^* & E' \end{pmatrix},$$

where E and E' are real numbers and a' is complex.

In the present case, H is not represented by the most general Hermitian 2×2 matrix. Indeed the two protons are identical. The mean value of the system's energy must therefore be the same whether the electron is around proton 1 or proton 2. Hence

---

[4] Several quantum systems (in particular, the H_2^+ ion) are treated with the help of two-dimensional state spaces in [85], Chaps. 9 and 10.

$$\langle H\rangle_{u_1} = \langle u_1|H|u_1\rangle = \langle H\rangle_{u_2} = \langle u_2|H|u_2\rangle . \tag{3.31}$$

Thus $H_{11} = H_{22}$, which means that $E = E'$. As regards the constant a', we can always choose the phase of $|u_1\rangle$ and $|u_2\rangle$ so that a' is real and negative. We write $a' = -a$ with a positive, so that the hamiltonian can be represented as

$$[H] = \begin{pmatrix} E & -a \\ -a & E \end{pmatrix} . \tag{3.32}$$

Eigenvalues and eigenvectors of H are obtained immediately. We find

$$\begin{aligned} E_1 &= E - a\,, \qquad |E_1\rangle = \frac{1}{\sqrt{2}}\{|u_1\rangle + |u_2\rangle\}\,; \\ E_2 &= E + a\,, \qquad |E_2\rangle = \frac{1}{\sqrt{2}}\{|u_1\rangle - |u_2\rangle\}\,. \end{aligned} \tag{3.33}$$

These vectors represent the system's two stationary states. The vector $|E_1\rangle$ represents the lowest-energy state.[5] A more complete treatment shows that in this state, the electron has a high probability to be found between the two protons (this is the meaning of the + sign between $|u_1\rangle$ and $|u_2\rangle$). The electrostatic energy thus receives an important negative contribution. In state $|E_2\rangle$, on the other hand, the electron is rarely between the two protons. Negative contributions to electrostatic energy are less important.

If the electron is not in a stationary state, the probability that it is near proton 1 or proton 2 changes with time. For example let $|\psi(t = 0)\rangle = |u_1\rangle$. One easily shows (Exercise 3.4) that $|\psi(t)\rangle$ oscillates from $|u_1\rangle$ to $|u_2\rangle$ to $|u_1\rangle$, and so on, with a period equal to $2\pi\hbar/a$.

We have solved the problem of the H_2^+ ion for a given value of the distance between the protons. To investigate the ion's stability, it is necessary to know how matrix elements E and a change with r. If r is very large, the state represented by $|u_1\rangle$ really corresponds to a hydrogen atom on one side (say, in its ground state) and a proton on the other side. Thus $E(r = \infty)$ is equal (up to a sign) to the binding energy of a hydrogen atom, which is about -13.6 eV. Proton 2 simply sees a neutral atom. If r is not infinite but nonetheless large, proton 2 sees an electric dipole. As r decreases, the electron less and less screens the proton's electrostatic repulsion. The constant E therefore increases, and is proportional to $1/r$ for small r. The constant a, on the other hand, is always positive and behaves more or less like $\mathrm{e}^{-\alpha r}$ if r is large.

The curves of $E_1 = E - a$ and $E_2 = E + a$ as functions of r will be computed in Chap. 20. They are plotted in Fig. 3.2. We see that E_2 is a strictly decreasing function of r, whereas E_1 has a minimum. So there is a value of r (call it r_0) for which the configuration of two protons and one electron in state $|E_1\rangle$ is stable. The constant r_0 represents the internuclear distance in the H_2^+ ion. The state $|E_2\rangle$ is stable at no value of r.

---

[5] The lowest energy state of a system is called the *ground* (or *fundamental*) *state*. Others are *excited states*.

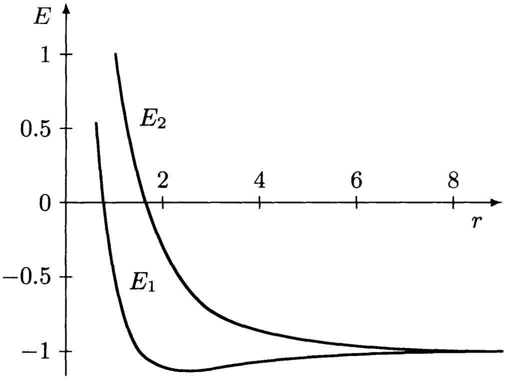

Fig. 3.2. E_1 and E_2 as functions of r. Energy and distance are expressed in atomic units of 13.6 eV and 0.53 Å respectively

The energy required to dissociate the H_2^+ ion into a proton and a hydrogen atom is given by

$$E_D = E_1(\infty) - E_1(r_0) = -13.6\,\mathrm{eV} - E_1(r_0). \tag{3.34}$$

This energy is positive, because $|E_1(r_0)| > 13.6\,\mathrm{eV}$.

Let us point out that the simplified treatment of the H_2^+ ion can be adapted to asymmetric ions such as HLi^+. In that case the Hamiltonian has the more general form

$$[H] = \begin{pmatrix} E & -a \\ -a & E' \end{pmatrix}, \tag{3.35}$$

where $E' \neq E$.

Exercises

3.1. A dynamical variable is represented in the $|u_i\rangle$ basis by the following Hermitian matrix, whose eigenvalues and eigenvectors were obtained in Exercise (2.13):

$$[A] = \begin{pmatrix} 0 & \mathrm{i}/\sqrt{2} & -\mathrm{i}/\sqrt{2} \\ -\mathrm{i}/\sqrt{2} & 1 & 1 \\ \mathrm{i}/\sqrt{2} & 1 & 1 \end{pmatrix}.$$

At time t, the system's state vector is given by $|\psi\rangle = |u_3\rangle$.

a) What are the possible results of a measurement of A and what are the corresponding probabilities?

b) What is the mean value of A and what is the standard deviation of A in state $|\psi\rangle$?

3.2. The Hermitian operator associated with a dynamical variable A is defined in the $|u_i\rangle$ basis as

$$[A] = \begin{pmatrix} 3 & \sqrt{6} & -\sqrt{3} \\ \sqrt{6} & 4 & \sqrt{2} \\ -\sqrt{3} & \sqrt{2} & 5 \end{pmatrix} .$$

a) Show that the eigenvalues and eigenvectors of the operator A are given by

$$a_1 = 0 , \qquad |a_1\rangle = \frac{1}{\sqrt{2}}|u_1\rangle - \frac{1}{\sqrt{3}}|u_2\rangle + \frac{1}{\sqrt{6}}|u_3\rangle ;$$

$$a_2 = 6 , \qquad |a_2\rangle = \frac{1}{\sqrt{3}}|u_2\rangle + \sqrt{\frac{2}{3}}|u_3\rangle ;$$

$$a_3 = 6 , \qquad |a_3\rangle = \frac{1}{\sqrt{2}}|u_1\rangle + \frac{1}{\sqrt{3}}|u_2\rangle - \frac{1}{\sqrt{6}}|u_3\rangle .$$

b) At $t = 0$, the state of the system is represented by the vector $|\psi\rangle = |u_3\rangle$. Find, at that time, the probability of obtaining each possible result of a measurement of A; find also the mean value and standard deviation of A.

3.3. Let H be a time-independent Hamiltonian and let $|E_1\rangle$ and $|E_2\rangle$ be two orthonormal eigenvectors of H with eigenvalues E_1 and E_2. Let $|\psi(t)\rangle$ be given by ($C_1 \neq 0 \neq C_2$):

$$|\psi(t)\rangle = C_1 \mathrm{e}^{-\mathrm{i}E_1 t/\hbar}|E_1\rangle + C_2 \mathrm{e}^{-\mathrm{i}E_2 t/\hbar}|E_2\rangle .$$

a) Compute the variance of H in $|\psi\rangle$.
b) Show that $\mathrm{Var}(H) = 0$ if and only if $E_1 = E_2$.

3.4. The simplified Hamiltonian of the H_2^+ ion is written as

$$[H] = \begin{pmatrix} E & -a \\ -a & E \end{pmatrix} ,$$

where E and a are real. At $t = 0$, the state of the system is represented by the vector $|\psi(t = 0)\rangle = |u_1\rangle$. What is the state vector at any time t?

3.5. The hamiltonian of a quantum system and a dynamical variable K are represented by the following matrices, whose eigenvalues and eigenvectors were obtained in Exercise (2.21):

$$[H] = \begin{pmatrix} E & 0 & 0 \\ 0 & -E & 0 \\ 0 & 0 & -E \end{pmatrix} , \qquad [K] = \begin{pmatrix} b & 0 & 0 \\ 0 & 0 & b \\ 0 & b & 0 \end{pmatrix} .$$

Suppose that at $t = 0$, the probability that a measurement of K yields value $-b$ is equal to 1.
a) What is the state vector at $t = 0$?
b) What is the probability that a measurement of K yields $-b$ at any other time t?

3.6. Let H and K be defined as in Exercise (3.5). Suppose that at $t = 0$ the state vector is given by

$$|\psi(t=0)\rangle = \frac{1}{\sqrt{2}}\left\{|u_1\rangle - |u_2\rangle\right\}.$$

a) Find $|\psi(t)\rangle$ at any time t.
b) Compute the mean value of K in $|\psi(t)\rangle$ and check that it does not depend on time.

3.7. The Hamiltonian of a system is represented by

$$[H] = \begin{pmatrix} 2 & -3\sqrt{2} & 3\sqrt{2} \\ -3\sqrt{2} & -1 & -3 \\ 3\sqrt{2} & -3 & -1 \end{pmatrix} \text{eV}.$$

a) Show that its eigenvalues and eigenvectors are given by

$$E_1 = 8\,\text{eV}\,, \qquad |E_1\rangle = \frac{-1}{\sqrt{2}}|u_1\rangle + \frac{1}{2}|u_2\rangle - \frac{1}{2}|u_3\rangle\,;$$

$$E_2 = -4\,\text{eV}\,, \qquad |E_2\rangle = \frac{1}{\sqrt{2}}|u_2\rangle + \frac{1}{\sqrt{2}}|u_3\rangle\,;$$

$$E_3 = -4\,\text{eV}\,, \qquad |E_3\rangle = \frac{1}{\sqrt{2}}|u_1\rangle + \frac{1}{2}|u_2\rangle - \frac{1}{2}|u_3\rangle\,.$$

b) At $t = 0$, the state of the system is represented by the vector $|\psi(t=0)\rangle = |u_1\rangle$. What is the state vector $|\psi(t)\rangle$ at time t?
c) What are the mean value $\langle H\rangle$ and standard deviation ΔH of the dynamical variable H in state $|\psi(t)\rangle$?

3.8. Let H, E_i and $|E_i\rangle$ be defined as in Exercise (3.7). Let K be an operator defined as

$$K = |E_1\rangle\langle E_1| + |E_2\rangle\langle E_2|\,.$$

a) What are the eigenvalues and eigenvectors of K?
b) Show that K and H make up a complete set of commuting operators.
c) What are the matrix elements of K in the $|u_i\rangle$ basis?

3.9. A quantum system is governed by a Hamiltonian H whose matrix representation in the $|u_i\rangle$ basis is the following:

$$[H] = \begin{pmatrix} E & 0 & 0 \\ 0 & E' & b \\ 0 & b & E' \end{pmatrix}.$$

Here E, E' and b are real.

a) Show that the eigenvalues and eigenvectors of H are given by

$$E_1 = E\,, \qquad |E_1\rangle = |u_1\rangle\,;$$

$$E_2 = E' + b\,, \qquad |E_2\rangle = \frac{1}{\sqrt{2}}\left\{|u_2\rangle + |u_3\rangle\right\}\,;$$

$$E_3 = E' - b\,, \qquad |E_3\rangle = \frac{1}{\sqrt{2}}\left\{|u_2\rangle - |u_3\rangle\right\}\,.$$

b) At $t = 0$, the state vector of the system is given by $|\psi(t = 0)\rangle = |u_2\rangle$. What is the state vector $|\psi(t)\rangle$ at time t?
c) What are the mean value and standard deviation of H in the state $|\psi(t)\rangle$ found in (b)?

3.10. A dynamical variable A is represented in the $|u_i\rangle$ basis by the following Hermitian matrix:

$$[A] = \begin{pmatrix} 0 & \mathrm{i} & 0 & 0 \\ -\mathrm{i} & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

a) What are the possible results of a measurement of A?
b) At time t_0, the state vector of the system is given by

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left\{|u_2\rangle + |u_3\rangle\right\}.$$

What is, at time t_0, the probability that a measurement of A yields the value 2?
c) What is the mean value of A in $|\psi\rangle$?
d) Find a Hermitian operator B that makes up with A (but not by itself) a complete set of commuting operators.

4 Spin and Magnetic Moment

Of all dynamical variables defined in a finite-dimensional state space, spin is no doubt the most important. Spin is associated with particles, atoms and molecules. It is connected with angular momentum and magnetic moment. Historically, atomic magnetic moments were revealed in the Stern–Gerlach experiment, which we will describe schematically. Analysis of this experiment will lead, through natural hypotheses and the investigation of spatial rotations, to a complete characterization of spin.

4.1 The Stern–Gerlach Experiment

In 1922 O. Stern and W. Gerlach set up an experiment to measure the magnetic moment of silver atoms. The arrangement is shown schematically in Figs. 4.1 and 4.2, which represent two different cross sections of the apparatus.[1]

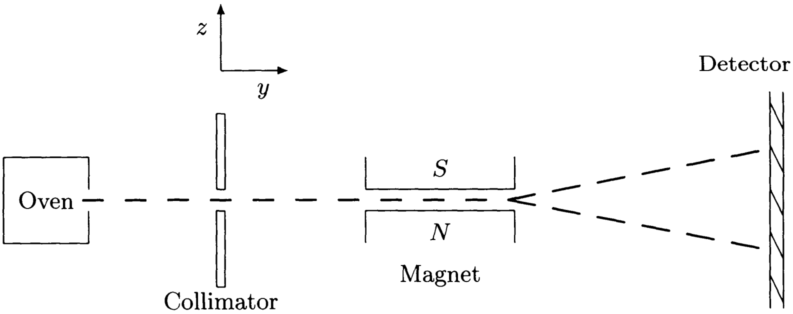

Fig. 4.1. Cross section of the Stern–Gerlach apparatus in the beam's plane

Silver atoms come out of the oven in a high vacuum and a collimator filters them into an essentially threadlike beam. The beam then passes between the

---

[1] A careful analysis of spin based on the Stern–Gerlach device can be found in [85], Chaps. 5 and 6. See also [53], Chap. 4 and, for historical aspects and reference to original papers, [125], Sect. 3.4.

L. Marchildon, *Quantum Mechanics*

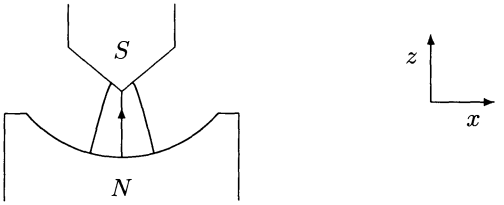

Fig. 4.2. Cross section of the Stern–Gerlach apparatus in a plane perpendicular to the incoming beam

poles of a magnet. The magnetic field configuration is shown in Fig. 4.2. The field has an intense B_z component, a not so strong B_x component (that vanishes in the middle) and a B_y component that vanishes for all practical purposes. Derivatives $\partial B_z/\partial z$ and $\partial B_x/\partial x$ are both appreciable so that $\boldsymbol{\nabla}\cdot\boldsymbol{B}$ everywhere vanishes.

Let us try to follow, between the magnet's poles, the trajectory of a silver atom carrying a magnetic moment $\boldsymbol{\mu}$. Classically, the atom's magnetic energy is given by $-\boldsymbol{\mu}\cdot\boldsymbol{B}$. This means that the magnetic force on the atom is equal to

$$\boldsymbol{F} = -\boldsymbol{\nabla}\left(-\boldsymbol{\mu}\cdot\boldsymbol{B}\right) = \boldsymbol{\nabla}\left\{\mu_x B_x + \mu_z B_z\right\}. \tag{4.1}$$

In Sect. 4.5 we will see that the component of magnetic moment perpendicular to the field direction rapidly oscillates about zero. Here the field is essentially along the z axis. It means that on the average μ_x vanishes, so that for all practical purposes this component does not contribute to the force exerted on the atom. The component μ_z, on the other hand, is essentially constant. Hence

$$\boldsymbol{F} = \mu_z \boldsymbol{\nabla} B_z = \mu_z \frac{\partial B_z}{\partial z}\hat{z}, \tag{4.2}$$

where variations of B_z along x and y have been neglected. As shown on Fig. 4.2, the z derivative of B_z is positive. A silver atom will thus be deflected upward if the μ_z component of its magnetic moment is positive, and downward if μ_z is negative.

Classically, we expect that silver atoms come out of the oven with magnetic moments $\boldsymbol{\mu}$ oriented in all directions. The μ_z component should therefore take all positive and negative values in an appropriate interval. Consequently, silver atoms in the incoming beam should, upon leaving the magnet, spread along a narrow stripe on the detector. But this is not what Stern and Gerlach observed. The incoming beam was rather separated in two threadlike beams respectively directed toward the upper and the lower part of the detector. Everything happened as if the μ_z component could take only two values. The rather symmetric deviation of the two beams with respect to the y axis suggested that the two possible values of the μ_z component have the same magnitude and opposite signs.

Evidently, the z axis here is not singled out in any absolute way. It is defined by the magnetic field direction. The Stern–Gerlach experiment indicates that the projection of a silver atom's magnetic moment on the magnetic field direction is quantized. Subsequently, the experiment was carried out with other atoms. In every case it was observed that the projection of $\boldsymbol{\mu}$ on $\boldsymbol{B}$ is quantized. The number of possible values of μ_z, however, is not always equal to two. For some atoms there is only one value (that is, only one beam coming out of the magnet), for some there are three, for others even more.

Let us try to interpret the results of the Stern–Gerlach experiment in terms of quantum-mechanical concepts. The dynamical variable μ_z, the component of an atom's magnetic moment in the direction of the magnetic field, can take a finite number of values only. Thus in quantum mechanics, it seems that magnetic moment can be described by means of a finite-dimensional state space. The dimension is certainly not smaller than the number of values μ_z can take.

The properties of magnetic moment emerge even better if the beam of atoms is directed toward two consecutive Stern–Gerlach devices. Assume the magnetic field in the first apparatus is oriented toward the z axis, while in the second one the field is oriented toward an axis u. Let us lead toward the second apparatus one of the beams coming out of the first magnet. What do we observe? In the particular case where $\hat{u} = \pm\hat{z}$ [2] (i.e. the field in both apparatus is, within a sign, oriented in the same direction), the beam falling on the second apparatus comes out undivided. Moreover if $\hat{u} = \hat{z}$, it is deflected exactly as in the first apparatus. If, on the other hand, $\hat{u} \neq \pm\hat{z}$, the beam falling on the second apparatus is itself divided into several beams, in general the same number (call it N) as the number of values of μ_z for the atom considered.

For atoms corresponding to a given value of N, one observes that the probability that an atom coming out of the first apparatus with a value μ_z subsequently comes out of the second one with a value μ_u depends only on the relative orientation of the unit vectors $\hat{z}$ and $\hat{u}$.

All these observations, at least from a qualitative point of view, fit the concepts of quantum mechanics rather well. Components of an atom's magnetic moment are dynamical variables associated with an N-dimensional state space. Denote by $|\mu_z^\alpha\rangle$ the vector associated with a state wherein μ_z is well-defined and has a value μ_z^α. Here the exponent α denotes each beam coming out of the Stern–Gerlach device. An atom that comes out in the α beam of an apparatus oriented toward $\hat{z}$ is in a state represented by the vector $|\mu_z^\alpha\rangle$.

The set $\{|\mu_z^\alpha\rangle,\ \alpha = 1,\ 2,\ \ldots,\ N\}$ makes up an orthonormal basis of the state space. Likewise we can denote by $\{|\mu_u^\beta\rangle,\ \beta = 1,\ 2,\ \ldots,\ N\}$ the orthonormal basis associated with states wherein the dynamical variable μ_u is well-defined. The vectors $|\mu_u^\beta\rangle$ can be written as linear combinations of the $|\mu_z^\alpha\rangle$, that is,

---

[2] The circumflex ($\wedge$) indicates a unit three-dimensional real vector.

$$|\mu_u^\beta\rangle = \sum_{\alpha=1}^{N} U_{\alpha\beta} |\mu_z^\alpha\rangle\,, \tag{4.3}$$

where the coefficients $U_{\alpha\beta}$ are matrix elements of a unitary operator that depends on the relative orientation of $\hat{z}$ and $\hat{u}$.

What is the probability that an atom coming out of the first apparatus with a value μ_z^α comes out of the second one with a value μ_u^β? Upon leaving the first apparatus the atom is in a state represented by $|\mu_z^\alpha\rangle$. One can think of the second apparatus as performing a measurement of the dynamical variable μ_u. If, therefore, quantum mechanics correctly describes phenomena associated with magnetic moment, the probability that an atom coming out of the first apparatus with a value μ_z^α comes out of the second one with a value μ_u^β should be equal to $|\langle\mu_u^\beta|\mu_z^\alpha\rangle|^2 = |U_{\alpha\beta}|^2$.

The purpose of the present chapter is to construct explicitly the Hermitian operators associated with the components of magnetic moment, in all state spaces where they can be defined. Moreover we shall investigate, in the simplest cases, the way to obtain the coefficients $U_{\alpha\beta}$ in (4.3), that is, the connection between eigenvectors of two different components of magnetic moment. The connection only depends on the relative orientation of the two axes which serve to define the components.

Operators associated with magnetic moment components are closely related to spatial rotations, which we will now investigate.

4.2 Spatial Rotations

Consider an active rotation by an angle θ about the z axis, which transforms a vector $\boldsymbol{r}$ into a vector $\boldsymbol{r}'$ (Fig. 4.3). The connection between the components of $\boldsymbol{r}$ and $\boldsymbol{r}'$ is given by

$$\begin{aligned} x' &= x\cos\theta - y\sin\theta\,, \\ y' &= x\sin\theta + y\cos\theta\,, \\ z' &= z\,, \end{aligned}$$

or, in matrix form,

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{4.4}$$

For an infinitesimal rotation by an angle $\delta\theta$ we have, to first order,

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} 1 & -\delta\theta & 0 \\ \delta\theta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \{I - \mathrm{i}\delta\theta[I_z]\} \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \tag{4.5}$$

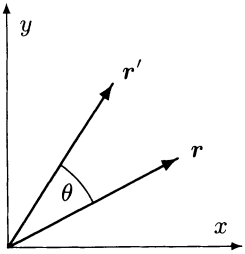

Fig. 4.3. Rotation by an angle θ about the z axis

where the matrix $[I_z]$ is defined as

$$[I_z] = \begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{4.6}$$

$[I_z]$ is called the *generator of a rotation* about the z axis.

The matrix $[I_z]$ is very important. It can parametrize not only an infinitesimal rotation about the z axis, but also a finite rotation. This is established in the following lemma.

Lemma The matrix corresponding to a rotation by θ about the z axis is equal to $\exp\{-\mathrm{i}\theta[I_z]\}$.

Proof One easily sees that

$$[I_z]^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad [I_z]^3 = [I_z]$$

and that for any positive integer n,

$$[I_z]^{2n} = [I_z]^2, \qquad [I_z]^{2n-1} = [I_z].$$

Making use of the series expansion of the exponential we find that

$$\begin{aligned} \exp\{-\mathrm{i}\theta[I_z]\} &= \sum_{l=0}^{\infty} \frac{1}{l!}\left(-\mathrm{i}\theta[I_z]\right)^l \\ &= I + \sum_{\text{odd } l} \frac{1}{l!}\left(-\mathrm{i}\theta[I_z]\right)^l + \sum_{\text{even } l>0} \frac{1}{l!}\left(-\mathrm{i}\theta[I_z]\right)^l \\ &= I + [I_z] \sum_{\text{odd } l} \frac{(-\mathrm{i}\theta)^l}{l!} + [I_z]^2 \sum_{\text{even } l>0} \frac{(-\mathrm{i}\theta)^l}{l!} \\ &= I + [I_z]\left(-\mathrm{i}\sin\theta\right) + [I_z]^2\left(\cos\theta - 1\right) \\ &= \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

♣

By a similar procedure, one defines two matrices $[I_x]$ and $[I_y]$ that are generators of rotations about the x and y axes:

$$[I_x] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix}, \qquad [I_y] = \begin{pmatrix} 0 & 0 & \mathrm{i} \\ 0 & 0 & 0 \\ -\mathrm{i} & 0 & 0 \end{pmatrix}. \tag{4.7}$$

One shows (Exercise 4.1) that

$$\exp\{-\mathrm{i}\theta[I_x]\} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \tag{4.8}$$

$$\exp\{-\mathrm{i}\theta[I_y]\} = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}. \tag{4.9}$$

Matrices $[I_x]$, $[I_y]$ and $[I_z]$ correspond to Hermitian operators acting in a three-dimensional vector space. Hence they can be associated with dynamical variables. These variables are, up to a multiplicative constant, the components μ_x, μ_y and μ_z of the magnetic moment of an atom which, going through a Stern–Gerlach device, is deflected in three different directions.

One easily checks that matrices $[I_x]$, $[I_y]$ and $[I_z]$ obey the following commutation relations:

$$[[I_x], [I_y]] = \mathrm{i}[I_z], \tag{4.10}$$

$$[[I_y], [I_z]] = \mathrm{i}[I_x], \tag{4.11}$$

$$[[I_z], [I_x]] = \mathrm{i}[I_y]. \tag{4.12}$$

Already we can say that these commutation relations are closely connected with spatial rotations. This statement, however, will be clarified fully in Chaps. 13 and 14 only. Meanwhile, Exercise (4.2) will help apprehending its meaning.

Hermitian matrices $[I_x]$, $[I_y]$ and $[I_z]$, obeying commutation relations (4.10)–(4.12), can be built not only in a three-dimensional vector space, but also in spaces of any finite dimension N. This is what we will do in Sect. 4.3. Suppose that an atom, going through a Stern–Gerlach device, is deflected in N different directions. Its dynamical variables μ_x, μ_y and μ_z are then proportional to operators corresponding to Hermitian matrices $[I_x]$, $[I_y]$ and $[I_z]$ defined in a vector space of N dimensions.

4.3 Generators of Rotation

Let $\mathcal{V}$ be an N-dimensional vector space. We shall investigate the possibility of finding three Hermitian operators I_x, I_y and I_z, defined in $\mathcal{V}$ and obeying the commutation relations (4.10)–(4.12). By extension, these operators are also called *generators of rotation.*

We first point out that the triplet of operators, if it exists, is certainly not unique. Indeed let U be a unitary operator defined in $\mathcal{V}$. One easily sees that the three operators I'_x, I'_y and I'_z defined as

$$I'_x = UI_xU^{-1}, \qquad I'_y = UI_yU^{-1}, \qquad I'_z = UI_zU^{-1}, \tag{4.13}$$

are also Hermitian and satisfy the relations (4.10)–(4.12). Moreover let $(I_x(1), I_y(1), I_z(1))$ and $(I_x(2), I_y(2), I_z(2))$ be two triplets of Hermitian operators defined in spaces $\mathcal{V}^1$ and $\mathcal{V}^2$ with dimensions N_1 and N_2, and satisfying the relations (4.10)–(4.12). One then checks that operators

$$I_x = I_x(1) \oplus I_x(2)\,, \tag{4.14}$$

$$I_y = I_y(1) \oplus I_y(2)\,, \tag{4.15}$$

$$I_z = I_z(1) \oplus I_z(2)\,, \tag{4.16}$$

defined in the space $\mathcal{V}^1 \oplus \mathcal{V}^2$ with dimension $N_1 + N_2$, are also Hermitian and satisfy (4.10)–(4.12).

A triplet (I_x, I_y, I_z) of Hermitian operators defined in $\mathcal{V}$ is called *irreducible* if no proper subspace $\mathcal{V}'$ of $\mathcal{V}$ is invariant under the action of the three operators.[3] Otherwise the triplet is *reducible.* Later in this section, we will show that all reducible triplets can be written in the form (4.14)–(4.16). Hence the problem stated at the beginning of this section can be reformulated as follows: Find, within a unitary transformation, all irreducible triplets of Hermitian operators obeying the commutation rules (4.10)–(4.12).

Let us first define two operators I_+ and I_- as

$$I_+ = I_x + \mathrm{i}I_y\,, \qquad I_- = I_x - \mathrm{i}I_y\,. \tag{4.17}$$

One immediately checks that I_+ and I_- are not Hermitian but that

$$I_+ = (I_-)^\dagger\,. \tag{4.18}$$

From (4.10)–(4.12) and (4.17) one easily finds that I_+, I_- and I_z obey the following commutation relations:

$$[I_z, I_+] = I_+\,, \qquad [I_z, I_-] = -I_-\,, \qquad [I_+, I_-] = 2I_z\,. \tag{4.19}$$

The operator I_z is Hermitian. Thus its eigenvalues are real and its eigenvectors corresponding to distinct eigenvalues are orthogonal. Let j be the largest eigenvalue of I_z and let $|j\rangle$ be the corresponding normalized eigenvector. For the time being we assume that j is not degenerate. We will examine later the case where it is. Consider the vector $I_-|j\rangle$. We have

---

[3] A subspace $\mathcal{V}'$ is *invariant under the action of an operator* L if for every $|\psi\rangle$ in $\mathcal{V}'$ the vector $L|\psi\rangle$ belongs to $\mathcal{V}'$.

$$\begin{aligned} I_z \{I_-|j\rangle\} &= \{I_-I_z + [I_z, I_-]\} |j\rangle \\ &= I_-j|j\rangle - I_-|j\rangle \\ &= (j-1)\{I_-|j\rangle\}. \end{aligned}$$

Thus $I_-|j\rangle$ is an eigenvector of I_z with eigenvalue $j-1$. We note, however, that $I_-|j\rangle$ can vanish. If it does not, it can be normalized. Let us denote by $|j-1\rangle$ the normalized vector which is proportional to $I_-|j\rangle$ and whose phase coincides with the phase of $I_-|j\rangle$. The proportionality factor between $|j-1\rangle$ and $I_-|j\rangle$ is then real and positive.

In a similar way, one shows that $I_-|j-1\rangle$ is an eigenvector of I_z with eigenvalue $j-2$, and so on. In this manner a sequence $|j\rangle$, $|j-1\rangle$, $|j-2\rangle, \dots,$ $|j_{\text{min}}\rangle$ of orthonormal vectors is constructed. The vector $|j_{\text{min}}\rangle$ is such that $|j_{\text{min}}\rangle \neq 0$ and

$$I_-|j_{\text{min}}\rangle = 0. \tag{4.20}$$

The existence of j_{min} is ensured from the fact that there cannot be more than N orthogonal vectors in an N-dimensional vector space. At this stage, however, we do not know whether the number of vectors built as above is equal to or lower than N.

Orthonormal vectors $\{|m\rangle\}$ make up a basis of a subspace $\mathcal{V}'$ of $\mathcal{V}$. In this basis, the matrix representation of I_z is diagonal. From the above discussion we know that

$$I_-|m\rangle = C_m|m-1\rangle, \tag{4.21}$$

where the constants C_m are real and positive, with the exception of $C_{j_{\text{min}}}$ which vanishes. Matrix elements of I_+ for any vectors $|m'\rangle$ et $|m\rangle$ are thus given by (2.13) as

$$\langle m'|I_+|m\rangle = \langle m|I_-|m'\rangle^* = C_{m'}\delta_{m,m'-1} = C_{m+1}\delta_{m',m+1}. \tag{4.22}$$

From (4.22) it is tempting to conclude that

$$I_+|m\rangle = C_{m+1}|m+1\rangle. \tag{4.23}$$

If such is the case it is clear that the subspace $\mathcal{V}'$ is invariant under the action of the three operators I_z, I_+ and I_-, and hence also of I_x and I_y.

To prove that (4.23) follows from (4.22), it is enough to show that $I_+|m\rangle$ belongs to $\mathcal{V}'$ for every $|m\rangle$. Let us proceed by induction. We first point out that $I_+|j\rangle$ vanishes (otherwise it would be an eigenvector of I_z with eigenvalue $j+1$), and thus that $I_+|j\rangle$ belongs to $\mathcal{V}'$. Suppose $I_+|m\rangle$ belongs to $\mathcal{V}'$. Then for $m \neq j_{\text{min}}$,

$$\begin{aligned} I_+|m-1\rangle &= I_+\left\{(C_m)^{-1} I_-|m\rangle\right\} \\ &= (C_m)^{-1}\{I_-I_+ + [I_+, I_-]\}|m\rangle \\ &= (C_m)^{-1}\{I_-C_{m+1}|m+1\rangle + 2I_z|m\rangle\} \\ &= (C_m)^{-1}\left\{(C_{m+1})^2 + 2m\right\}|m\rangle\,. \end{aligned}$$

This completes the proof by induction. Thus (4.23) is true for all m. As a by-product of the proof we have shown that

$$(C_m)^2 = (C_{m+1})^2 + 2m\,. \tag{4.24}$$

We will soon obtain the values of the coefficients C_m. But let us first examine the questions pending. The triplet (I_x, I_y, I_z) is irreducible in the space $\mathcal{V}'$. This is proved as in Exercise (4.3), where one checks that repeated action of I_+ and I_- on any vector of $\mathcal{V}'$ generates a basis of $\mathcal{V}'$. If $\mathcal{V}'$ coincides with $\mathcal{V}$, we have found an irreducible triplet in $\mathcal{V}$. If $\mathcal{V}'$ is strictly included in $\mathcal{V}$, then $\mathcal{V} = \mathcal{V}' \oplus \mathcal{V}_0$, where $\mathcal{V}_0$ is the orthogonal complement of $\mathcal{V}'$, i.e. the space made up of vectors orthogonal to all vectors in $\mathcal{V}'$. Let $|\bar{\jmath}\rangle$ be the eigenvector (or one of the eigenvectors) of I_z, in $\mathcal{V}_0$, which corresponds to the largest eigenvalue $\bar{\jmath}$. Evidently $\bar{\jmath} \leq j$. Repeated action of I_- on $|\bar{\jmath}\rangle$ produces vectors $|\bar{\jmath}-1\rangle, \ldots, |\bar{\jmath}_{\min}\rangle$ that make up a basis of a subspace $\mathcal{V}''$ of $\mathcal{V}_0$, in which the triplet (I_x, I_y, I_z) is irreducible. Keeping up, we decompose $\mathcal{V}$ into a direct sum $\mathcal{V}' \oplus \mathcal{V}'' \oplus \cdots$ of vector spaces, on each of which the triplet (I_x, I_y, I_z) is irreducible. This takes care of an eventual degeneracy of the eigenvalue j and proves that all reducible triplets can be written in the form (4.14)–(4.16).

Let us now turn to the coefficients C_m. From (4.23) we see that $C_{j+1} = 0$. This is enough to solve the recurrence equation (4.24). We find

$$C_m = \{(j+m)(j+1-m)\}^{1/2}\,. \tag{4.25}$$

From the above discussion $C_{j_{\min}} = 0$, which is possible only if $j_{\min} = -j$. Allowed values of m are thus included between $-j$ and j and are each separated by one unit. In all, there are $2j+1 = N$ values of m. This entails that allowed values of j are nonnegative integers or half-integers.

We can now put together the results we have found. For any positive integer N there exists an N-dimensional vector space in which an irreducible triplet (I_x, I_y, I_z) can be defined. This triplet is unique up to a unitary transformation. Eigenvalues of I_z go from $-j$ to $+j$ by unit steps, where $j = (N-1)/2$. Matrix elements of operators I_z, I_+ and I_- in the basis of eigenvectors of I_z are given by

$$\langle m'|I_z|m\rangle = m\delta_{m'm}\,, \tag{4.26}$$

$$\langle m'|I_-|m\rangle = \{(j+m)(j+1-m)\}^{1/2}\,\delta_{m',m-1}\,, \tag{4.27}$$

$$\langle m'|I_+|m\rangle = \{(j-m)(j+1+m)\}^{1/2}\,\delta_{m',m+1}\,. \tag{4.28}$$

Matrix elements of operators I_x and I_y easily follow from (4.17), (4.27) and (4.28).

In closing this section we give matrix elements of operators I_x, I_y and I_z in the $|m\rangle$ basis, for vector spaces of 1, 2 and 3 dimensions. Rows and columns are indexed from $+j$ to $-j$. For $j = 0$ we trivially have

$$[I_x] = [0], \qquad [I_y] = [0], \qquad [I_z] = [0].$$

For $j = 1/2$,

$$[I_x] = \frac{1}{2}\sigma_x \equiv \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{4.29}$$

$$[I_y] = \frac{1}{2}\sigma_y \equiv \frac{1}{2}\begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \tag{4.30}$$

$$[I_z] = \frac{1}{2}\sigma_z \equiv \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{4.31}$$

Here we have introduced three 2×2 matrices denoted by σ_x, σ_y and σ_z and called the *Pauli spin matrices.* In the forthcoming chapters we will use them repeatedly. For $j = 1$ we finally have (Exercise 4.4)

$$[I_x] = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{4.32}$$

$$[I_y] = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & -\mathrm{i} & 0 \\ \mathrm{i} & 0 & -\mathrm{i} \\ 0 & \mathrm{i} & 0 \end{pmatrix}, \tag{4.33}$$

$$[I_z] = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \tag{4.34}$$

Note that these matrices are related to the three generators defined in Sect. 4.2. Indeed if $[I'_x]$, $[I'_y]$ and $[I'_z]$ denote the matrices given in (4.7) and (4.6), then these are connected with matrices in (4.32)–(4.34) through (4.13), where $[U]$ is the unitary matrix given by

$$[U] = \frac{1}{\sqrt{2}}\begin{pmatrix} \mathrm{i} & 0 & -\mathrm{i} \\ -1 & 0 & -1 \\ 0 & -\sqrt{2}\mathrm{i} & 0 \end{pmatrix}.$$

4.4 Spin, Magnetic Moment and the Zeeman Effect

We will denote by $\boldsymbol{I}$ the triplet of Hermitian operators (I_x, I_y, I_z). It is useful to introduce a second triplet of operators $\boldsymbol{J} = (J_x, J_y, J_z)$ related to the first one by

$$\boldsymbol{J} = \hbar \boldsymbol{I} . \tag{4.35}$$

The three dynamical variables associated with the operators J_x, J_y and J_z are the components of *spin.*

Spin is an intrinsic property of a particle. We will see later that it is akin to the angular momentum of a quantum particle. A particle has spin j if its operators J_x, J_y and J_z make up an irreducible triplet in a space of dimension $N = 2j + 1$.

The argument in Sect. 4.3 showed that a measurement of the dynamical variable J_z of a particle with spin j can only yield values $m\hbar$, where m is an integer or half-integer included between $-j$ and j. So is it with every dynamical variable $\hat{u} \cdot \boldsymbol{J}$, where $\hat{u}$ is a unit vector with real components. To prove this consider a triplet $(\hat{u}, \hat{v}, \hat{w})$ of orthonormal vectors oriented right-handedly. One can show (Exercise 4.6) that the three operators $\hat{u} \cdot \boldsymbol{I}$, $\hat{v} \cdot \boldsymbol{I}$ and $\hat{w} \cdot \boldsymbol{I}$ are Hermitian and obey the same commutation rules as I_x, I_y and I_z. These, however, are the only properties of I_x, I_y and I_z that were used to obtain the spectrum of I_z. Eigenvalues of $\hat{u} \cdot \boldsymbol{I}$ therefore coincide with those of I_z.

The dynamical variable $\hat{u} \cdot \boldsymbol{J}$ is the spin component on $\hat{u}$. Possible results of a measurement of $\hat{u} \cdot \boldsymbol{J}$ coincide with those of a measurement of J_z. It is clear, however, that eigenvectors of $\hat{u} \cdot \boldsymbol{J}$ do not, in general, coincide with eigenvectors of J_z (unless $\hat{u} = \pm \hat{z}$). Indeed for any $\hat{u}$ and any $\hat{v} \neq \pm \hat{u}$

$$[\hat{u} \cdot \boldsymbol{J}, \hat{v} \cdot \boldsymbol{J}\,] \neq 0 . \tag{4.36}$$

This means that spin components on different axes cannot share a basis of eigenvectors.

The operator $\boldsymbol{J} \cdot \boldsymbol{J} = J_x J_x + J_y J_y + J_z J_z$, on the other hand, commutes with all spin components. Indeed one easily checks (Exercise 4.7) that, for any $\hat{u}$,

$$[\boldsymbol{J} \cdot \boldsymbol{J}, \hat{u} \cdot \boldsymbol{J}\,] = 0 . \tag{4.37}$$

The following equation shows that $\boldsymbol{J} \cdot \boldsymbol{J}$ is, in fact, a multiple of the identity operator in the space of vectors $|m\rangle$:

$$\begin{aligned}
\boldsymbol{J} \cdot \boldsymbol{J}\,|m\rangle &= \hbar^2 \left(I_x I_x + I_y I_y + I_z I_z\right)|m\rangle \\
&= \hbar^2 \left(I_+ I_- - I_z + I_z I_z\right)|m\rangle \\
&= \hbar^2 \left\{\left(C_m\right)^2 - m + m^2\right\}|m\rangle \\
&= \hbar^2 \left\{(j + m)(j + 1 - m) - m + m^2\right\}|m\rangle \\
&= \hbar^2 j\,(j + 1)\,|m\rangle .
\end{aligned} \tag{4.38}$$

With the tools developed in the last two sections, we can now come back to the Stern–Gerlach experiment. Suppose a given atom can be deflected in

N different directions. The atom's magnetic moment components are then associated with an N-dimensional vector space. It is natural to think that these components are proportional to those of a spin operator for which $N = 2j + 1$. In fact[4]

$$\boldsymbol{\mu} = \frac{q_e}{2m_e} g \boldsymbol{J}, \tag{4.39}$$

where q_e and m_e are the electron's charge and mass and g is a number of order 1 called the *Landé factor*. Formally, this equation looks like the one that connects a charge's classical magnetic moment to its angular momentum. The reason why the electron's charge and mass (instead of the proton's or the nucleus's) enter (4.39) is that the electrons are mainly responsible for the magnetic moment of atoms. The Landé factor changes from one atom to another and, for a given atom, from one electronic state to another.

Magnetic moment is a multiple of spin. This means that eigenvectors of the operator associated with a component $\mu_u = \hat{u} \cdot \boldsymbol{\mu}$ of magnetic moment coincide with eigenvectors of $\hat{u} \cdot \boldsymbol{J}$. In principle we have all we need to express eigenvectors of $\hat{u} \cdot \boldsymbol{J}$ in terms of vectors $|m\rangle$ (i.e. in terms of eigenvectors of J_z). This calculation will be made in Sect. 4.6 for $j = 1/2$ and in Exercise (4.9) for $j = 1$. The general result for $j > 1$ will be obtained in Chap. 14 with the help of Lie groups theory.

Knowledge of eigenvectors of $\hat{u} \cdot \boldsymbol{J}$ in terms of vectors $|m\rangle$ allows one to predict the probability that an atom coming out of a Stern–Gerlach apparatus with a given value of μ_z comes out of a second apparatus with a given value of μ_u. These are the probabilities which in Sect. 4.1 were denoted by $|U_{\alpha\beta}|^2$.

The presence of a magnetic field changes energy levels in atoms and molecules. We will investigate this phenomenon in Chap. 18. It is nonetheless useful, at this stage, to examine a simple case. Let H_0 be the Hamiltonian of an atom in the absence of a magnetic field. Let $\mathcal{V}$ be the vector space made up of all eigenvectors of H_0 that correspond to a given eigenvalue E. With certain hypotheses the Hamiltonian (restricted to $\mathcal{V}$) of the atom in the presence of a magnetic field is given by

$$H = E - \boldsymbol{\mu} \cdot \boldsymbol{B} = E - \frac{q_e}{2m_e} g \boldsymbol{J} \cdot \boldsymbol{B}. \tag{4.40}$$

It is understood here that E multiplies the identity operator. The analogy between the quantum Hamiltonian and the classical energy of a dipole $\boldsymbol{\mu}$ in a magnetic field $\boldsymbol{B}$ is apparent. It is, however, only formal. In quantum mechanics components of $\boldsymbol{\mu}$ are operators, whereas they are real numbers in classical mechanics.

---

[4] In truth $\boldsymbol{J}$ should stand here for the atom's total angular momentum, restricted to an appropriate subspace of the state space. This concept will really be clarified in Chap. 14. But algebraic properties of total angular momentum are identical with those of spin.

Eigenvalues of the Hamiltonian in the presence of the field are given by

$$E - \left(\frac{q_e \hbar}{2m_e}\right) gm|\boldsymbol{B}|\,. \tag{4.41}$$

The second term in (4.41) represents the correction to the atom's energy levels produced by the magnetic field. This correction is proportional to the magnetic field intensity and it can take $2j+1$ equally-spaced values. It gives rise to the *Zeeman effect* in atomic spectra, to which we will come back.

The quantity $(-q_e)\hbar/2m_e$, which appears in (4.41), is called the *Bohr magneton* and is denoted by μ_B. Its numerical value is equal to 9.27×10^{-24} J T^{-1}.

4.5 The Larmor Precession

The mean value of each spin component satisfies a very simple evolution equation whenever a quantum system's Hamiltonian is given by an expression like (4.40). More generally, let us write

$$H = H_0 - \frac{\mu}{\hbar}\boldsymbol{J}\cdot\boldsymbol{B}\,, \tag{4.42}$$

where H_0, the system's Hamiltonian in the absence of a magnetic field, commutes with the components of $\boldsymbol{J}$. The magnetic field $\boldsymbol{B}$ is constant and space independent and the constant μ has dimensions of a magnetic moment.

Mean values of the three components J_x, J_y and J_z in state $|\psi(t)\rangle$ can collectively be denoted as $\langle \boldsymbol{J}\rangle$. We have

$$\langle \boldsymbol{J}\rangle = \langle \psi(t)|\boldsymbol{J}|\psi(t)\rangle\,. \tag{4.43}$$

Let us compute the time derivative of $\langle \boldsymbol{J}\rangle$. From (3.29),

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \boldsymbol{J}\rangle = \frac{1}{\mathrm{i}\hbar}\langle \psi(t)|\,[\boldsymbol{J},H]\,|\psi(t)\rangle\,. \tag{4.44}$$

It is easy to check that

$$[\boldsymbol{J},H] = \left[\boldsymbol{J}, -\frac{\mu}{\hbar}\boldsymbol{J}\cdot\boldsymbol{B}\right] = -\mathrm{i}\mu\boldsymbol{B}\times\boldsymbol{J}\,. \tag{4.45}$$

Therefore we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \boldsymbol{J}\rangle = -\frac{\mu}{\hbar}\boldsymbol{B}\times\langle \boldsymbol{J}\rangle\,. \tag{4.46}$$

Call $\hat{z}$ the magnetic field direction. Components of (4.46) can then be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle J_x \rangle = \frac{\mu B}{\hbar}\langle J_y \rangle \,, \tag{4.47}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle J_y \rangle = -\frac{\mu B}{\hbar}\langle J_x \rangle \,, \tag{4.48}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle J_z \rangle = 0 \,. \tag{4.49}$$

The solution of these equations is given by

$$\langle J_x \rangle_t = \langle J_x \rangle_0 \cos \omega t + \langle J_y \rangle_0 \sin \omega t \,, \tag{4.50}$$

$$\langle J_y \rangle_t = -\langle J_x \rangle_0 \sin \omega t + \langle J_y \rangle_0 \cos \omega t \,, \tag{4.51}$$

$$\langle J_z \rangle_t = \langle J_z \rangle_0 \,, \tag{4.52}$$

where the index zero denotes the mean value at time $t = 0$ and where

$$\omega = \frac{\mu B}{\hbar} \,. \tag{4.53}$$

The constant ω is called the *Larmor frequency.* From (4.50)–(4.52), we see that the mean value of the spin component along the magnetic field direction is constant. The mean value of the normal component of spin, on the other hand, performs a rotation, with angular frequency ω, in the plane perpendicular to the magnetic field. This phenomenon is called the *Larmor precession.* It is the analogue of the precession of a classical angular momentum in a constant and uniform magnetic field.

At this stage it is instructive to come back to the Stern–Gerlach experiment and to justify better the transition from (4.1) to (4.2). We assumed that on the average, the μ_x component of the magnetic moment vanishes between the magnet's poles. Between the poles the magnetic field is not strictly parallel to the z axis and it is not uniform. The magnetic moment of an atom is of order of the Bohr magneton and a typical magnetic field in a Stern–Gerlach device is approximately 1 T. Thus the Larmor frequency is about 10^{11} Hz. Atoms come out of the oven at approximately 500 m s^{-1}, which means they go through 3×10^{-8} m during one Larmor period. At that scale, the field is essentially uniform. Thus (4.50)–(4.52) correctly describe the time evolution of mean values of spin and magnetic moment components. An atom will remain approximately 2×10^{-4} s in a Stern–Gerlach device 0.1 m long. The $\langle \mu_x \rangle$ component will then oscillate 3×10^6 times about zero. Hence for all practical purposes we can say that $\langle \mu_x \rangle$ vanishes.

4.6 Spin 1/2

In Sect. 4.4 we saw that the operator associated with the $\hat{u}$ component of magnetic moment is proportional to $\hat{u} \cdot \boldsymbol{J}$ or, equivalently, to $\hat{u} \cdot \boldsymbol{I}$. The probability that an atom coming out of a Stern–Gerlach apparatus with a given value of μ_z comes out of a second apparatus with a given value of μ_u is related to the coefficients of the expansion of eigenvectors of $\hat{u}\cdot\boldsymbol{I}$ in linear combinations of eigenvectors of $I_z = \hat{z}\cdot\boldsymbol{I}$. These coefficients are the $U_{\alpha\beta}$ in (4.3).

We will now obtain these coefficients explicitly for $j = 1/2$. In this case I_z has two distinct eigenvectors denoted by $|+\frac{1}{2}\rangle$ and $|-\frac{1}{2}\rangle$. We know that eigenvalues of $\hat{u}\cdot\boldsymbol{I}$ coincide with those of I_z. Let us denote eigenvectors of $\hat{u}\cdot\boldsymbol{I}$ by $|m;\hat{u}\rangle$, where $m = \pm\frac{1}{2}$. Thus we have

$$(\hat{u}\cdot\boldsymbol{I})\,|m;\hat{u}\rangle = m|m;\hat{u}\rangle\,. \tag{4.54}$$

From (4.29)–(4.31) we see that matrix elements of the operator $\hat{u}\cdot\boldsymbol{I}$ in a basis $\{|m\rangle\}$ of eigenvectors of I_z are given by

$$[\hat{u}\cdot\boldsymbol{I}\,] = \frac{1}{2}\begin{pmatrix} u_z & u_x - \mathrm{i}u_y \\ u_x + \mathrm{i}u_y & -u_z \end{pmatrix}. \tag{4.55}$$

It is useful to introduce the polar angle $\bar{\theta}$ and the azimuthal angle $\bar{\phi}$ of the unit vector $\hat{u}$ with respect to the z axis. Thus

$$u_x = \sin\bar{\theta}\cos\bar{\phi}\,, \tag{4.56}$$
$$u_y = \sin\bar{\theta}\sin\bar{\phi}\,, \tag{4.57}$$
$$u_z = \cos\bar{\theta}\,, \tag{4.58}$$

so that (4.55) becomes

$$[\hat{u}\cdot\boldsymbol{I}\,] = \frac{1}{2}\begin{pmatrix} \cos\bar{\theta} & \sin\bar{\theta}\,\mathrm{e}^{-\mathrm{i}\bar{\phi}} \\ \sin\bar{\theta}\,\mathrm{e}^{\mathrm{i}\bar{\phi}} & -\cos\bar{\theta} \end{pmatrix}. \tag{4.59}$$

Making use of methods developed in Chap. 2, one easily finds that vectors $|m;\hat{u}\rangle$ are related to vectors $|m\rangle$ as

$$|m;\hat{u}\rangle = U|m\rangle = \sum_{m'=-1/2}^{1/2} U_{m'm}|m'\rangle\,, \tag{4.60}$$

where the coefficients $U_{m'm}$ are given by (Exercise 4.8)

$$[U] = \begin{pmatrix} \cos\frac{\bar{\theta}}{2} & -\sin\frac{\bar{\theta}}{2}\,\mathrm{e}^{-\mathrm{i}\bar{\phi}} \\ \sin\frac{\bar{\theta}}{2}\,\mathrm{e}^{\mathrm{i}\bar{\phi}} & \cos\frac{\bar{\theta}}{2} \end{pmatrix}. \tag{4.61}$$

Recall that these coefficients contain all the information on probabilities of going through two consecutive Stern–Gerlach devices. For instance, the probability that a spin 1/2 atom coming out of an apparatus with an eigenvalue of I_z equal to $+\frac{1}{2}$ comes out of a second apparatus with an eigenvalue of $\hat{u}\cdot\boldsymbol{I}$ equal to $+\frac{1}{2}$ is given by $|U_{1/2,1/2}|^2 = \cos^2(\bar{\theta}/2)$.

The discussion in Sect. 2.5 implies that we can write

$$[\hat{u}\cdot\boldsymbol{I}] = [U][I_z][U]^\dagger = [U]\,[\hat{z}\cdot\boldsymbol{I}]\,[U]^\dagger , \tag{4.62}$$

which can at any rate be checked directly. It is very instructive to rewrite matrix (4.61) as

$$\begin{aligned}[U] &= \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\frac{\bar\phi}{2}} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\frac{\bar\phi}{2}} \end{pmatrix} \begin{pmatrix} \cos\frac{\bar\theta}{2} & -\sin\frac{\bar\theta}{2} \\ \sin\frac{\bar\theta}{2} & \cos\frac{\bar\theta}{2} \end{pmatrix} \begin{pmatrix} \mathrm{e}^{\mathrm{i}\frac{\bar\phi}{2}} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}\frac{\bar\phi}{2}} \end{pmatrix} \\ &= \exp\left(-\mathrm{i}\bar\phi\frac{\sigma_z}{2}\right)\exp\left(-\mathrm{i}\bar\theta\frac{\sigma_y}{2}\right)\exp\left(\mathrm{i}\bar\phi\frac{\sigma_z}{2}\right) . \end{aligned} \tag{4.63}$$

The second equality is proved through a series expansion of the exponentials, matrices σ_x, σ_y and σ_z being given in (4.29)–(4.31).

In Sect. 4.2 we saw that a matrix like $\exp\{-\mathrm{i}\bar\theta[I_z]\}$ represents a rotation by an angle $\bar\theta$ about the z axis. It was, of course, a rotation in a three-dimensional space, since $[I_z]$ was then a 3×3 matrix. We can now see that this observation also applies to the two-dimensional space of spin 1/2. The operator U in (4.3) transforms eigenvectors of I_z into eigenvectors of $\hat{u}\cdot\boldsymbol{I}$. In other words it transforms state vectors wherein dynamical variable $\hat{z}\cdot\boldsymbol{I}$ is well-defined into vectors wherein $\hat{u}\cdot\boldsymbol{I}$ is well-defined. Recall that $\bar\theta$ and $\bar\phi$ denote the polar and azimuthal angles of vector $\hat{u}$ with respect to vector $\hat{z}$. It is easy to see that $\hat{u}$ is obtained from $\hat{z}$ through the successive application of three rotations: (i) a rotation by an angle $-\bar\phi$ about the z axis; (ii) a rotation by an angle $\bar\theta$ about the y axis; and (iii) a rotation by an angle $\bar\phi$ about the z axis.[5] Let us examine the matrix representation of U given in (4.63). Matrix $[U]$ is the product of three matrices that precisely represent the three rotations we just mentioned, provided the exponentials of $\sigma_y/2$ and $\sigma_z/2$ represent rotations about the y and z axes, respectively. Matrices $\sigma_x/2$, $\sigma_y/2$ and $\sigma_z/2$ are thus connected with rotations in spin 1/2 state space in the same way as matrices $[I_x]$, $[I_y]$ and $[I_z]$, given in (4.6) and (4.7), are connected with rotations in ordinary three-dimensional space.

It is interesting to push the identification further and to write down matrix elements of an arbitrary rotation in a spin 1/2 state space. This can be done in two ways. The first one consists in parametrizing an arbitrary rotation by means of *Euler angles* (α,β,γ). We have, here, a rotation by an angle α about the z axis, followed by a rotation β about y, and finally by a rotation γ about z. Denote by $D^{(1/2)}_{mm'}(\alpha,\beta,\gamma)$ matrix elements of the corresponding rotation operator in the $|m\rangle$ basis. We have

$$D^{(1/2)}(\alpha,\beta,\gamma) = \exp\left(-\mathrm{i}\gamma\frac{\sigma_z}{2}\right)\exp\left(-\mathrm{i}\beta\frac{\sigma_y}{2}\right)\exp\left(-\mathrm{i}\alpha\frac{\sigma_z}{2}\right) . \tag{4.64}$$

---

[5] We point out that vector $\hat{u}$ will still be obtained from vector $\hat{z}$ if we replace rotation (i) by a rotation by an arbitrary angle about the z axis. This corresponds to the fact that the columns of matrix (4.61) can be multiplied by complex numbers of unit norm.

Evaluating the exponentials and matrix product yields the following result:

$$D^{(1/2)}(\alpha,\beta,\gamma) = \begin{pmatrix} \exp\left\{-\frac{\mathrm{i}}{2}(\alpha+\gamma)\right\}\cos\frac{\beta}{2} & -\exp\left\{\frac{\mathrm{i}}{2}(\alpha-\gamma)\right\}\sin\frac{\beta}{2} \\ \exp\left\{-\frac{\mathrm{i}}{2}(\alpha-\gamma)\right\}\sin\frac{\beta}{2} & \exp\left\{\frac{\mathrm{i}}{2}(\alpha+\gamma)\right\}\cos\frac{\beta}{2} \end{pmatrix}. \tag{4.65}$$

The second way to parametrize an arbitrary rotation consists in specifying by a unit vector $\hat{n}$ the axis about which the rotation is performed, and in specifying the rotation angle θ. Denote by $D^{(1/2)}_{mm'}(\hat{n};\theta)$ matrix elements of the corresponding rotation operator in the $|m\rangle$ basis. We have

$$D^{(1/2)}(\hat{n};\theta) = \exp\left(-\mathrm{i}\theta\hat{n}\cdot\frac{\boldsymbol{\sigma}}{2}\right), \tag{4.66}$$

where $\hat{n}\cdot\boldsymbol{\sigma}/2$ has been identified with the generator of a rotation about $\hat{n}$.[6] Observing that $(\hat{n}\cdot\boldsymbol{\sigma})^2$ is the identity matrix, one easily evaluates the exponential and finds

$$D^{(1/2)}(\hat{n};\theta) = \begin{pmatrix} \cos\frac{\theta}{2} - \mathrm{i}n_z\sin\frac{\theta}{2} & -(\mathrm{i}n_x+n_y)\sin\frac{\theta}{2} \\ (-\mathrm{i}n_x+n_y)\sin\frac{\theta}{2} & \cos\frac{\theta}{2}+\mathrm{i}n_z\sin\frac{\theta}{2} \end{pmatrix}. \tag{4.67}$$

Comparison of matrices (4.65) and (4.67) allows one to express Euler angles in terms of parmeters $(\hat{n};\theta)$ or vice versa.

We will come back to rotation operators in Chap. 14, from the point of view of Lie groups theory. In particular we will obtain the representation of rotation matrices in a state space of arbitrary spin j. From these follow the probabilities that atoms with arbitrary spin go through Stern–Gerlach devices oriented in any directions.

Exercises

4.1. Compute $\exp\{-\mathrm{i}\theta[I_x]\}$ and $\exp\{-\mathrm{i}\theta[I_y]\}$ where matrices $[I_x]$ and $[I_y]$ are given in (4.7).

4.2. Let A and B be two operators and α and β two infinitesimal parameters.
a) Expanding exponentials up to quadratic terms in infinitesimal quantities, show that approximately

$$\exp(\alpha A)\exp(\beta B)\exp(-\alpha A)\exp(-\beta B) = \exp\{\alpha\beta[A,B]\}.$$

b) Deduce that the product of (i) a rotation $\delta\theta$ about x, (ii) a rotation $\delta\bar{\theta}$ about y, (iii) a rotation $-\delta\theta$ about x and (iv) a rotation $-\delta\bar{\theta}$ about y coincides with a rotation about z.

---

[6] This will be justified in Sect. 14.1.

4.3. Let $\mathcal{V}'$ be the subspace defined just before (4.21). Show that repeated action of I_+ and I_- on any nonzero vector of $\mathcal{V}'$ generates a basis of $\mathcal{V}'$, and hence that the triplet (I_x, I_y, I_z) is irreducible in $\mathcal{V}'$.

4.4. Check that for $j = 1$, matrices $[I_x]$, $[I_y]$ and $[I_z]$ are given by (4.32)–(4.34). Show that these matrices obey the relation

$$[I_x]^2 + [I_y]^2 + [I_z]^2 = 2[I],$$

where $[I]$ is the identity.

4.5. Compute matrices $[I_x]$, $[I_y]$ and $[I_z]$ for $j = 3/2$ and show that

$$[I_x]^2+[I_y]^2+[I_z]^2 = \frac{3}{2}\left(\frac{3}{2}+1\right)[I] = \frac{15}{4}[I].$$

4.6. Let I_x, I_y and I_z be three Hermitian operators obeying the commutation rules (4.10)–(4.12); let $\boldsymbol{a}$ and $\boldsymbol{b}$ be two real vectors in three dimensions.
a) Show that

$$[\boldsymbol{a} \cdot \boldsymbol{I}, \boldsymbol{b} \cdot \boldsymbol{I}] = \mathrm{i}\,(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{I},$$

where $\boldsymbol{a} \times \boldsymbol{b}$ denotes the vector product.
b) Let $(\hat{u}, \hat{v}, \hat{w})$ be a triplet of orthonormal real vectors oriented right-handedly. Show that operators $\hat{u} \cdot \boldsymbol{I}$, $\hat{v} \cdot \boldsymbol{I}$ and $\hat{w} \cdot \boldsymbol{I}$ are Hermitian and obey the same commutation rules as I_x, I_y and I_z.

4.7. From (2.55)–(2.57) show that, for any vector $\hat{u}$,

$$[I_x I_x + I_y I_y + I_z I_z, \hat{u} \cdot \boldsymbol{I}] = 0.$$

4.8. Obtain (4.61).

4.9. Let $[I_x]$, $[I_y]$ and $[I_z]$ be spin matrices for $j = 1$ and let $\hat{u} = (u_x, u_y, u_z)$ be a unit vector. We have

$$\begin{aligned} \hat{u} \cdot [\boldsymbol{I}] &= u_x[I_x] + u_y[I_y] + u_z[I_z] \\ &= \begin{pmatrix} u_z & \frac{1}{\sqrt{2}}(u_x - \mathrm{i}u_y) & 0 \\ \frac{1}{\sqrt{2}}(u_x + \mathrm{i}u_y) & 0 & \frac{1}{\sqrt{2}}(u_x - \mathrm{i}u_y) \\ 0 & \frac{1}{\sqrt{2}}(u_x + \mathrm{i}u_y) & -u_z \end{pmatrix}. \end{aligned}$$

a) Show that eigenvalues of $\hat{u} \cdot [\boldsymbol{I}]$ are equal to $+1$, 0 and -1. Find the normalized eigenvectors of $\hat{u} \cdot [\boldsymbol{I}]$.
b) Let atoms come out in the $+1$ beam of a Stern–Gerlach apparatus oriented with respect to $\hat{z}$. Find the probabilities that these atoms end up in beams $+1$, 0 and -1 of a Stern–Gerlach apparatus oriented with respect to $\hat{y}$.

5 Particle in One Dimension

With a particle restricted to one space dimension, we begin the study of quantum systems with infinite-dimensional state spaces. The state space of a particle in one dimension is first introduced intuitively. We then carefully examine the dynamical variables position, momentum and energy, which leads to a precise definition of the state space. The Schrödinger equation is stated and its general solution obtained in terms of eigenvectors of the Hamiltonian. For potential wells and barriers and for the harmonic oscillator potential, eigenvalues and eigenvectors of the Hamiltonian are computed exactly. We finally investigate the time evolution of spatially localized states.

5.1 Basis Vectors $|x\rangle$

This chapter is devoted to quantum systems made of one spinless particle of mass m, with no internal degrees of freedom and restricted to one space dimension.

First we must ask, What is the state space of such a system? This question, in fact, is rather delicate. The answer will require an elaborate discussion, which we begin in this section and pursue in the next two.

At a given time, a quantum particle can be localized in an arbitrarily small spatial interval. Let us denote by $|x; \Delta x\rangle$ a vector representing a state wherein the particle is localized in the interval $(x, x + \Delta x)$. Clearly, two vectors $|x_1; \Delta x\rangle$ and $|x_2; \Delta x\rangle$ are orthogonal if the corresponding intervals do not overlap, that is, if $|x_1 - x_2| > \Delta x$.

Among all vectors $|x; \Delta x\rangle$, one can find an infinite number of mutually orthogonal ones. This, for instance, is the case with vectors like $|n\Delta x; \Delta x\rangle$, where n is an integer. Hence the vector space generated by vectors $|x; \Delta x\rangle$, and thus the state space of a particle in one dimension, have infinite dimension.

On closer look, however, we can see that vectors $|x; \Delta x\rangle$ are not entirely adequate to a precise treatment of a quantum particle. Indeed one does not really know where the particle is inside the specified interval. This may not be so bad if Δx is small enough. Nevertheless, it is much more convenient to introduce an idealization. Define a vector, say $|x\rangle$, which represents a state wherein the particle is precisely localized at point x. There is thus a vector

L. Marchildon, *Quantum Mechanics*

$|x\rangle$ associated with each spatial position x. Two vectors $|x\rangle$ and $|x'\rangle$ are orthogonal if $x \neq x'$, that is,

$$\langle x|x'\rangle = 0 \quad \text{if} \quad x \neq x' .$$

Note that we have introduced bras $\langle x|$, in one-to-one correspondence with kets $|x\rangle$, to write down the scalar products.

As a first approximation, we will say that the state space of a quantum particle in one dimension is generated by basis vectors $|x\rangle$. We shall soon realize, however, that this statement needs revision. The space so defined is much too large. In particular, it has a basis which is not only infinite but also nondenumerable, with a distinct vector $|x\rangle$ corresponding to each real number x. For the moment, however, we will consider formally the vector space generated by all kets $|x\rangle$.

The question of vector normalization is more delicate with kets $|x\rangle$ than in the case of finite-dimensional vector spaces. It is not appropriate here to set $\langle x|x\rangle = 1$. It is very convenient, however, to define normalization so that the resolution of the identity operator is written in terms of bras $\langle x|$ and kets $|x\rangle$ similarly as in the finite-dimensional case. We thus let

$$\int_{-\infty}^{\infty} \mathrm{d}x \, |x\rangle\langle x| = I . \tag{5.1}$$

Integration over $\mathrm{d}x$ has replaced discrete summation here because x is a continuous variable. The operator I is the identity in the space of vectors $|x\rangle$.[1]

Let us see how (5.1) determines the normalization of vectors $|x\rangle$. Projecting that equation on a ket $|x'\rangle$ and manipulating bras and kets in the usual way, we get

$$\left\{ \int_{-\infty}^{\infty} \mathrm{d}x \, |x\rangle\langle x| \right\} |x'\rangle = \int_{-\infty}^{\infty} \mathrm{d}x \, \{\langle x|x'\rangle\} \, |x\rangle = I|x'\rangle = |x'\rangle .$$

This equation is satisfied if

$$\langle x|x'\rangle = \delta(x - x') , \tag{5.2}$$

where $\delta(x - x')$ is the Dirac delta function.[2] Equation (5.2) defines the normalization of vectors $|x\rangle$ and is consistent with the orthogonality relation we obtained earlier.

The vector space generated by basis vectors $|x\rangle$ is, as mentioned before, a first approximation to the state space of a quantum particle in one dimension. That space has infinite dimension.[3] We will now define operators that will correspond to the most important dynamical variables of a quantum particle.

---

[1] Equation (5.1) entails that kets $|x\rangle$ and bras $\langle x|$ have dimensions of $\mathrm{m}^{-1/2}$.

[2] The main properties of the delta function are summarized in Sect. 5.9.1.

[3] Sect. 5.9.2 is devoted to elementary properties of infinite-dimensional vector spaces.

5.2 Position, Momentum and Energy

The position of a particle is a dynamical variable. Which Hermitian operator corresponds to it? According to the quantum-mechanical interpretation rules, eigenvectors of the position operator X should represent states wherein position is well-defined and corresponding eigenvalues should be possible results of position measurement. Hence the *position operator* is defined by the relation

$$X|x\rangle = x|x\rangle\,. \tag{5.3}$$

The position operator is diagonal in the $|x\rangle$ basis.[4]

As in classical mechanics, in quantum mechanics the momentum of a particle is a dynamical variable. Denote by P the momentum operator. What precisely is the form of P? The answer to this question is not at all obvious. The association of a specific operator with momentum is, again, an interpretation rule of the quantum-mechanical formalism. The postulate is the following: Matrix elements of the *momentum operator* in the $|x\rangle$ basis are given by

$$\langle x|P|x'\rangle = -\mathrm{i}\hbar\frac{\partial}{\partial x}\delta(x-x')\,. \tag{5.4}$$

It is very hard to motivate this postulate convincingly by an a priori argument. It finds justification in all the consequences derived from it.

At first sight operators X and P are Hermitian, as it should be with operators associated with dynamical variables. Indeed,

$$\begin{aligned}
\langle x|X|x'\rangle^* &= \{x'\delta(x-x')\}^* = x'\delta(x-x') \\
&= x\delta(x-x') = \langle x'|X|x\rangle\,, \\
\langle x|P|x'\rangle^* &= \left\{-\mathrm{i}\hbar\frac{\partial}{\partial x}\delta(x-x')\right\}^* = \mathrm{i}\hbar\frac{\partial}{\partial x}\delta(x-x') \\
&= -\mathrm{i}\hbar\frac{\partial}{\partial x'}\delta(x'-x) = \langle x'|P|x\rangle\,.
\end{aligned}$$

The question of the hermiticity of X and P is, however, more delicate than this argument may indicate. It will be investigated in detail in Sect. 5.3.

In contrast with the position operator, the momentum operator is not diagonal in the $|x\rangle$ basis. But we will see that there exists a basis where P is diagonal. The eigenvalue equation for P is written as

$$P|p\rangle = p|p\rangle\,. \tag{5.5}$$

The ket $|p\rangle$ represents a state wherein the particle's momentum is well-defined and equal to p. Two kets $|p\rangle$ and $|p'\rangle$ are orthogonal if $p \neq p'$.

---

[4] The position operator, like its eigenvalues, has dimensions of m.

What is the relation between vectors $|x\rangle$ and $|p\rangle$? If both sets of vectors belong to the same vector space, the latter should be expandable in linear combinations of the former. Thus

$$|p\rangle = I|p\rangle = \left\{ \int_{-\infty}^{\infty} \mathrm{d}x \, |x\rangle\langle x| \right\} |p\rangle = \int_{-\infty}^{\infty} \mathrm{d}x \, \langle x|p\rangle |x\rangle . \tag{5.6}$$

To determine coefficients $\langle x|p\rangle$, we will compute the matrix element $\langle x|P|p\rangle$ in two different ways. On the one hand

$$\langle x|P|p\rangle = \langle x|p|p\rangle = p\langle x|p\rangle .$$

On the other hand, introducing the identity operator between P and $|p\rangle$ and rearranging, we find

$$\begin{aligned}\langle x|P|p\rangle &= \int_{-\infty}^{\infty} \mathrm{d}x' \langle x|P|x'\rangle\langle x'|p\rangle \\ &= -\mathrm{i}\hbar \int_{-\infty}^{\infty} \mathrm{d}x' \left\{ \frac{\partial}{\partial x}\delta(x-x') \right\} \langle x'|p\rangle \\ &= -\mathrm{i}\hbar \frac{\partial}{\partial x} \int_{-\infty}^{\infty} \mathrm{d}x' \, \delta(x-x')\langle x'|p\rangle \\ &= -\mathrm{i}\hbar \frac{\partial}{\partial x}\langle x|p\rangle .\end{aligned}$$

Comparing both results, one sees that $\langle x|p\rangle$ obeys the following differential equation:

$$\frac{\partial}{\partial x}\langle x|p\rangle = \frac{\mathrm{i}p}{\hbar}\langle x|p\rangle . \tag{5.7}$$

The general solution of that equation is given by

$$\langle x|p\rangle = c_p \exp\left\{ \frac{\mathrm{i}px}{\hbar} \right\} , \tag{5.8}$$

where c_p is a constant that can depend on p. As there are no constraints on this variable, there exists a ket $|p\rangle$ for each value of p. Hence a particle's momentum can take any value between $-\infty$ and $+\infty$.

The value of c_p is related to the normalization condition on vectors $|p\rangle$. Indeed

$$\begin{aligned}\langle p|p'\rangle &= \int_{-\infty}^{\infty} \mathrm{d}x \, \langle p|x\rangle\langle x|p'\rangle = c_p^* c_{p'} \int_{-\infty}^{\infty} \mathrm{d}x \, \exp\left\{ \frac{\mathrm{i}x}{\hbar}(p'-p) \right\} \\ &= c_p^* c_{p'} 2\pi\delta\left\{ \frac{p'-p}{\hbar} \right\} = |c_p|^2 2\pi\hbar\delta(p-p') .\end{aligned} \tag{5.9}$$

Setting

$$\langle p|p'\rangle = \delta(p-p'), \tag{5.10}$$

we find that

$$|c_p|^2 = (2\pi\hbar)^{-1}. \tag{5.11}$$

One can write $c_p = (2\pi\hbar)^{-1/2}$, which corresponds to specifying the phase of kets $|p\rangle$. One thus obtains

$$\langle x|p\rangle = (2\pi\hbar)^{-1/2}\exp\left\{\frac{\mathrm{i}px}{\hbar}\right\}. \tag{5.12}$$

Equations (5.1) and (5.12) entail that (Exercise 5.3)

$$\int_{-\infty}^{\infty} \mathrm{d}p\,|p\rangle\langle p| = I. \tag{5.13}$$

This is the resolution of the identity in terms of vectors $|p\rangle$. Kets $|p\rangle$ therefore make up a basis of the state space.

We now prove a very important result about the commutator of operators X and P.

Theorem The commutator of X and P is given by

$$[X,P] = \mathrm{i}\hbar I. \tag{5.14}$$

Proof Let us compute matrix elements of the operator $[X,P]$. Making use of properties of the delta function, we have

$$\begin{aligned}\langle x|\,[X,P]\,|x'\rangle &= \langle x|XP-PX|x'\rangle = \langle x|xP|x'\rangle - \langle x|Px'|x'\rangle \\ &= x\langle x|P|x'\rangle - x'\langle x|P|x'\rangle = (x-x')(-\mathrm{i}\hbar)\frac{\partial}{\partial x}\delta(x-x') \\ &= \mathrm{i}\hbar\delta(x-x').\end{aligned}$$

The last term coincides with matrix elements of the operator $i\hbar I$. ♣

The most important dynamical variable of a quantum system is energy. In terms of position and momentum operators, the *Hamiltonian* is written in the same way as energy, in classical mechanics, is written in terms of the position and momentum coordinates. Thus

$$H = \frac{1}{2m}P^2 + V(X), \tag{5.15}$$

where $V(X)$ is a function of the position operator.[5] Matrix elements of H in the $|x\rangle$ basis are given by

---

[5] The function V can also depend on time, so that $V = V(X,t)$. Because time is a parameter (not an operator), formulas for matrix elements and commutators remain the same.

$$\begin{aligned}
\langle x|H|x'\rangle &= \frac{1}{2m}\langle x|P^2|x'\rangle + \langle x|V(X)|x'\rangle \\
&= \frac{1}{2m}\int_{-\infty}^{\infty} \mathrm{d}x''\langle x|P|x''\rangle\langle x''|P|x'\rangle + \langle x|V(x)|x'\rangle \\
&= \frac{1}{2m}\int_{-\infty}^{\infty} \mathrm{d}x'' \left[-\mathrm{i}\hbar\frac{\partial}{\partial x}\delta(x-x'')\right]\left[-\mathrm{i}\hbar\frac{\partial}{\partial x''}\delta(x''-x')\right] \\
&\quad + V(x)\delta(x-x') \\
&= \frac{-\hbar^2}{2m}\frac{\partial}{\partial x}\int_{-\infty}^{\infty} \mathrm{d}x''\,\delta(x-x'')\frac{\partial}{\partial x''}\delta(x''-x') + V(x)\delta(x-x') \\
&= \frac{-\hbar^2}{2m}\frac{\partial}{\partial x}\frac{\partial}{\partial x}\delta(x-x') + V(x)\delta(x-x') \\
&= \left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x)\right]\delta(x-x')\,. \qquad (5.16)
\end{aligned}$$

It is easy to show that

$$[H,X] = -\frac{\mathrm{i}\hbar}{m}P\,. \qquad (5.17)$$

Furthermore, one shows in Exercise (5.4) that

$$[H,P] = \mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}X}V(X)\,. \qquad (5.18)$$

Here it is understood that the derivative applies to the functional form of V.

5.3 State Space

As a first approximation, the state space of a quantum particle in one dimension is generated by basis vectors $|x\rangle$ or, equivalently, by basis vectors $|p\rangle$. This means that an arbitrary vector $|\psi\rangle$ can be written as

$$|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\,|x\rangle\langle x|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\,\psi(x)|x\rangle\,, \qquad (5.19)$$

$$|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}p\,|p\rangle\langle p|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}p\,\tilde{\psi}(p)|p\rangle\,. \qquad (5.20)$$

Here we have introduced notations $\psi(x) \equiv \langle x|\psi\rangle$ and $\tilde{\psi}(p) \equiv \langle p|\psi\rangle$. The latter is motivated by the fact that

$$\tilde{\psi}(p) = \langle p|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\,\langle p|x\rangle\langle x|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} \mathrm{d}x\,\mathrm{e}^{-\mathrm{i}px/\hbar}\psi(x)\,. \qquad (5.21)$$

Thus $\tilde{\psi}(p)$ is the Fourier transform of $\psi(x)$. We say that $\psi(x)$ gives the state vector in the *coordinate representation*, whereas $\tilde{\psi}(p)$ gives it in the *momentum representation*.

A number of matrix elements will be used repeatedly in this and later chapters. One has

$$\begin{aligned}\langle x|V(X)|\psi\rangle &= \int_{-\infty}^{\infty} \mathrm{d}x' \langle x|V(X)|x'\rangle\langle x'|\psi\rangle \\ &= \int_{-\infty}^{\infty} \mathrm{d}x'\, V(x)\delta(x-x')\psi(x') \\ &= V(x)\psi(x)\,. \end{aligned} \tag{5.22}$$

Similarly, one finds that

$$\langle x|P|\psi\rangle = -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}\psi(x)\,, \tag{5.23}$$

$$\langle x|H|\psi\rangle = \left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)\right]\psi(x)\,. \tag{5.24}$$

Finally,

$$\langle\phi|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\,\langle\phi|x\rangle\langle x|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\,\phi^*(x)\psi(x)\,, \tag{5.25}$$

$$\langle\phi|V(X)|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\,\phi^*(x)V(x)\psi(x)\,, \tag{5.26}$$

$$\langle\phi|P|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\,(-\mathrm{i}\hbar)\phi^*(x)\frac{\mathrm{d}}{\mathrm{d}x}\psi(x)\,, \tag{5.27}$$

$$\langle\phi|H|\psi\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\,\phi^*(x)\left[-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + V(x)\right]\psi(x)\,. \tag{5.28}$$

One can define two operators P_{coor} and H_{coor}, which act in the space of functions $\psi(x)$:

$$P_{\text{coor}} \equiv -\mathrm{i}\hbar\frac{\partial}{\partial x}, \qquad H_{\text{coor}} \equiv -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x). \tag{5.29}$$

Equations (5.23) and (5.24) can then be written as

$$\langle x|P|\psi\rangle = P_{\text{coor}}\psi(x), \qquad \langle x|H|\psi\rangle = H_{\text{coor}}\psi(x). \tag{5.30}$$

P_{coor} and H_{coor} are called the *momentum* and the *Hamiltonian* in the *coordinate representation.* The index "coor" is usually suppressed. The context should indicate whether one deals with P and H or with P_{coor} and H_{coor}.

We now examine more closely the question of the hermiticity of operators X, P and H. We know that an operator O is Hermitian if for all vectors $|\phi\rangle$ and $|\psi\rangle$, $\langle\phi|O|\psi\rangle = \langle\psi|O|\phi\rangle^*$. Let us write this condition in terms of functions $\phi(x)$ and $\psi(x)$.[6] For X we have

---

[6] One can also write it in terms of functions $\tilde{\phi}(p)$ and $\tilde{\psi}(p)$, in which case the following discussion must be adapted. The hermiticity of an operator in general depends on the space in which it acts.

$$\begin{aligned}&\langle\phi|X|\psi\rangle-\langle\psi|X|\phi\rangle^*\\&\quad=\int_{-\infty}^{\infty}\mathrm{d}x\,\phi^*(x)x\psi(x)-\left\{\int_{-\infty}^{\infty}\mathrm{d}x\,\psi^*(x)x\phi(x)\right\}^*=0\,,\end{aligned}$$

so that X is Hermitian. For P the hermiticity condition is

$$\begin{aligned}0&=\langle\phi|P|\psi\rangle-\langle\psi|P|\phi\rangle^*\\&=\int_{-\infty}^{\infty}\mathrm{d}x\,\phi^*(x)(-\mathrm{i}\hbar)\frac{\mathrm{d}}{\mathrm{d}x}\psi(x)-\left\{\int_{-\infty}^{\infty}\mathrm{d}x\,\psi^*(x)(-\mathrm{i}\hbar)\frac{\mathrm{d}}{\mathrm{d}x}\phi(x)\right\}^*\\&=-\mathrm{i}\hbar\int_{-\infty}^{\infty}\mathrm{d}x\,\frac{\mathrm{d}}{\mathrm{d}x}\left\{\phi^*(x)\psi(x)\right\},\end{aligned}\tag{5.31}$$

while for H,

$$\begin{aligned}0&=\langle\phi|H|\psi\rangle-\langle\psi|H|\phi\rangle^*\\&=-\frac{\hbar^2}{2m}\int_{-\infty}^{\infty}\mathrm{d}x\,\frac{\mathrm{d}}{\mathrm{d}x}\left\{\phi^*(x)\frac{\mathrm{d}}{\mathrm{d}x}\psi(x)-\psi(x)\frac{\mathrm{d}}{\mathrm{d}x}\phi^*(x)\right\}.\end{aligned}\tag{5.32}$$

To interpret these equations, we first restrict integrals to an interval going from $-L$ to $+L$. We will eventually take the limit $L\longrightarrow\infty$. For L finite, (5.31) and (5.32) imply that

$$\phi^*(L)\psi(L)=\phi^*(-L)\psi(-L)\,,\tag{5.33}$$

$$\begin{aligned}&\phi^*(L)\left(\frac{\mathrm{d}\psi}{\mathrm{d}x}\right)_L-\psi(L)\left(\frac{\mathrm{d}\phi^*}{\mathrm{d}x}\right)_L\\&\quad=\phi^*(-L)\left(\frac{\mathrm{d}\psi}{\mathrm{d}x}\right)_{-L}-\psi(-L)\left(\frac{\mathrm{d}\phi^*}{\mathrm{d}x}\right)_{-L}.\end{aligned}\tag{5.34}$$

Equations (5.33) and (5.34) are satisfied if the functions $\phi(x)$ and $\psi(x)$ obey the following boundary conditions:[7]

$$\phi(L)=\phi(-L)\,,\qquad\left(\frac{\mathrm{d}\phi}{\mathrm{d}x}\right)_L=\left(\frac{\mathrm{d}\phi}{\mathrm{d}x}\right)_{-L};\tag{5.35}$$

$$\psi(L)=\psi(-L)\,,\qquad\left(\frac{\mathrm{d}\psi}{\mathrm{d}x}\right)_L=\left(\frac{\mathrm{d}\psi}{\mathrm{d}x}\right)_{-L}.\tag{5.36}$$

What are the consequences of these conditions on the nature of eigenvectors of P and H? Let $|\psi_p\rangle=|p\rangle$ be an eigenvector of P with eigenvalue p. We know that

---

[7] One should note that (5.35) and (5.36) are sufficient (not necessary) conditions for the validity of (5.33) and (5.34). Other sufficient conditions are given by the equations $\phi(L)=0=\phi(-L)$, $\psi(L)=0=\psi(-L)$. Their use would not change the main conclusions of the forthcoming discussion.

$$\psi_p(x) = \langle x|p\rangle = c\exp\left\{\frac{\mathrm{i}px}{\hbar}\right\},$$

where c is a constant. It is easy to see that (5.35) and (5.36) entail that $p = n\pi\hbar/L$, where n is an integer. This means that for L finite, p only takes discrete values. Values of p, however, merge into a continuum as $L \to \infty$.

Let us now turn to eigenvectors of the Hamiltonian. Let $|\psi_E\rangle$ be an eigenvector of H with eigenvalue E. One has

$$\langle x|H|\psi_E\rangle = E\langle x|\psi_E\rangle, \tag{5.37}$$

which implies the following differential equation for $\psi_E(x)$:

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi_E(x)+V(x)\psi_E(x) = E\psi_E(x). \tag{5.38}$$

Assume that, as x goes toward $\pm\infty$, $V(x)$ goes toward a finite value V_∞. Let $\psi_E(x)$ be a continuous solution of (5.38) that satisfies conditions (5.35) and (5.36). We will examine cases where $E < V_\infty$ and where $E > V_\infty$.

i) $E < V_\infty$.
Here (5.38) and the boundary conditions can be satisfied for certain values of E only, denoted by $E_i(L)$. One can show that as $L \to \infty$, $E_i(L)$ goes toward a well-defined limit, say E_i. The interval between consecutive values of E_i is finite. One always has $E_i > V_{\min}$, where $V_{\min}$ is the minimum of the function $V(x)$. One also shows that $\psi_{E_i}(x)$ and its derivative go to zero as $|x| \to \infty$ and that[8]

$$\langle\psi_{E_i}|\psi_{E_i}\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\, |\psi_{E_i}(x)|^2 < \infty. \tag{5.39}$$

ii) $E > V_\infty$.
Here again, for L finite, (5.38) and the boundary conditions can be satisfied for certain values of E only. The interval between consecutive values of E is, however, of order L^{-1}, so that the spectrum of E becomes continuous as $L \to \infty$. Functions $\psi_E(x)$ do not vanish as $|x| \to \infty$. They are oscillating and $\langle\psi_E|\psi_E\rangle = \infty$.

Thus as L becomes infinite, eigenvalues of H divide into two groups: a set of discrete values, denoted by E_i, which all obey the inequality $E_i < V_\infty$; and a set of continuous values, which we can denote by E_α, that all obey the inequality $E_\alpha > V_\infty$. Corresponding functions (called *eigenfunctions* of H) can be denoted by $\psi_i(x)$ and $\psi_\alpha(x)$, and corresponding kets by $|E_i\rangle$ and $|E_\alpha\rangle$. Eigenvectors associated with distinct eigenvalues are all orthogonal. It is very convenient to use the following normalization:

---

[8] All these statements, like those in the following paragraph, will be illustrated in Chap. 8.

$$\langle E_i|E_j\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\, \psi_i^*(x)\psi_j(x) = \delta_{ij}\,, \tag{5.40}$$

$$\langle E_\alpha|E_\beta\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\, \psi_\alpha^*(x)\psi_\beta(x) = \delta(\alpha-\beta)\,. \tag{5.41}$$

On the basis of this analysis, we can now define more precisely the state space of a quantum particle in one dimension. The state space is made up of all vectors $|\psi\rangle$ that obey the following conditions:

i) $|\psi\rangle$ can be expanded as a linear combination of kets $|E_i\rangle$ and $|E_\alpha\rangle$, that is,

$$|\psi\rangle = \sum_i C_i|E_i\rangle + \int \mathrm{d}\alpha\, C_\alpha|E_\alpha\rangle\,.$$

ii) The norm of $|\psi\rangle$ is finite, that is,

$$\langle\psi|\psi\rangle = \sum_i |C_i|^2 + \int \mathrm{d}\alpha\, |C_\alpha|^2 < \infty\,.$$

iii) $\psi(x) = \langle x|\psi\rangle$ is everywhere a finite and continuous function.

In (i) and (ii) the sum and integral are carried over all values of i and α.

In the state space we have just defined, the identity operator can be written as

$$\sum_i |E_i\rangle\langle E_i| + \int \mathrm{d}\alpha\, |E_\alpha\rangle\langle E_\alpha| = I\,. \tag{5.42}$$

The state vector $|\psi\rangle$ can always be normalized, so that $\langle\psi|\psi\rangle = 1$. In this case, according to interpretation rule 5 on p. 35, $|\langle x|\psi\rangle|^2$ represents the probability density that a position measurement in state $|\psi\rangle$ yields value x. This means that $|\psi(x)|^2\mathrm{d}x$ is equal to the probability that the result of a position measurement falls between x and $x + \mathrm{d}x$. Likewise $|\langle p|\psi\rangle|^2\mathrm{d}p = |\tilde{\psi}(p)|^2\mathrm{d}p$ is equal to the probability that a momentum measurement yields a result between p and $p + \mathrm{d}p$. In the case where energy is nondegenerate, $|\langle E_i|\psi\rangle|^2 = |C_i|^2$ is equal to the probability that an energy measurement yields value E_i, and $|\langle E_\alpha|\psi\rangle|^2\mathrm{d}\alpha = |C_\alpha|^2\mathrm{d}\alpha$ is equal to the probability that the result of an energy measurement falls between E_α and $E_{\alpha+\mathrm{d}\alpha}$. If energy is degenerate, these expressions should be modified appropriately.

Strictly speaking vectors $|x\rangle$, $|p\rangle$ and $|E_\alpha\rangle$ do not belong to the state space, because their norm is infinite. This means that a vector with perfectly well-defined p, for instance, does not belong to the state space. It is nonetheless always possible to build finite-norm vectors representing states wherein momentum is arbitrarily well-defined. Thus

$$|\psi_{p_0}\rangle = \int_{-\infty}^{\infty} \mathrm{d}p\, c_p|p\rangle\,, \tag{5.43}$$

where $c_p = 0$ if $|p - p_0| > \varepsilon$, represents a state wherein momentum is necessarily between $p_0 - \varepsilon$ and $p_0 + \varepsilon$. This state is normalized if $\int_{-\infty}^{\infty} \mathrm{d}p \, |c_p|^2 = 1$.

In spite of that remark, one frequently treats vectors $|x\rangle$, $|p\rangle$ and $|E_\alpha\rangle$ as if they were part of the state space. Great simplification often ensues in computations. Any ambiguity should be resolved by going back to normalized vectors. They can be vectors like $|\psi_{p_0}\rangle$, or similar ones associated with a well-defined energy. The problem can also be treated by restricting values of x to an interval $(-L, L)$, in which case vectors $|p\rangle$ and $|E\rangle$ all have finite norm. This method is called *box normalization.* The limit $L \longrightarrow \infty$ is taken at the end of calculations.

We point out that if $V_\infty = \infty$, the Hamiltonian has a discrete spectrum only. The case where $V(\infty) \neq V(-\infty)$ can be treated by a straightforward generalization of the foregoing analysis.

5.4 The Schrödinger Equation

At any time t, the state of a quantum particle can be represented by a vector $|\psi(t)\rangle$ in the state space. The time evolution of $|\psi(t)\rangle$ satisfies the *Schrödinger equation*:

$$\mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi(t)\rangle = H(t) |\psi(t)\rangle . \tag{5.44}$$

Expanded in the basis of kets $|x\rangle$, the state vector $|\psi(t)\rangle$ of a particle in one dimension becomes a function $\langle x|\psi(t)\rangle = \psi(x, t)$, called the *wave function.* Projecting the Schrödinger equation on $\langle x|$ and making use of (5.24), one easily finds that the wave function satisfies the following equation:

$$\mathrm{i}\hbar \frac{\partial}{\partial t} \psi(x, t) = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x, t) + V(x, t) \psi(x, t) . \tag{5.45}$$

Here we have to deal with a second-order linear differential equation. The time derivative, however, shows up in first order only. One can show that this equation determines $\psi(x, t)$ at any time t if $\psi(x, t_0)$ is known at a given time t_0.

It is easy to show that if the Hamiltonian does not depend on time, the Schrödinger equation has particular solutions of the form

$$|\psi(t)\rangle = \mathrm{e}^{-\mathrm{i}Et/\hbar} |E\rangle , \tag{5.46}$$

where E is an eigenvalue of H, belonging to the discrete or the continuous spectrum. Equation (5.44) is linear. This means that any linear combination of solutions is itself a solution. For all values of the constants C_i and C_α, therefore, the following expression is a solution of the Schrödinger equation:

$$\begin{aligned}|\psi(t)\rangle &= \sum_i C_i \mathrm{e}^{-\mathrm{i}E_i t/\hbar}|E_i\rangle + \int \mathrm{d}\alpha\, C_\alpha \mathrm{e}^{-\mathrm{i}E_\alpha t/\hbar}|E_\alpha\rangle \\ &= \mathsf{S}_E C_E \mathrm{e}^{-\mathrm{i}Et/\hbar}|E\rangle\,. \end{aligned} \tag{5.47}$$

We have introduced a condensed notation S to cover at once a summation on the discrete and an integral on the continuous spectrum. The summation and the integral are carried on the set of eigenvectors of H.

At a given time t_0 let us define

$$C_E = \mathrm{e}^{\mathrm{i}Et_0/\hbar}\langle E|\psi(t_0)\rangle\,, \tag{5.48}$$

and let us substitute this in (5.47). We get

$$|\psi(t)\rangle = \mathsf{S}_E \langle E|\psi(t_0)\rangle \mathrm{e}^{-\mathrm{i}E(t-t_0)/\hbar}|E\rangle\,. \tag{5.49}$$

We have thus obtained a solution of Schrödinger's equation that satisfies an arbitrary initial condition. Hence it is the general solution. In terms of the wave function it can be written as

$$\langle x|\psi(t)\rangle = \psi(x,t) = \mathsf{S}_E \langle E|\psi(t_0)\rangle \mathrm{e}^{-\mathrm{i}E(t-t_0)/\hbar}\psi_E(x)\,. \tag{5.50}$$

Once again, the determination of the time evolution of $|\psi(t)\rangle$ is equivalent to the complete solution of the eigenvalue equation of H. This, in the present case, is (5.38), together with the boundary conditions developed in Sect. 5.3.

We can rewrite (5.49) as

$$|\psi(t)\rangle = \left\{\mathsf{S}_E \mathrm{e}^{-\mathrm{i}E(t-t_0)/\hbar}|E\rangle\langle E|\right\}|\psi(t_0)\rangle\,. \tag{5.51}$$

This means that the evolution operator, defined in (3.13), is given for a time-independent Hamiltonian by

$$U(t,t_0) = \mathsf{S}_E \mathrm{e}^{-\mathrm{i}E(t-t_0)/\hbar}|E\rangle\langle E|\,. \tag{5.52}$$

The *propagator* $U(x,t;x_0,t_0)$ is defined in terms of matrix elements of the evolution operator in the basis of kets $|x\rangle$:

$$U(x,t;x_0,t_0) = \langle x|U(t,t_0)|x_0\rangle\,. \tag{5.53}$$

By means of the propagator one can write the wave function at time t in terms of the wave function at time t_0 as

$$\psi(x,t) = \int_{-\infty}^{\infty} \mathrm{d}x_0\, U(x,t;x_0,t_0)\psi(x_0,t_0)\,. \tag{5.54}$$

5.5 Square Well Potential

The simplest of all potentials is the one for a free particle, in which case $V(x) = 0$. The Hamiltonian is then given by

$$H = \frac{1}{2m} P^2 , \tag{5.55}$$

and eigenvectors of H coincide with eigenvectors of P. Thus

$$\langle x | E_p \rangle = \langle x | p \rangle = \frac{1}{\sqrt{2\pi\hbar}} \mathrm{e}^{\mathrm{i}px/\hbar} , \tag{5.56}$$

$$E_p = \frac{1}{2m} p^2 . \tag{5.57}$$

Each value of energy is twofold degenerate (except for the null value), and energy is never negative. We recall that the Hamiltonian cannot have eigenvalues lower than the potential's minimum. Indeed it is then impossible to satisfy the boundary conditions.

The eigenvalue equation for H is written in general as

$$-\frac{\hbar^2}{2m} \frac{\mathrm{d}^2 \psi}{\mathrm{d}x^2} + V(x)\psi(x) = E\psi(x) . \tag{5.58}$$

In this section and the following we will consider sectionally constant potentials. It is very useful to write down the general solution of (5.58) in a region of space where V is constant. Let us first define two positive constants k and κ as

$$k^2 = \frac{2m}{\hbar^2}(E - V) \quad \text{if} \quad E > V , \tag{5.59}$$

$$\kappa^2 = \frac{2m}{\hbar^2}(V - E) \quad \text{if} \quad E < V . \tag{5.60}$$

It is easy to check that in a region where V does not depend on x, the general solution of (5.58) can be written as

$$\psi(x) = \begin{cases} A \cos kx + B \sin kx & \text{if } E > V , \\ A\mathrm{e}^{\kappa x} + B\mathrm{e}^{-\kappa x} & \text{if } E < V . \end{cases} \tag{5.61}$$

Here A and B are two arbitrary complex constants.

If $V(x)$ is sectionally constant, the function $\psi(x)$ can be written as in (5.61) in each region where V is constant. Coefficients A and B differ from one region to another. Their values are determined so that at a point where $V(x)$ is discontinuous, $\psi(x)$ is as smooth as it can be. To see what this means, suppose that $V(x)$ only has finite discontinuities. Equation (5.58) then implies that the second derivative of ψ has finite discontinuities. This means that the first derivative of ψ, and ψ itself, are continuous. Hence in the case where

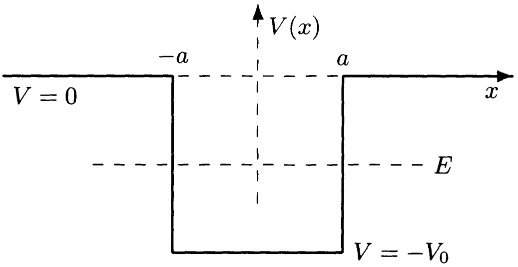

Fig. 5.1. Square well potential

$V(x)$ only has finite discontinuities, the coefficients A and B must be chosen so that ψ and its first derivative are continuous.

To illustrate this let us consider a square well potential, of width $2a$ and depth V_0, shown in Fig. 5.1. We have

$$V(x)=\begin{cases}0 & \text{if } x<-a\,,\\ -V_0 & \text{if } -a<x<a\,,\\ 0 & \text{if } x>a\,.\end{cases} \tag{5.62}$$

Here $V(\pm\infty)=0$. According to the argument in Sect. 5.3, eigenvalues E of the Hamiltonian divide into two groups: a discrete spectrum if $-V_0<E<0$ and a continuous spectrum if $E>0$. Let us focus on the discrete spectrum. In this case, $\psi(x)$ and its derivative must vanish as $|x|\to\infty$. Hence the problem can be formulated as follows: find values of E, between $-V_0$ and 0, for which there is a function $\psi(x)$ that satisfies (5.58), that, together with its derivative, is everywhere continuous and that vanishes as $|x|\to\infty$.

In the three regions where V is constant, the general solution of (5.58) can be written as

$$\psi(x)=\begin{cases}A_-\mathrm{e}^{\kappa x}+B_-\mathrm{e}^{-\kappa x} & \text{if } x<-a\,,\\ A\cos kx+B\sin kx & \text{if } -a<x<a\,,\\ A_+\mathrm{e}^{\kappa x}+B_+\mathrm{e}^{-\kappa x} & \text{if } x>a\,.\end{cases}$$

Here, E is negative and

$$\kappa^2=-\frac{2mE}{\hbar^2}\,,\qquad k^2=\frac{2m}{\hbar^2}(E+V_0)\,.$$

For $\psi(x)$ to vanish as $|x|\to\infty$, it is necessary that $A_+=0=B_-$. Let us now enforce the conditions of continuity of $\psi(x)$ and its derivative at the points $x=a$ and $x=-a$. At $x=a$ we must have

$$\lim_{x\to a^+}\psi(x)=B_+\mathrm{e}^{-\kappa a}=\lim_{x\to a^-}\psi(x)=A\cos ka+B\sin ka\,,$$

$$\lim_{x\to a^+}\frac{\mathrm{d}\psi}{\mathrm{d}x}=-\kappa B_+\mathrm{e}^{-\kappa a}=\lim_{x\to a^-}\frac{\mathrm{d}\psi}{\mathrm{d}x}=k(-A\sin ka+B\cos ka)\,.$$

These two equations entail that

$$\begin{aligned} B_+ &= \mathrm{e}^{\kappa a}(A\cos ka + B\sin ka)\,, \\ \kappa &= k\left\{\frac{A\sin ka - B\cos ka}{A\cos ka + B\sin ka}\right\}. \end{aligned}$$

Similarly, the continuity conditions at $x = -a$ imply that

$$\begin{aligned} A_- &= \mathrm{e}^{\kappa a}(A\cos ka - B\sin ka)\,, \\ \kappa &= k\left\{\frac{A\sin ka + B\cos ka}{A\cos ka - B\sin ka}\right\}. \end{aligned}$$

It is not difficult to see that the two equations for κ are consistent only if $A = 0$ or $B = 0$. These two cases give rise to two types of solutions for $\psi(x)$.

i) $B = 0$.

$$\psi(x) = \begin{cases} A(\cos ka)\mathrm{e}^{\kappa a}\mathrm{e}^{\kappa x} & \text{if } x < -a\,, \\ A\cos kx & \text{if } -a < x < a\,, \\ A(\cos ka)\mathrm{e}^{\kappa a}\mathrm{e}^{-\kappa x} & \text{if } x > a\,. \end{cases} \tag{5.63}$$

In the present case, $\psi(x)$ is an even function and k and κ satisfy the following constraint:

$$\kappa = k\tan ka\,. \tag{5.64}$$

ii) $A = 0$.

$$\psi(x) = \begin{cases} -B(\sin ka)\mathrm{e}^{\kappa a}\mathrm{e}^{\kappa x} & \text{if } x < -a\,, \\ B\sin kx & \text{if } -a < x < a\,, \\ B(\sin ka)\mathrm{e}^{\kappa a}\mathrm{e}^{-\kappa x} & \text{if } x > a\,. \end{cases} \tag{5.65}$$

Here $\psi(x)$ is an odd function and k and κ satisfy the constraint

$$\kappa = -k\cot ka\,. \tag{5.66}$$

The two constraints on k and κ can be rewritten as

$$\text{i)}\quad \left(\frac{-E}{E+V_0}\right)^{1/2} = \tan\left[\left(\frac{2m(E+V_0)a^2}{\hbar^2}\right)^{1/2}\right], \tag{5.67}$$

$$\text{ii)}\quad \left(\frac{-E}{E+V_0}\right)^{1/2} = -\cot\left[\left(\frac{2m(E+V_0)a^2}{\hbar^2}\right)^{1/2}\right]. \tag{5.68}$$

It is clear that these are in fact constraints on E. The upshot is that discrete eigenvalues of H, for the square well potential, are the solutions of (5.67) and (5.68).

These equations can be solved graphically or numerically. For V_0 and a finite, one always finds a finite number of solutions. The larger the parameter V_0a^2 is, the more solutions there are. The lowest energy always corresponds to an even function, the next energy to an odd function, and so on alternately. No eigenvalue in the discrete spectrum is degenerate.

It is important to notice that the continuity conditions together with boundary conditions for large $|x|$ have completely determined eigenvalues E and have determined eigenfunctions $\psi(x)$ up to a constant multiplicative factor. This factor can be chosen so that $\psi(x)$ is normalized.

We point out that some of our results are more general than the context in which they have been derived. Thus the discrete energy spectrum of a spinless particle in one dimension (with no internal degrees of freedom) is never degenerate. This, however, no longer holds in two or more dimensions. Furthermore, for any potential $V(x)$ that is an even function, eigenfunctions of H in the discrete spectrum are even or odd. The lowest energy always corresponds to an even function, the next lowest to an odd function, and so on.

Eigenfunctions of H in the continuous spectrum of the square well potential can be obtained easily. Let us just point out that each positive value of energy is twofold degenerate.

The limiting case of an infinitely deep potential well is interesting. It is convenient to set the zero of potential at the bottom of the well, so that

$$V(x) = \begin{cases} \infty & \text{if } x < -a\,, \\ 0 & \text{if } -a < x < a\,, \\ \infty & \text{if } x > a\,. \end{cases} \tag{5.69}$$

Outside the well the constant κ is infinite, hence $\psi(x)$ can only vanish. Eigenfunctions inside the well are therefore given by the following expressions:

i) Even functions.

$$\psi(x) = A\cos kx \quad \text{if} \quad -a < x < a, \tag{5.70}$$

with the constraint $\cos ka = 0$, that is, $ka = (2n+1)\pi/2$ where n is an integer.

ii) Odd functions.

$$\psi(x) = B\sin kx \quad \text{if} \quad -a < x < a, \tag{5.71}$$

with the constraint $\sin ka = 0$, that is, $ka = n\pi$ where n is an integer.

In both cases we have

$$E = \frac{\hbar^2k^2}{2m} = \frac{\hbar^2\pi^2}{8ma^2}n^2\,. \tag{5.72}$$

Only positive values of n should be kept, as negative values reproduce the same eigenfunctions up to a sign. In (5.72), even and odd values of n correspond to odd and even functions, respectively.

For the infinitely deep well, the derivative of $\psi(x)$ is discontinuous at $x = \pm a$. The function ψ, on the other hand, is everywhere continuous. This corresponds to the fact that the second derivative of ψ has an infinite discontinuity at $x = \pm a$.

5.6 Square Potential Barrier

We just obtained the discrete energy spectrum associated with a square well potential. To highlight the importance of the continuous spectrum, we consider a square potential barrier, illustrated in Fig. 5.2. We have

$$V(x) = \begin{cases} 0 & \text{if } x < -a\,, \\ V_0 & \text{if } -a < x < a\,, \\ 0 & \text{if } x > a\,. \end{cases} \tag{5.73}$$

Since the potential minimum coincides with V_∞, one immediately concludes that the Hamiltonian has a continuous spectrum only. We will focus on values of E between 0 and V_0.

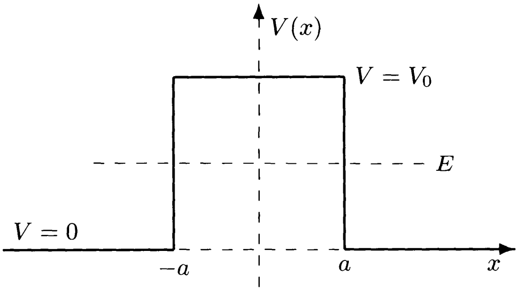

Fig. 5.2. Square potential barrier

Here we have

$$k = \left\{\frac{2mE}{\hbar^2}\right\}^{1/2}, \qquad \kappa = \left\{\frac{2m}{\hbar^2}(V_0 - E)\right\}^{1/2}. \tag{5.74}$$

We look for eigenfunctions of H of the following form:

$$\psi(x) = \begin{cases} A_- \mathrm{e}^{\mathrm{i}kx} + B_- \mathrm{e}^{-\mathrm{i}kx} & \text{if } x < -a\,, \\ A\mathrm{e}^{\kappa x} + B\mathrm{e}^{-\kappa x} & \text{if } -a < x < a\,, \\ A_+ \mathrm{e}^{\mathrm{i}kx} & \text{if } x > a\,. \end{cases} \tag{5.75}$$

It is clear that the form of $\psi(x)$ in regions where $x > a$ and $x < -a$ is equivalent to a sum of terms $\sin kx$ and $\cos kx$.

We will show in Sect. 7.2 that wave function (5.75) represents a flux of incoming particles and a flux of reflected particles in the region where $x < -a$, and a flux of transmitted particles in the region where $x > a$. The number of incoming particles per unit area and time is proportional to $|A_-|^2$. The number of reflected particles and the number of transmitted particles are proportional to $|B_-|^2$ and $|A_+|^2$ respectively.

Enforcing the continuity conditions at $x = -a$ and $x = a$ allows one to express coefficients A_- and B_- in terms of A and B, and these in terms of A_+. A relatively simple calculation shows that (Exercise 5.10)

$$A_- = A_+ \mathrm{e}^{2\mathrm{i}ka} \frac{1}{4\mathrm{i}k\kappa} \left\{ (\kappa + \mathrm{i}k)^2 \mathrm{e}^{-2\kappa a} - (\kappa - \mathrm{i}k)^2 \mathrm{e}^{2\kappa a} \right\}, \tag{5.76}$$

$$B_- = A_+ \frac{\kappa^2 + k^2}{4\mathrm{i}k\kappa} \left\{ -\mathrm{e}^{-2\kappa a} + \mathrm{e}^{2\kappa a} \right\}. \tag{5.77}$$

In the case where $\kappa a \gg 1$ we find that

$$\left| \frac{B_-}{A_-} \right|^2 \approx 1, \qquad \left| \frac{A_+}{A_-} \right|^2 \approx \left[\frac{4k\kappa \mathrm{e}^{-2\kappa a}}{k^2 + \kappa^2} \right]^2.$$

The first of these relations means that if $\kappa a \gg 1$, the number of reflected particles is approximately equal to the number of incoming particles. The second relation means that the ratio of number of transmitted particles to number of incoming particles is very small, of order $\mathrm{e}^{-4\kappa a}$. Nevertheless, it is always possible for a quantum particle to reach beyond a potential barrier higher than its energy. The probability, however, gets smaller and smaller as a and $V_0 - E$ become large.

The possibility of crossing a barrier higher than its energy is a hallmark of the quantum, as opposed to the classical, particle. It gives rise to many important phenomena, for instance α decay and the Josephson effect. We will come back to barriers in Sect. 16.7.

The methods we have used, in this and the last section, to obtain eigenvalues and eigenfunctions of the Hamiltonian can be applied to any sectionally constant potential. Such problems always admit, in principle, a complete and exact solution. Needless to say, the computations become more and more cumbersome as the number of different values of $V(x)$ increases.

5.7 The Harmonic Oscillator

A *harmonic oscillator* is a quantum system governed by the Hamiltonian

$$H = \frac{1}{2m} P^2 + \frac{1}{2} m\omega^2 X^2, \tag{5.78}$$

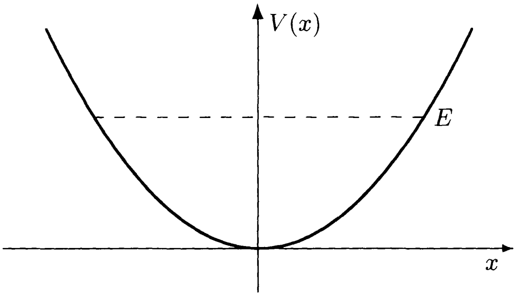

Fig. 5.3. Harmonic oscillator potential

where m is the particle's mass and ω is a constant, called *angular frequency*, that has dimensions of s^{-1}. The potential is illustrated in Fig. 5.3. We see that $V(\infty) = \infty$. Hence the Hamiltonian has a discrete spectrum only and all eigenvectors of H have finite norm.

Eigenvalues and eigenvectors of H can be found exactly, in several different ways. The most popular method consists in looking for values of E for which the differential equation

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + \frac{1}{2}m\omega^2 x^2\psi(x) = E\psi(x)$$

has solutions that go to zero as $|x| \to \infty$. That approach is developed in Exercise (5.17).

Instead of solving the differential equation directly, we will use here an algebraic method. In this way a type of operators will be introduced that have many applications in quantum mechanics as well as in quantum field theory.

We introduce a dimensionless operator A and its adjoint $A^\dagger$ defined as

$$A = \frac{\mathrm{i}}{(2m\hbar\omega)^{1/2}}(P - \mathrm{i}\omega m X)\,, \tag{5.79}$$

$$A^\dagger = \frac{-\mathrm{i}}{(2m\hbar\omega)^{1/2}}(P + \mathrm{i}\omega m X)\,. \tag{5.80}$$

One easily checks that the commutation relation (5.14) for X and P translates into a commutation relation for A and $A^\dagger$ given by

$$[A, A^\dagger] = I\,. \tag{5.81}$$

Operators X and P can be expressed in terms of A and $A^\dagger$. We have

$$X = \left(\frac{\hbar}{2\omega m}\right)^{1/2}(A + A^\dagger)\,, \tag{5.82}$$

$$P = -\mathrm{i}\left(\frac{m\hbar\omega}{2}\right)^{1/2}(A - A^\dagger)\,. \tag{5.83}$$

Substituting these relations in the expression for the Hamiltonian, one finds that

$$H = \hbar\omega \left\{ A^\dagger A + \frac{1}{2} I \right\} . \tag{5.84}$$

From (5.84), one easily obtains the commutation relations of H with A and $A^\dagger$:

$$[H, A] = -\hbar\omega A \,, \tag{5.85}$$

$$\left[H, A^\dagger\right] = \hbar\omega A^\dagger \,. \tag{5.86}$$

Let us denote by $|E_i\rangle$ a normalized eigenvector of H with eigenvalue E_i. We will not take it for granted here that eigenvalues of H are not degenerate. Let us consider the vector $A|E_i\rangle$ and apply to it the operator H. We see that

$$\begin{aligned} H\left\{A|E_i\rangle\right\} &= AH|E_i\rangle + [H, A]|E_i\rangle \\ &= AE_i|E_i\rangle - \hbar\omega A|E_i\rangle \\ &= (E_i - \hbar\omega)\left\{A|E_i\rangle\right\} . \end{aligned} \tag{5.87}$$

Thus $A|E_i\rangle$, if it does not vanish, is an eigenvector of H with eigenvalue $E_i - \hbar\omega$. The action of A on an eigenvector of H yields another eigenvector whose energy is reduced by $\hbar\omega$. For this reason A is called an *annihilation operator*. $A^\dagger$, on the other hand, is called a *creation operator*. One easily shows that the action of $A^\dagger$ on $|E_i\rangle$ yields an eigenvector of H with eigenvalue $E_i + \hbar\omega$.

The vector $A^l|E_i\rangle$ is the result of applying l times the operator A on $|E_i\rangle$. If it does not vanish, it is an eigenvector of H with eigenvalue $E_i - l\hbar\omega$. But $A^l|E_i\rangle$ will necessarily vanish for l large enough. This follows from the existence of a lower bound on eigenvalues of H. Indeed for any normalized eigenvector $|E_i\rangle$,

$$\langle E_i| \left\{ H - \frac{\hbar\omega}{2} I \right\} |E_i\rangle = E_i - \frac{1}{2}\hbar\omega = \hbar\omega\langle E_i|A^\dagger A|E_i\rangle = \hbar\omega\langle\phi|\phi\rangle, \tag{5.88}$$

where $|\phi\rangle = A|E_i\rangle$. The last term in this equation is obviously nonnegative, so that $E_i \geq \hbar\omega/2$. It is clear that from a given l up, $E_i - l\hbar\omega$ will not satisfy this inequality.

We now show that up to a multiplicative factor, the vector $|E_0\rangle$ for which $A|E_0\rangle = 0$ is unique. Denoting $\langle x|E_0\rangle$ by $\psi_0(x)$, we have

$$\begin{aligned} 0 = \langle x|A|E_0\rangle &= \frac{\mathrm{i}}{(2m\hbar\omega)^{1/2}} \langle x|P - \mathrm{i}m\omega X|E_0\rangle \\ &= \frac{\mathrm{i}}{(2m\hbar\omega)^{1/2}} \left\{ -\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x}\psi_0(x) - \mathrm{i}m\omega x\psi_0(x) \right\} . \end{aligned}$$

The general solution of this differential equation for $\psi_0(x)$ is given by

$$\psi_0(x) = c\exp\left\{-\frac{m\omega}{2\hbar}x^2\right\},$$

where c is a constant. This proves the uniqueness of $|E_0\rangle$ and yields the explicit form of $\psi_0(x)$. The multiplicative constant is partly determined by the normalization condition on $\psi_0(x)$, which is satisfied if one lets

$$\psi_0(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} \exp\left\{-\frac{m\omega}{2\hbar}x^2\right\}. \tag{5.89}$$

The energy E_0 associated with the ket $|E_0\rangle$ is given by

$$E_0 = \langle E_0|H|E_0\rangle = \hbar\omega\langle E_0|\left\{A^\dagger A + \frac{1}{2}I\right\}|E_0\rangle = \frac{1}{2}\hbar\omega, \tag{5.90}$$

which is the lowest possible value of the lower bound found earlier. Eigenvalues of H all have the form $\hbar\omega(n+\frac{1}{2})$, because any eigenvector can be brought to $|E_0\rangle$ by n applications of the annihilation operator. No eigenvalue is degenerate. Indeed if $|E_n\rangle$ and $|\bar{E}_n\rangle$ correspond to the same eigenvalue $E_n = \hbar\omega(n+\frac{1}{2})$, there exist constants α and $\bar{\alpha}$ such that

$$A^n|E_n\rangle = \alpha|E_0\rangle, \qquad A^n|\bar{E}_n\rangle = \bar{\alpha}|E_0\rangle.$$

Thus

$$A^n\left\{\bar{\alpha}|E_n\rangle - \alpha|\bar{E}_n\rangle\right\} = 0.$$

But this is possible only if $\bar{\alpha}|E_n\rangle - \alpha|\bar{E}_n\rangle = 0$, that is, if $|E_n\rangle$ and $|\bar{E}_n\rangle$ differ by a multiplicative factor only.

The foregoing discussion entails that all vectors in the state space are obtained through repeated action of the creation operator $A^\dagger$ on $|E_0\rangle$. The proportionality constant between $A|E_n\rangle$ and $|E_{n-1}\rangle$ can be determined by writing $A|E_n\rangle = \alpha_n|E_{n-1}\rangle$ and showing that

$$\begin{aligned}|\alpha_n|^2 &= \langle E_n|A^\dagger A|E_n\rangle = \langle E_n|\left\{\frac{1}{\hbar\omega}H - \frac{1}{2}I\right\}|E_n\rangle \\ &= (\hbar\omega)^{-1}(\hbar\omega)\left(n+\frac{1}{2}\right) - \frac{1}{2} = n.\end{aligned}$$

Hence by an appropriate choice of phase for kets $|E_n\rangle$,

$$A|E_n\rangle = \sqrt{n}|E_{n-1}\rangle. \tag{5.91}$$

Moreover, for any l

$$\langle E_l|A^\dagger|E_n\rangle = \langle E_n|A|E_l\rangle^* = \sqrt{l}\,\delta_{n,l-1} = \sqrt{n+1}\,\delta_{l,n+1},$$

and therefore

$$A^\dagger|E_n\rangle = \sqrt{n+1}|E_{n+1}\rangle\,. \tag{5.92}$$

Matrix elements of operators H, A and $A^\dagger$ in the basis of eigenvectors of H are thus completely determined.

There remains to find the coefficients $\psi_n(x) = \langle x|E_n\rangle$ of the expansion of vectors $|E_n\rangle$ in the $|x\rangle$ basis. Let us first derive a recurrence formula for $\psi_{n+1}(x)$:

$$\begin{aligned}
\psi_{n+1}(x) = \langle x|E_{n+1}\rangle &= \frac{1}{\sqrt{n+1}}\langle x|A^\dagger|E_n\rangle \\
&= \frac{1}{\sqrt{n+1}}\frac{(-\mathrm{i})}{(2m\hbar\omega)^{1/2}}\langle x|P+\mathrm{i}\omega mX|E_n\rangle \\
&= \frac{1}{\sqrt{n+1}}\frac{(-\hbar)}{(2m\hbar\omega)^{1/2}}\left\{\frac{\mathrm{d}}{\mathrm{d}x}\psi_n(x) - \frac{m\omega x}{\hbar}\psi_n(x)\right\} \\
&= -\left(\frac{\hbar}{(n+1)2m\omega}\right)^{1/2}\exp\left\{\frac{m\omega}{2\hbar}x^2\right\}\frac{\mathrm{d}}{\mathrm{d}x}\left[\exp\left\{-\frac{m\omega}{2\hbar}x^2\right\}\psi_n(x)\right].
\end{aligned}$$

By induction one easily checks that

$$\begin{aligned}
\psi_n(x) = (-1)^n\left(\frac{\hbar}{2m\omega}\right)^{n/2}&\frac{1}{\sqrt{n!}}\exp\left\{\frac{m\omega}{2\hbar}x^2\right\} \\
&\times\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left[\exp\left\{-\frac{m\omega}{2\hbar}x^2\right\}\psi_0(x)\right].
\end{aligned} \tag{5.93}$$

Letting

$$\xi = \left(\frac{m\omega}{\hbar}\right)^{1/2} x \tag{5.94}$$

and substituting in (5.93) the expression (5.89) for $\psi_0(x)$, we finally obtain

$$\begin{aligned}
\psi_n(\xi) &= (-1)^n\left[\sqrt{\frac{m\omega}{\pi\hbar}}\frac{1}{2^n n!}\right]^{1/2}\mathrm{e}^{\xi^2/2}\frac{\mathrm{d}^n}{\mathrm{d}\xi^n}e^{-\xi^2} \\
&= \left[\sqrt{\frac{m\omega}{\pi\hbar}}\frac{1}{2^n n!}\right]^{1/2}\mathrm{e}^{-\xi^2/2}H_n(\xi)\,,
\end{aligned} \tag{5.95}$$

where

$$H_n(\xi) = (-1)^n\mathrm{e}^{\xi^2}\frac{\mathrm{d}^n}{\mathrm{d}\xi^n}\mathrm{e}^{-\xi^2}\,. \tag{5.96}$$

$H_n(\xi)$ is a polynomial of degree n in ξ, called the *Hermite polynomial* of order n. Properties of these polynomials are summarized in Sect. 5.9.3.

We have found, exactly and explicitly, the eigenvalues and eigenfunctions of the quantum harmonic oscillator Hamiltonian. Our results, valid in one space dimension, can be generalized easily to the three-dimensional case and even to the case of several coupled harmonic oscillators. In this guise the harmonic oscillator has numerous applications in molecular physics, in solid state physics and in quantum field theory ([53], Chap. 5). We will come back to this in Sects. 10.7 and 20.6.

5.8 Evolution Operator and Wave Packets

In the last three sections we have explicitly obtained, for different potentials, the Hamiltonian's eigenvalues and eigenvectors. When the Hamiltonian does not depend on time, knowledge of eigenvalues and eigenvectors yields, in principle, the time evolution of the quantum system. The time evolution is entirely determined by the evolution operator or, equivalently, by its matrix elements, in other words the propagator.

Let us first compute the propagator of a free particle. We have in general

$$\begin{aligned} U(x,t;x_0,t_0) &= \langle x| \left\{ \mathsf{S}_E \mathrm{e}^{-\mathrm{i}E(t-t_0)/\hbar} |E\rangle\langle E| \right\} |x_0\rangle \\ &= \mathsf{S}_E \mathrm{e}^{-\mathrm{i}E(t-t_0)/\hbar} \psi_E(x)\psi_E^*(x_0) \,. \end{aligned} \tag{5.97}$$

Eigenvalues and eigenvectors of the free-particle Hamiltonian were obtained in (5.56) and (5.57). There is a continuous spectrum only, and eigenfunctions are indexed by p, the eigenvalue of the momentum operator. Thus

$$\begin{aligned} &U_{\mathrm{free}}(x,t;x_0,t_0) \\ &= \int_{-\infty}^{\infty} \mathrm{d}p \exp\left[\frac{-\mathrm{i}(t-t_0)}{\hbar}\frac{p^2}{2m} \right] \frac{1}{\sqrt{2\pi\hbar}} \mathrm{e}^{\mathrm{i}px/\hbar} \frac{1}{\sqrt{2\pi\hbar}} \mathrm{e}^{-\mathrm{i}px_0/\hbar} \\ &= \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} \mathrm{d}p \exp\left[-\left(\frac{\mathrm{i}(t-t_0)}{2m\hbar}\right)\left(p - \frac{m(x-x_0)}{(t-t_0)}\right)^2 \right. \\ &\qquad \left. + \frac{\mathrm{i}m(x-x_0)^2}{2\hbar(t-t_0)} \right] . \end{aligned}$$

By an elementary change of variable we obtain

$$U_{\mathrm{free}}(x,t;x_0,t_0) = \frac{1}{2\pi\hbar} \exp\left[\frac{\mathrm{i}m(x-x_0)^2}{2\hbar(t-t_0)} \right] \int_{-\infty}^{\infty} \mathrm{d}q \exp\left[-\frac{\mathrm{i}(t-t_0)}{2m\hbar} q^2 \right] .$$

For real and positive α we have

$$\int_{-\infty}^{\infty} \mathrm{d}q \exp\left\{-\alpha q^2\right\} = \sqrt{\frac{\pi}{\alpha}} \,. \tag{5.98}$$

In Sect. 16.2 we will show that this result remains true if α is complex, provided $-\pi/2 \le \arg(\alpha) \le \pi/2$. Hence we find

$$U_{\text{free}}(x,t;x_0,t_0) = \left[\frac{m}{2\pi i\hbar(t-t_0)}\right]^{1/2} \exp\left[\frac{im(x-x_0)^2}{2\hbar(t-t_0)}\right]. \tag{5.99}$$

A wave function that is localized in a small region is called a *wave packet.* To some extent, the quantum state represented by a wave packet resembles what one intuitively means by a particle. An example of a wave packet at t_0 is a normalized Gaussian function, that is,

$$\psi(x_0,t_0) = \pi^{-1/4}\Delta^{-1/2} \exp\left[-\frac{x_0^2}{2\Delta^2}\right]. \tag{5.100}$$

It is instructive to follow the time evolution of such a wave packet associated with a free particle. We then have

$$\psi(x,t) = \int_{-\infty}^{\infty} \mathrm{d}x_0 \, U_{\text{free}}(x,t;x_0,t_0)\psi(x_0,t_0).$$

Substituting (5.99) and (5.100) in that expression, one finds with some manipulations that (Exercise 5.18)

$$\begin{aligned}\psi(x,t) &= \pi^{-1/4}\Delta^{-1/2}\left[\frac{m}{2\pi i\hbar(t-t_0)}\right]^{1/2} \\ &\quad \times \int_{-\infty}^{\infty} \mathrm{d}x_0 \, \exp\left[-\alpha\left(x_0 + \frac{imx}{2\hbar\alpha(t-t_0)}\right)^2 + \frac{imx^2}{4\hbar\alpha(t-t_0)\Delta^2}\right],\end{aligned}$$

where

$$\alpha = \frac{1}{2\Delta^2} - \frac{im}{2\hbar(t-t_0)}.$$

A change of variable allows evaluating the integral as[9]

$$\begin{aligned}\psi(x,t) &= \pi^{-1/4}\Delta^{-1/2}\left[\frac{m}{2\pi i\hbar(t-t_0)}\right]^{1/2} \\ &\quad \times \exp\left[\frac{imx^2}{4\hbar\alpha(t-t_0)\Delta^2}\right]\int_{-\infty}^{\infty} \mathrm{d}q \exp\left\{-\alpha q^2\right\} \\ &= \pi^{-1/4}\Delta^{-1/2}\left[\frac{m}{2\pi i\hbar(t-t_0)}\right]^{1/2} \exp\left[\frac{imx^2}{4\hbar\alpha(t-t_0)\Delta^2}\right]\sqrt{\frac{\pi}{\alpha}}.\end{aligned}$$

Rearranging, we obtain

---

[9] Here the change of variable brings the path of integration in the complex plane. One easily shows, however, that the path can be brought back to the real axis.

$$\psi(x,t) = \pi^{-1/4}\Delta^{-1/2}\left[1+\frac{\mathrm{i}\hbar(t-t_0)}{m\Delta^2}\right]^{-1/2} \times \exp\left[-\frac{x^2}{2\Delta^2}\left(1+\frac{\mathrm{i}\hbar(t-t_0)}{m\Delta^2}\right)^{-1}\right]. \tag{5.101}$$

The probability density to find, at time t, the particle at point x is equal to $|\psi(x,t)|^2$ and is given by

$$|\psi(x,t)|^2 = \frac{1}{\sqrt{\pi}\Delta}\left[1+\frac{\hbar^2(t-t_0)^2}{m^2\Delta^4}\right]^{-1/2} \times \exp\left[-\frac{x^2}{\Delta^2}\left(1+\frac{\hbar^2(t-t_0)^2}{m^2\Delta^4}\right)^{-1}\right]. \tag{5.102}$$

This means that for large $t-t_0$, the standard deviation of the probability density grows as $t-t_0$. It is interesting to observe that the wave packet of a particle with a mass equal to the electron's grows from an initial standard deviation of 1 Å to a standard deviation of 1 m in 10^{-6} s.

Rapid spread of the wave packet is a characteristic of the free particle. The behavior of a bound particle is different. To see this, let us consider a harmonic oscillator of mass m and angular frequency ω. Assume the initial wave function is a Gaussian function centered on $\bar{x}$ and given as

$$\begin{aligned}\psi(x_0,t_0) &= \left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\exp\left\{-\frac{m\omega}{2\hbar}(x_0-\bar{x})^2\right\} \\ &= \left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\exp\left\{-\frac{1}{2}\left(\xi_0-\bar{\xi}\right)^2\right\}.\end{aligned} \tag{5.103}$$

Here ξ_0 and $\bar{\xi}$ are related to x_0 and $\bar{x}$ by (5.94).

Eigenvalues and eigenfunctions of the harmonic oscillator were obtained in Sect. 5.7. The propagator is given by

$$\begin{aligned}U_{\text{osc}}(x,t;x_0,t_0) &= \sum_{n=0}^{\infty}\mathrm{e}^{-\mathrm{i}E_n(t-t_0)/\hbar}\psi_n(x)\psi_n^*(x_0) \\ &= \sum_{n=0}^{\infty}\exp\left\{-\mathrm{i}\omega\left(n+\frac{1}{2}\right)(t-t_0)\right\}\left(\frac{m\omega}{\pi\hbar}\right)^{1/2} \\ &\quad\times\frac{1}{2^n n!}\mathrm{e}^{-(\xi^2+\xi_0^2)/2}H_n(\xi)\,H_n(\xi_0)\ .\end{aligned} \tag{5.104}$$

This series is evaluated in Exercise (5.21), with the result that

$$U_{\text{osc}}(x,t;x_0,t_0) = \left[\frac{\omega m}{2\pi\mathrm{i}\hbar\sin\left[\omega(t-t_0)\right]}\right]^{1/2} \times\exp\left\{\frac{\mathrm{i}}{2\sin\left[\omega(t-t_0)\right]}\left\{(\xi^2+\xi_0^2)\cos\left[\omega(t-t_0)\right]-2\xi\xi_0\right\}\right\}. \tag{5.105}$$

To obtain the time evolution of the wave packet, we will use properties of the Hermite polynomials and their generating function. By means of the latter, let us first evaluate the following integral:

$$I=\int_{-\infty}^{\infty}\mathrm{d}\xi_0\,\mathrm{e}^{-\xi_0^2/2}H_n\left(\xi_0\right)\exp\left\{-\frac{1}{2}\left(\xi_0-\bar{\xi}\right)^2\right\}.$$

One has

$$\begin{aligned}
I&=\int_{-\infty}^{\infty}\mathrm{d}\xi_0\,\exp\left\{-\frac{1}{2}\xi_0^2-\frac{1}{2}\left(\xi_0-\bar{\xi}\right)^2\right\}\left[\frac{\partial^n}{\partial s^n}\exp\left\{-s^2+2s\xi_0\right\}\right]_{s=0}\\
&=\left[\frac{\partial^n}{\partial s^n}\int_{-\infty}^{\infty}\mathrm{d}\xi_0\,\exp\left\{-\left(\xi_0-\frac{\bar{\xi}+2s}{2}\right)^2+s\bar{\xi}-\frac{1}{4}\bar{\xi}^2\right\}\right]_{s=0}\\
&=\left[\frac{\partial^n}{\partial s^n}\left\{\sqrt{\pi}\exp\left(s\bar{\xi}-\frac{1}{4}\bar{\xi}^2\right)\right\}\right]_{s=0}\\
&=\sqrt{\pi}\bar{\xi}^n\mathrm{e}^{-\bar{\xi}^2/4}.
\end{aligned}$$

With this result and expression (5.104) for the propagator, one finds that

$$\begin{aligned}
\psi(x,t)&=\int_{-\infty}^{\infty}\mathrm{d}x_0\,U_{\mathrm{osc}}(x,t;x_0,t_0)\psi(x_0,t_0)\\
&=\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\exp\left\{-\frac{1}{4}\bar{\xi}^2-\frac{1}{2}\xi^2\right\}\\
&\quad\times\sum_{n=0}^{\infty}\frac{1}{n!2^n}\exp\left\{-\mathrm{i}\omega\left(n+\frac{1}{2}\right)(t-t_0)\right\}\bar{\xi}^nH_n(\xi)\\
&=\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\exp\left\{-\frac{1}{4}\bar{\xi}^2-\frac{1}{2}\xi^2-\frac{\mathrm{i}\omega}{2}(t-t_0)\right\}\\
&\quad\times\sum_{n=0}^{\infty}\frac{1}{n!}\left[\frac{\bar{\xi}}{2}\mathrm{e}^{-\mathrm{i}\omega(t-t_0)}\right]^nH_n(\xi)\\
&=\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\exp\left\{-\frac{1}{4}\bar{\xi}^2-\frac{1}{2}\xi^2-\frac{\mathrm{i}\omega}{2}(t-t_0)\right\}\\
&\quad\times\exp\left[-\frac{\bar{\xi}^2}{4}\mathrm{e}^{-2\mathrm{i}\omega(t-t_0)}+\xi\bar{\xi}\mathrm{e}^{-\mathrm{i}\omega(t-t_0)}\right].
\end{aligned}\tag{5.106}$$

Properties of the generating function have again been used. Equation (5.106) represents the wave packet at time t.

The meaning of this result becomes clearer upon evaluation of $|\psi(x,t)|^2$, the probability density to find, at time t, the particle at point x. One easily obtains

$$\begin{aligned}
|\psi(x,t)|^2&=\left(\frac{m\omega}{\pi\hbar}\right)^{1/2}\exp\left\{-\left[\xi-\bar{\xi}\cos[\omega(t-t_0)]\right]^2\right\}\\
&=\left(\frac{m\omega}{\pi\hbar}\right)^{1/2}\exp\left\{-\frac{m\omega}{\hbar}\left[x-\bar{x}\cos[\omega(t-t_0)]\right]^2\right\}.
\end{aligned}\tag{5.107}$$

Thus the wave packet center oscillates as a classical oscillator with angular frequency ω. The wave packet does not spread.

One can express the wave packet differently by going back to the next-to-last equality in (5.106). Making use of (5.95) and rearranging, one obtains

$$\psi(x,t) = \exp\left\{-\frac{1}{4}\bar{\xi}^2 - \frac{\mathrm{i}\omega}{2}(t-t_0)\right\}\sum_{n=0}^{\infty}\frac{1}{\sqrt{n!}}\left\{\frac{\bar{\xi}}{\sqrt{2}}\mathrm{e}^{-\mathrm{i}\omega(t-t_0)}\right\}^n \psi_n(\xi).$$

Letting

$$\alpha = \frac{\bar{\xi}}{\sqrt{2}}\mathrm{e}^{-\mathrm{i}\omega(t-t_0)} = \frac{\bar{\xi}}{\sqrt{2}}\mathrm{e}^{-\mathrm{i}\theta_\alpha}$$

and switching to the ket notation, one finds that

$$|\psi(t)\rangle = \mathrm{e}^{-\mathrm{i}\theta_\alpha/2}\mathrm{e}^{-|\alpha|^2/2}\sum_{n=0}^{\infty}\frac{1}{\sqrt{n!}}\alpha^n|E_n\rangle. \tag{5.108}$$

By means of (5.91) it is easy to check that

$$A|\psi(t)\rangle = \alpha|\psi(t)\rangle. \tag{5.109}$$

Eigenvectors of the annihilation operator are called *coherent states*. We have shown that the wave packet (5.106) is, at all t, a coherent state.

One can show that in a coherent state, the product of standard deviations of X and P assumes its minimum value $\hbar/2$. Coherent states are particularly important in quantum optics, where each degree of freedom of the electromagnetic field is a harmonic oscillator.[10]

5.9 Appendix

5.9.1 The Dirac Delta Function

The *Dirac delta function*,[11] denoted by $\delta(x-x_0)$, is defined so that for any infinitely differentiable function $f(x)$

$$\int_{-\infty}^{\infty}\mathrm{d}x\, f(x)\delta(x-x_0) = f(x_0). \tag{5.110}$$

---

[10] Good discussions of coherent states can be found in [53], Compl. $\mathrm{G_V}$, [120] and [132]. See also the bibliography of [216].

[11] Properties of the Dirac delta function are presented more rigorously in [65], Sect. 3.13 and in [160], App. A. Maple V produces $\delta(x)$ and its derivative of order n through the commands `Dirac(x)` and `Dirac(n,x)`, respectively. Mathematica produces them by means of the package ``Calculus`DiracDelta` ``.

Consider the set of nonnegative functions $f(x)$ that vanish at x_0. One must have

$$\int_{-\infty}^{\infty} \mathrm{d}x\, f(x)\delta(x-x_0) = 0\,.$$

This will hold for all these functions only if $\delta(x-x_0)$ vanishes for $x \neq x_0$. On the other hand, letting $f(x) = 1$, one finds that

$$\int_{-\infty}^{\infty} \mathrm{d}x\, \delta(x-x_0) = 1\,.$$

This entails that $\delta(x-x_0)$ is infinite at $x = x_0$. As a matter of fact, Dirac's delta is not a genuine function. It can be seen as a limit, as $\varepsilon \longrightarrow 0$, of functions $\delta_\varepsilon(x-x_0)$, where

$$\delta_\varepsilon(x-x_0) = \frac{1}{\varepsilon\sqrt{\pi}} \exp\left(-\frac{(x-x_0)^2}{\varepsilon^2}\right). \tag{5.111}$$

In other words, $\delta(x-x_0)$ is the limit of a Gaussian function centered on x_0, infinitely high and infinitely narrow, and covering an area equal to 1.

The limiting process considered should really be understood in a particular way. Every time $\delta(x-x_0)$ appears in an integral, the limit must be taken outside the integral. For example

$$\int_{-\infty}^{\infty} \mathrm{d}x\, f(x)\delta(x-x_0) = \lim_{\varepsilon\to 0} \int_{-\infty}^{\infty} \mathrm{d}x\, f(x)\delta_\varepsilon(x-x_0). \tag{5.112}$$

For any nonvanishing real constant c one has

$$\delta\left\{c(x-x_0)\right\} = \frac{1}{|c|}\delta(x-x_0). \tag{5.113}$$

To prove this it is enough to show that for any function $f(x)$,

$$\int_{-\infty}^{\infty} \mathrm{d}x\, f(x)\delta\left\{c(x-x_0)\right\} = \frac{1}{|c|}f(x_0).$$

The proof simply follows from a change of variable $y = cx$, the integration limits being correctly taken care of.

Let $g(x)$ be a function that has only one zero, of order one, at $x = \bar{x}$. Then

$$\delta\left[g(x)\right] = \left|\left(\frac{\mathrm{d}g}{\mathrm{d}x}\right)_{\bar{x}}\right|^{-1} \delta(x-\bar{x}). \tag{5.114}$$

To prove this one simply considers the behavior of $g(x)$ around $x = \bar{x}$, that is,

$$g(x) = g(\bar{x})+(x-\bar{x})\left(\frac{\mathrm{d}g}{\mathrm{d}x}\right)_{\bar{x}} = (x-\bar{x})\left(\frac{\mathrm{d}g}{\mathrm{d}x}\right)_{\bar{x}}, \qquad |x-\bar{x}| \ll 1.$$

If $g(x)$ has several zeros of order one at points $\bar{x}_i$,

$$\delta\left[g(x)\right] = \sum_i \left|\left(\frac{\mathrm{d}g}{\mathrm{d}x}\right)_{\bar{x}_i}\right|^{-1} \delta(x-\bar{x}_i). \tag{5.115}$$

One can introduce derivatives of the delta function of all orders. They are defined so that properties of integration by parts are satisfied. Thus for any $f(x)$,

$$\begin{aligned}\int_{-\infty}^{\infty} \mathrm{d}x\, f(x)\frac{\mathrm{d}}{\mathrm{d}x}\delta(x-x_0) &= \left[f(x)\delta(x-x_0)\right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \mathrm{d}x\,\frac{\mathrm{d}f}{\mathrm{d}x}\delta(x-x_0)\\ &= -\left(\frac{\mathrm{d}f}{\mathrm{d}x}\right)_{x_0}.\end{aligned} \tag{5.116}$$

From (5.116) one shows that for any function $F(x)$,

$$\int_{-\infty}^{\infty} \mathrm{d}x\, F(x)\left\{(x-x_0)\frac{\mathrm{d}}{\mathrm{d}x}\delta(x-x_0)\right\} = -F(x_0),$$

from which we get

$$(x-x_0)\frac{\mathrm{d}}{\mathrm{d}x}\delta(x-x_0) = -\delta(x-x_0). \tag{5.117}$$

Higher-order derivatives of the delta function are defined similarly.

There are many representations of the delta function in terms of limits, series or integrals. Let $\tilde{f}(k)$ be the Fourier transform of a function $f(x)$. One has

$$\begin{aligned}f(x) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}k\, \mathrm{e}^{\mathrm{i}kx}\tilde{f}(k),\\ \tilde{f}(k) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}x\, \mathrm{e}^{-\mathrm{i}kx}f(x).\end{aligned}$$

With some manipulations one obtains

$$\begin{aligned}f(x) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}k\, \mathrm{e}^{\mathrm{i}kx}\left[\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{d}x'\, \mathrm{e}^{-\mathrm{i}kx'}f(x')\right]\\ &= \int_{-\infty}^{\infty} \mathrm{d}x' f(x')\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}k\, \mathrm{e}^{\mathrm{i}k(x-x')}\right].\end{aligned}$$

Comparing this equation with the one that defines the delta function, one finds that

$$\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{d}k\, \mathrm{e}^{\mathrm{i}k(x-x')} = \delta(x-x'). \tag{5.118}$$

Furthermore, we rather easily convince ourselves that

$$\delta(x-x') = \frac{d}{dx}\theta(x-x'), \tag{5.119}$$

where $\theta(x-x')$, the *Heaviside step function*, is defined as

$$\theta(x-x') = \begin{cases} 0 & \text{if } x < x', \\ 1 & \text{if } x > x'. \end{cases} \tag{5.120}$$

It is clear that the delta function can be generalized to several variables. Let V be a d-dimensional volume and let α be a set of variables representing a point in V. If $\mathrm{d}\alpha$ stands for a volume element, the function $\delta(\alpha - \alpha_0)$ is defined so that for any infinitely differentiable function $f(\alpha)$

$$\int_V \mathrm{d}\alpha\, f(\alpha)\delta(\alpha-\alpha_0) = f(\alpha_0). \tag{5.121}$$

5.9.2 Infinite-Dimensional Vector Spaces

The dimension of a vector space is infinite if there are an infinite number of linearly independent vectors.[12] Such is the case, for instance, with the state space of a harmonic oscillator.

Let $\mathcal{H}$ be an infinite-dimensional vector space equipped with a scalar product. We say that an infinite sequence of vectors $|\phi_1\rangle$, $|\phi_2\rangle$, ... *converges* if the norm of $|\phi_n\rangle - |\phi_m\rangle$ goes to zero whenever n and m go to infinity. A convergent sequence has a *limit* if there exists a vector $|\phi\rangle$ in $\mathcal{H}$ such that the norm of $|\phi_n\rangle - |\phi\rangle$ goes to zero whenever n goes to infinity. The vector space is a *Hilbert space* if every convergent sequence has a limit. Hilbert spaces normally dealt with in quantum mechanics are *separable*, which means they have a denumerable basis.

Examples Consider the set of complex functions over $\mathcal{R}$ such that

$$\int_{-\infty}^{\infty} \mathrm{d}x\, |f(x)|^2 < \infty. \tag{5.122}$$

This set makes up a vector space. Define the scalar product of $|f\rangle$ and $|g\rangle$ as[13]

---

[12] For detailed treatment of infinite-dimensional vector spaces, the reader is referred to [65], Chap. 3, [81] and [167]. See also [13], Sect. 1.4, [53], Chap. 2 and [127], Chaps. 1–4.

[13] It happens that $(|f\rangle, |f\rangle) = 0$ without $f(x)$ vanishing everywhere. This is really a minor problem, solved by identifying functions that differ only on a set of measure zero.

$$(|f\rangle,|g\rangle) \equiv \int_{-\infty}^{\infty} \mathrm{d}x f^*(x)g(x). \tag{5.123}$$

Then the vector space becomes a Hilbert space, denoted by $L^2(\mathcal{R})$. For any positive integer n one similarly defines $L^2(\mathcal{R}^n)$. ♣

It is tempting to define bras in a Hilbert space as linear functionals, just as in finite-dimensional spaces. In this case the correspondence between bras and kets is not one-to-one. Take for instance $L^2(\mathcal{R})$ and let $f(x) = \mathrm{e}^{\mathrm{i}px}$ $(p > 0)$. The right-hand side of (5.123) defines a linear functional, hence a bra. But there is no ket associated with $f(x)$, since inequality (5.122) is not satisfied. One way to solve the problem consists in enlarging the vector space so as to include these kets. But one can also define bras as *continuous* linear functionals, i.e. linear functionals such that for some $c > 0$, $\mathrm{norm}(f|g\rangle) \leq c\,\mathrm{norm}(|g\rangle)$ for all $|g\rangle$. Then the correspondence between bras and kets in $L^2(\mathcal{R}^n)$ is one-to-one.

An operator A is, in general, defined on a subset D_A of $\mathcal{H}$, called the *domain* of A. The adjoint $A^\dagger$ is defined so that $(|\phi\rangle, A|\psi\rangle) = (A^\dagger|\phi\rangle, |\psi\rangle)$. In general $D_{A^\dagger} \neq D_A$. The operator A is *Hermitian* if $(|\phi\rangle, A|\psi\rangle) = (|\psi\rangle, A|\phi\rangle)^*$ for all $|\phi\rangle$ and $|\psi\rangle$ in D_A. Furthermore, A is *self-adjoint* if it is Hermitian and if $D_{A^\dagger} = D_A$. Then $A^\dagger = A$. In infinite-dimensional spaces some Hermitian operators are not self-adjoint. Dynamical variables are associated with self-adjoint operators.

Operators X and P are self-adjoint in $L^2(\mathcal{R})$. Nonetheless, they do not have eigenvectors in that space. Here again the problem is solved by enlarging the vector space so as to include functions $\mathrm{e}^{\mathrm{i}px}$ and Dirac's deltas. We should point out that the spectral decomposition theorem holds even within $L^2(\mathcal{R})$. For operator X, for instance, one defines projectors $P(\bar{x})$ such that

$$P(\bar{x})|f\rangle = |f'\rangle, \quad \text{where} \quad f'(x) = \begin{cases} f(x) & \text{if } x \leq \bar{x}\,, \\ 0 & \text{if } x > \bar{x}\,. \end{cases} \tag{5.124}$$

With a suitable definition of the integral, one can show that

$$I = \int_{-\infty}^{\infty} \mathrm{d}P(\bar{x}), \qquad X = \int_{-\infty}^{\infty} \mathrm{d}P(\bar{x})\,\bar{x}. \tag{5.125}$$

By means of mathematical objects introduced in Sect. 5.1, the projector $P(\bar{x})$ is represented by $\int_{-\infty}^{\bar{x}} \mathrm{d}x\, |x\rangle\langle x|$. Intuitively, our strategy has been to introduce a vector space large enough to include all vectors that can be useful and to define the state space as a subspace of that big space.

5.9.3 Hermite Polynomials

The *Hermite polynomial* of order n,[14] denoted by $H_n(\xi)$, is defined as

---

[14] Detailed discussion of Hermite polyomials can be found in [143], Chap. 4. In Maple V, the polynomial $H_n(x)$ is produced by the command `H(n,x)` from

$$H_n(\xi) = (-1)^n \mathrm{e}^{\xi^2} \frac{\mathrm{d}^n}{\mathrm{d}\xi^n} \mathrm{e}^{-\xi^2} . \tag{5.126}$$

For $n = 0$, 1, 2 and 3 one has

$$\begin{aligned} H_0(\xi) &= 1 , \\ H_1(\xi) &= 2\xi , \\ H_2(\xi) &= 4\xi^2 - 2 , \\ H_3(\xi) &= 8\xi^3 - 12\xi . \end{aligned}$$

In general, $H_n(\xi)$ is a polynomial of degree n in ξ. $H_n(\xi)$ has the same parity as n, that is,

$$H_n(-\xi) = (-1)^n H_n(\xi) . \tag{5.127}$$

Let $W(\xi, s)$ be the analytic function of ξ and s defined as

$$W(\xi, s) = \exp\left\{-s^2 + 2s\xi\right\} . \tag{5.128}$$

We will show that

$$W(\xi, s) = \sum_{n=0}^{\infty} \frac{1}{n!} H_n(\xi) s^n . \tag{5.129}$$

To do so, it is enough to show that $H_n(\xi)$ coincides with

$$\left(\frac{\partial^n W}{\partial s^n}\right)_{s=0} .$$

But

$$\begin{aligned} \left(\frac{\partial^n W}{\partial s^n}\right)_{s=0} &= \mathrm{e}^{\xi^2} \left(\frac{\partial^n}{\partial s^n} \mathrm{e}^{-(s-\xi)^2}\right)_{s=0} = \mathrm{e}^{\xi^2} \left(\frac{\mathrm{d}^n}{\mathrm{d}u^n} \mathrm{e}^{-u^2}\right)_{u=-\xi} \\ &= \mathrm{e}^{\xi^2} (-1)^n \left(\frac{\mathrm{d}^n}{\mathrm{d}u^n} \mathrm{e}^{-u^2}\right)_{u=\xi} = H_n(\xi) . \end{aligned}$$

$W(\xi, s)$ is called the *generating function* of Hermite polynomials.

The equation

$$\exp\left\{-s^2 + 2\xi s\right\} = \sum_{n=0}^{\infty} \frac{1}{n!} H_n(\xi) s^n$$

can be differentiated with respect to ξ, and the coefficients of a given power of s on both sides must be equal. This implies that

---

the package `orthopoly`. In Mathematica, $H_n(x)$ is produced by the command `HermiteH[n,x]`.

$$\frac{\mathrm{d}}{\mathrm{d}\xi}H_n(\xi) = 2nH_{n-1}(\xi). \tag{5.130}$$

A similar differentiation with respect to s entails that

$$H_{n+1}(\xi) = 2\xi H_n(\xi) - 2nH_{n-1}(\xi). \tag{5.131}$$

Combining these two equations, one obtains a second-order differential equation for Hermite polynomials:

$$\frac{\mathrm{d}^2 H_n}{\mathrm{d}\xi^2} - 2\xi\frac{\mathrm{d}H_n}{\mathrm{d}\xi} + 2nH_n = 0. \tag{5.132}$$

The generating function of Hermite polynomials is used to compute integrals like

$$\int_{-\infty}^{\infty} \mathrm{d}\xi\, \xi^l H_m(\xi)H_n(\xi)\mathrm{e}^{-\xi^2}. \tag{5.133}$$

Let us work out the case where $l = 0$. It is easy to show that

$$\int_{-\infty}^{\infty} \mathrm{d}\xi \exp\left\{-s^2 + 2\xi s\right\} \exp\left\{-t^2 + 2\xi t\right\} \mathrm{e}^{-\xi^2}$$
$$= \sqrt{\pi}\mathrm{e}^{2st} = \sqrt{\pi}\sum_{n=0}^{\infty}\frac{1}{n!}(2st)^n .$$

On the other hand

$$\int_{-\infty}^{\infty} \mathrm{d}\xi \exp\left\{-s^2 + 2\xi s\right\} \exp\left\{-t^2 + 2\xi t\right\} \mathrm{e}^{-\xi^2}$$
$$= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{s^n t^m}{n!m!}\int_{-\infty}^{\infty} \mathrm{d}\xi\, H_m(\xi)H_n(\xi)\mathrm{e}^{-\xi^2} .$$

Comparing, in the right-hand sides of the last two equations, the same powers of s and t, one sees that

$$\int_{-\infty}^{\infty} \mathrm{d}\xi\, H_m(\xi)H_n(\xi)\mathrm{e}^{-\xi^2} = \sqrt{\pi}2^n n!\delta_{nm}. \tag{5.134}$$

For any complex number z such that $|z| < 1$, one can show that

$$\sum_{n=0}^{\infty}\frac{z^n}{2^n n!}H_n(\xi)H_n(\xi_0) = \frac{1}{\sqrt{1-z^2}}\exp\left\{\frac{2\xi\xi_0 z - (\xi^2+\xi_0^2)z^2}{1-z^2}\right\}. \tag{5.135}$$

The Hermite polynomials make up a complete set of functions in the interval $(-\infty, \infty)$. Specifically, any real function $f(\xi)$ that is piecewise smooth in the interval $(-\infty, \infty)$ and such that

$$\int_{-\infty}^{\infty} \mathrm{d}\xi \, |f(\xi)|^2 \mathrm{e}^{-\xi^2} < \infty \tag{5.136}$$

can be expanded as

$$f(\xi) = \sum_{n=0}^{\infty} c_n H_n(\xi) , \tag{5.137}$$

where

$$c_n = \frac{1}{\sqrt{\pi} 2^n n!} \int_{-\infty}^{\infty} \mathrm{d}\xi \, f(\xi) H_n(\xi) \mathrm{e}^{-\xi^2} . \tag{5.138}$$

The series converges to $f(\xi)$ at every point where f is continuous.

Exercises

5.1. Prove the following two identities for the second derivative of the delta function:

$$\int_{-\infty}^{\infty} \mathrm{d}x \, f(x) \frac{\mathrm{d}^2}{\mathrm{d}x^2} \delta(x - x_0) = \left(\frac{\mathrm{d}^2 f}{\mathrm{d}x^2} \right)_{x_0} ,$$

$$(x - x_0)^2 \frac{\mathrm{d}^2}{\mathrm{d}x^2} \delta(x - x_0) = 2\delta(x - x_0) .$$

5.2. Let ε be real and positive and let

$$\Delta_\varepsilon(x - x_0) = \begin{cases} 0 & \text{if } x - x_0 < -\varepsilon , \\ (2\varepsilon)^{-1} & \text{if } -\varepsilon < x - x_0 < \varepsilon , \\ 0 & \text{if } x - x_0 > \varepsilon . \end{cases}$$

Show that

$$\int_{-\infty}^{\infty} \mathrm{d}x \, \Delta_\varepsilon(x - x_0) = 1 .$$

Can one write

$$\delta(x - x_0) = \lim_{\varepsilon \to 0} \Delta_\varepsilon(x - x_0) ?$$

5.3. Show that (5.1) and (5.12) imply (5.13).

5.4. Check (5.18).

5.5. Let $\psi(x,t)$ be the normalized wave function of a particle. Show that

$$\langle \psi(t) | P^n | \psi(t) \rangle = \int \mathrm{d}x \, \psi^*(x,t) \left(-\mathrm{i}\hbar \frac{\partial}{\partial x} \right)^n \psi(x,t) = \int \mathrm{d}p \left| \tilde{\psi}(p,t) \right|^2 p^n .$$

5.6. Show that for all functions $\phi(x)$ and $\psi(x)$ that go to 0 as $|x| \to \infty$,

$$\langle\phi|P|\psi\rangle - \langle\psi|P|\phi\rangle^* = 0\,,$$

thereby checking the hermiticity of P on that function space.

5.7. Let Δ be positive and let

$$\psi(x) = \langle x|\psi\rangle = \pi^{-1/4}\Delta^{-1/2}\exp\left\{-\frac{x^2}{2\Delta^2}\right\}.$$

Compute $\langle\psi|X|\psi\rangle$ and $\langle\psi|P|\psi\rangle$.

5.8. Even eigenfunctions associated with the square well potential's discrete spectrum are given in (5.63), where k and κ satisfy (5.64). Find a constant A such that $\int_{-\infty}^{\infty} \mathrm{d}x\, |\psi(x)|^2 = 1$.

5.9. Find, graphically or numerically, the two lowest values of E/V_0 that satisfy (5.67), where $2ma^2V_0/\hbar^2 = 100$. Compute E/V_0 with an error less than 0.005.

5.10. Derive (5.76) and (5.77).

5.11. The square potential barrier is illustrated in Fig. 5.2. Find a solution of the Hamiltonian's eigenvalue equation which is continuous and has a continuous derivative everywhere, and which for $-a < x < a$ is given by $\psi(x) = A\cosh(\kappa x)$, where

$$\kappa = \left\{\frac{2m}{\hbar^2}(V_0 - E)\right\}^{1/2}.$$

5.12. Let $V(x)$ be the step potential shown in Fig. 5.4, that is,

$$V(x) = V_0\,\theta(x) = \begin{cases} 0 & \text{if } x < 0\,, \\ V_0 & \text{if } x > 0\,, \end{cases}$$

and let $E > V_0 > 0$. Let k and k' be defined as

$$k = \left\{\frac{2mE}{\hbar^2}\right\}^{1/2}, \qquad k' = \left\{\frac{2m}{\hbar^2}(E - V_0)\right\}^{1/2}.$$

a) In each region, write the most general solution of the eigenvalue equation for H. Next consider the particular case where, in region $x > 0$, there is a transmitted wave only.
b) Use the continuity of the wave function and its derivative at $x = 0$ to express all coefficients in terms of only one.
c) What is the probability that a particle coming from the region $x < 0$ is reflected by the potential step?

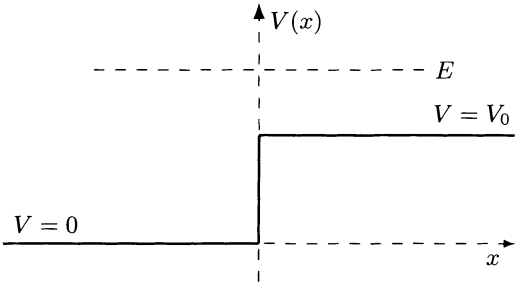

Fig. 5.4. Step potential

5.13. Show that the first five Hermite polynomials are given by

$$\begin{aligned}
H_0(\xi) &= 1\,,\\
H_1(\xi) &= 2\xi\,,\\
H_2(\xi) &= 4\xi^2 - 2\,,\\
H_3(\xi) &= 8\xi^3 - 12\xi\,,\\
H_4(\xi) &= 16\xi^4 - 48\xi^2 + 12\,.
\end{aligned}$$

5.14. Check that functions $H_2(\xi)\mathrm{e}^{-\xi^2/2}$ and $H_3(\xi)\mathrm{e}^{-\xi^2/2}$, with ξ as in (5.94), satisfy the eigenvalue equation of H for the harmonic oscillator.

5.15. Let W be the generating function of Hermite polynominals. Consider the integral

$$\int_{-\infty}^{\infty} \mathrm{d}\xi\, W(\xi, s) W(\xi, t) \xi \mathrm{e}^{-\xi^2}\,.$$

By evaluating this expression in two different ways, find the value of the following integrals, where m and n are nonnegative integers:

$$\int_{-\infty}^{\infty} \mathrm{d}\xi\, H_m(\xi) H_n(\xi) \xi \mathrm{e}^{-\xi^2}\,.$$

5.16. Let $|E_n\rangle$ be the normalized state vectors of the harmonic oscillator.
a) Show that for all kets $|E_n\rangle$,

$$\langle E_n|X|E_n\rangle = 0 = \langle E_n|P|E_n\rangle\,.$$

b) Compute $X^2|E_n\rangle$ and $P^2|E_n\rangle$.
c) Compute the standard deviations ΔX and ΔP in state $|E_n\rangle$.

5.17. In this problem we will obtain the eigenvalues and eigenvectors of the harmonic oscillator Hamiltonian by solving the differential equation. Let us first define

$$\xi = \sqrt{\frac{m\omega}{\hbar}}\,x, \qquad \lambda = \frac{2E}{\hbar\omega}, \qquad f(\xi) = \psi(x)e^{\xi^2/2}.$$

a) Show that in terms of these variables, the differential equation becomes

$$\frac{d^2 f}{d\xi^2} - 2\xi\frac{df}{d\xi} + (\lambda - 1)f = 0.$$

b) Show that this equation has series solutions of the form

$$f(\xi) = \sum_{n=0}^{\infty} a_n \xi^{s+n},$$

where $a_0 \neq 0$, $s = 0$ or 1, and where coefficients a_n satisfy the following recurrence equation:

$$a_{n+2} = \left(\frac{2s + 2n + 1 - \lambda}{(s+n+1)(s+n+2)}\right) a_n .$$

There are thus two independent power series, corresponding to even or odd values of the index n.

c) These power series are close to the function $\exp(\xi^2)$, and consequently $\psi(\xi)$ is not normalisable. Show that the series of even indices stops after a finite number of terms only if

$$\lambda = \begin{cases} 2n+1 & \text{if } s = 0\,, \\ 2n+3 & \text{if } s = 1\,. \end{cases}$$

The series of odd indices then cannot stop, unless $a_1 = 0$. With the above condition one obtains one normalizable solution.

d) For this solution show that $f(\xi)$ is a Hermite polynomial and that E coincides with the corresponding eigenvalue of H.

5.18. Fill in the steps in the calculation of the spreading of the free-particle Gaussian wave packet.

5.19.

a) Show that the propagator satisfies the equation

$$U(x,t;x_0,t_0) = \int_{-\infty}^{\infty} dx'\, U(x,t;x',t')U(x',t';x_0,t_0).$$

b) Check this relation for a free particle.

5.20. Show that

$$\lim_{t-t_0\to 0} \left[\frac{m}{2\pi i\hbar(t-t_0)}\right]^{1/2} \exp\left(\frac{im(x-x_0)^2}{2\hbar(t-t_0)}\right) = \delta(x-x_0).$$

5.21. Derive (5.105) from (5.135). Comment on the use of (5.135) at the limit of its domain of validity.

6 The Interpretation of Quantum Mechanics

In the first few chapters, we have introduced the formal objects and laws at the heart of quantum mechanics: state vector, Hermitian operators, eigenvalues, Schrödinger's equation, etc. In simple cases we have shown how to use them to account for situations which, although idealized, are not unlike the ones studied in the laboratory.

We will now look at the notions we have introduced, and try to understand better their logical and epistemological status. Quantum mechanics is much harder to interpret than earlier physical theories. This comes from the fact that its objects are not directly related to experience and that their behavior often runs against common sense.

6.1 Formalism and Interpretation Rules

To put the interpretation of quantum mechanics in proper context, we should first recall certain general features of physical theories. Much effort was paid, in the twentieth century, to understand the nature of a physical theory.[1] Although significant progress has been made, the problem is certainly not fully elucidated. One can say, however, that the ultimate purpose of a physical theory is to account for a set of phenomena.

By *phenomena* here we mean natural events directly or almost directly perceptible by sense organs. A theory accounts for a set of phenomena if it brings understanding and, to a certain point, allows prediction. The notion of understanding is intricate, but it involves an intuitive grasp of order or unity in the apparent diversity or chaos of phenomena. Prediction means the capability to describe in advance, and quantitatively, observations that will be made in specified circumstances or following specific experimental procedures.

Every physical theory attempting to account for a sufficiently large set of phenomena is necessarily based on a *mathematical formalism*. The mathematical formalism is made up of a number of symbols that obey precise rules. The symbols stand for abstract objects, in the sense that the formalism itself sets no connection between them and the world of phenomena.

---

[1] See for instance [44], [151] and [221].

L. Marchildon, *Quantum Mechanics*

But a physical theory does not end with a mathematical formalism. It brings something more by endowing the symbols with a concrete *interpretation.* This means it relates the symbols, either directly or sometimes very indirectly, to the set of phenomena it tries to account for. At a level close to phenomena, this relation is often effected by means of entities defined in an *operational* way, that is, defined by the process carried out to measure them. Directly related to the world of phenomena, such entities are also related in a more or less complicated way to symbols in the mathematical formalism and to abstract objects they stand for.

Classical mechanics provides a particularly simple illustration of these ideas. Briefly, the mathematical formalism there is made up of constants and real variables that obey differential equations. Time is the independent variable while dependent variables represent, for example, the coordinates of sufficiently small particles. Constants are associated with particle masses, friction coefficients, spring constants, etc. All these parameters can be defined operationally.

In classical electrodynamics, the connection between some symbols of the formalism and phenomena is less straightforward. The symbols $\boldsymbol{E}$ and $\boldsymbol{B}$ of the electromagnetic field, for instance, belong to the mathematical formalism. They are related to the world of phenomena indirectly, through entities defined operationally, such as a charged particle and its acceleration in a given situation.

In quantum mechanics, the relation between mathematical symbols and phenomena is much more indirect. The mathematical formalism, we recall, is based on vector spaces, in particular Hilbert spaces, and their vectors; on Hermitian operators, their eigenvalues and their eigenvectors; on unitary operators, etc. At the other extreme, phenomena are processes or macroscopic events such as trails in a bubble chamber, dots on a photographic plate, a light spectrum or, in general, the position of a pointer or the numbers shown by a piece of equipment.

The interpretation of the quantum-mechanical formalism is effected through intermediate entities that can be called *microobjects.* They are, for example, electrons, nuclei, atoms, molecules, photons, etc. The fundamental symbols of the formalism, such as state vector, Hermitian operators, etc., are most closely related to these microobjects, through the interpretation rules we have given in Chap. 3.[2] To eventually connect mathematical symbols with the world of phenomena, one assumes that microobjects interact one way or another with macroscopic apparatus. The nature of these interactions will be examined in Chap. 21. For the moment, let us say that the interaction is supposed to set up a correlation between the value of a dynamical va-

---

[2] Microobjects, or entities with which symbols of the formalism are directly connected, are not necessarily microscopic. One could call them *quantons* ([47], Chap. 5).

riable associated with the microobject and a numerical result shown by the measurement device.

It is no overstatement to say that all specialists of quantum mechanics agree on the mathematical formalism and on the formalism's predictions for numerical values shown by macroscopic measurement apparatus. These are certainly the most important aspects of the theory. But the consensus vanishes as soon as one turns to questions of interpretation. In the next section we will see that symbols of the formalism can be interpreted in terms of microobjects in different ways. Moreover, investigators do not agree on the proper way to describe a measurement apparatus. Following Bohr, many believe that the behavior of a measurement apparatus is irreducibly classical, while others would like to treat it ultimately as a quantum system. Finally there are some who believe, like Einstein, that quantum mechanics does not provide an epistemologically satisfying explanation of the phenomena it tries to account for.

6.2 Interpretation of the State Vector

In quantum mechanics the state vector $|\psi\rangle$ refers to microobjects. We have hitherto met it associated with spin particles, H_2^+ ions, harmonic oscillators, etc. Let us try to delineate better the nature of this association.

In the *individual* interpretation, the one more widely used, the vector $|\psi\rangle$ represents the state of a specific microobject. One says, for instance, that a given hydrogen atom or harmonic oscillator is in the state represented by the vector $|\psi\rangle$ or, more simply, in state $|\psi\rangle$. Although quantum-mechanical predictions, statistical by nature, can be verified only by experiments made, for instance, on a large number of harmonic oscillators, one nonetheless says that each oscillator is in a state $|\psi\rangle$. The state vector is assigned to a microobject taken individually.

By contrast, in the *statistical* interpretation, the vector $|\psi\rangle$ represents an ensemble of identical and similarly prepared microobjects ([8], [177], [184], [210]). The three terms "ensemble," "identical" and "similarly prepared" call for explanation. *Ensemble* is understood in a probabilistic or statistical sense. It is a conceptual collection of an infinite number of objects endowed with certain properties that define this collection. An ensemble can be realized more or less adequately by a finite collection of real objects. By *identical microobjects*, one means for instance particles of mass M in a given potential $V(x)$. The particles need not, however, have the same energy. Or else one means particles carrying a spin $\boldsymbol{J}$, but not with the same value m of spin projection along a given axis. In other words, identical microobjects are characterized by identical state spaces. As for the predicate *similarly prepared*, it refers to certain events in the previous history of microobjects. We will examine this in detail in Sect. 6.3.

In the statistical interpretation of the state vector, statements like "The harmonic oscillator is in state $|\psi\rangle$" are rejected. The nature of the state vector is statistical, which entails that it can properly refer only to an ensemble of microobjects.

One can bring forward a third interpretation of the state vector, that can be called *epistemic* ([107], Chap. 3; [181]; [182]). The vector $|\psi\rangle$ does not then represent the state of a microobject or of an ensemble of microobjects, but rather the knowledge we have about the microobject, or the information we have on it. Of course, "we" does not refer to specific experimentalists, but to an ideal and abstract observer. The information one has on a microobject is linked to its previous history, that is, to the preparation process. But here one no longer refers to states, so that even the statement "An ensemble of identical harmonic oscillators is in state $|\psi\rangle$" must be rejected.

No matter what interpretation of the state vector is adopted, the predictions of quantum mechanics are always statistical. One can thus reasonably ask: Is there a theory, more fundamental than quantum mechanics or simply complementary to it, that allows categorical predictions of measurement results for arbitrary dynamical variables?

Following Bohr and most of the founders of quantum mechanics, the majority of contemporary physicists tend to answer this question negatively. Others, however, inspired by de Broglie and Einstein, try to develop theories that can predict individual measurement results. Such theories bring forth so-called *hidden variables* ([23]; [35]; [67], Chap. 13; [114]; [124], Chap. 7). They are parameters that characterize the state of a microobject more completely than does the state vector. Knowing the values of hidden variables would allow to predict measurement results of dynamical variables categorically.

In quantum mechanics the use of probabilistic language originates from a fundamental indeterminacy. In classical statistical physics, on the other hand, it comes from an incomplete knowledge of the values of some parameters. The introduction of hidden variables in the quantum-mechanical formalism brings quantum probabilities back to classical probabilities.

The example of a spin 1/2 particle shows how hidden variables work [22]. Suppose a spin 1/2 particle is prepared so that its spin projection on the z axis is positive. Quantum mechanics then predicts that a measurement of dynamical variable σ_z [equals to $(2/\hbar)J_z$] necessarily yields a $+1$ result. If $\hat{u}$ is an arbitrary unit vector, quantum mechanics makes no categorical prediction on the result of measuring dynamical variable $\hat{u}\cdot\boldsymbol{\sigma}$. It does predict, however, that the mean value of $\hat{u}\cdot\boldsymbol{\sigma}$ is given by

$$\langle\hat{u}\cdot\boldsymbol{\sigma}\rangle_{\mathrm{qm}} = \hat{u}\cdot\hat{z}\,. \tag{6.1}$$

We now show that one can introduce a hidden variable λ such that (i) knowledge of λ entails knowledge of the value of $\hat{u}\cdot\boldsymbol{\sigma}$ and (ii) the mean value of $\hat{u}\cdot\boldsymbol{\sigma}$, on the distribution of λ, coincide with $\langle\hat{u}\cdot\boldsymbol{\sigma}\rangle_{\mathrm{qm}}$, that is,

$$\int \mathrm{d}\lambda\, p(\lambda)\xi(\hat{u};\lambda) = \hat{u}\cdot\hat{z}\,. \tag{6.2}$$

Here $p(\lambda)$ is the probability density that the hidden variable has value λ, and $\xi(\hat{u};\lambda)$ is the value of $\hat{u}\cdot\boldsymbol{\sigma}$ when the hidden variable has value λ.

To show this, let us pick as hidden variable a unit vector $\hat{\lambda}$ oriented in an arbitrary direction and let us set

$$p(\hat{\lambda}) = \begin{cases} (2\pi)^{-1} & \text{if } \hat{\lambda}\cdot\hat{z} > 0\,, \\ 0 & \text{if } \hat{\lambda}\cdot\hat{z} \leq 0\,. \end{cases} \tag{6.3}$$

Clearly,

$$\int \mathrm{d}\hat{\lambda}\, p(\hat{\lambda}) = 1\,, \tag{6.4}$$

where the integral is taken on all directions of $\hat{\lambda}$, in other words on all solid angle. We specify that

$$\xi(\hat{u};\hat{\lambda}) = \begin{cases} 1 & \text{if } \hat{\lambda}\cdot\hat{u}' > 0\,, \\ -1 & \text{if } \hat{\lambda}\cdot\hat{u}' \leq 0\,. \end{cases} \tag{6.5}$$

Here, $\hat{u}'$ is a function of $\hat{u}$ that will soon be defined. Equation (6.5) will then give the value of $\hat{u}\cdot\boldsymbol{\sigma}$ as a function of the hidden variable $\hat{\lambda}$.

It is easy to show that (Exercise 6.1)

$$\int \mathrm{d}\hat{\lambda}\, p(\hat{\lambda})\xi(\hat{u};\hat{\lambda}) = 1 - \frac{2\theta'}{\pi}\,, \tag{6.6}$$

where θ' is the angle between $\hat{z}$ and $\hat{u}'$. For a given $\hat{u}$ one can always find a unit vector $\hat{u}'$ in the plane of $\hat{z}$ and $\hat{u}$ and such that[3]

$$\theta' = \frac{\pi}{2}(1 - \hat{u}\cdot\hat{z})\,. \tag{6.7}$$

The choice of $\hat{u}'$ is illustrated in Fig. 6.1. One sees immediately that $\xi(\hat{u};\lambda)$ specified as in (6.5) reproduces the statistical predictions of quantum mechanics.

In principle, hidden variables can be introduced no matter what state vector interpretation is adopted. Clearly, however, they are closer in spirit to the statistical and epistemic interpretations.

The purpose of hidden variables, we recall, is to provide a theory that allows more categorical predictions than quantum mechanics while reproducing, either exactly or very closely, the quantum-mechanical statistical predic-

---

[3] In general two vectors $\hat{u}'$ satisfy this requirement. Anyone can be picked, for instance the one whose angle with $\hat{u}$ is smaller.

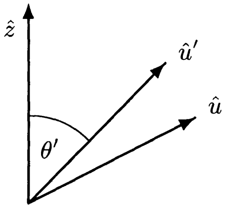

Fig. 6.1. Vectors $\hat{z}$, $\hat{u}$ and $\hat{u}'$ in the case where $\hat{u} \cdot \hat{z} = 1/2$

tions. In Chap. 21 we will see that this last requirement significantly restricts the way the value of a dynamical variable can depend on a hidden variable.

For the time being let us show that even if hidden variables would categorically determine measurement results of all dynamical variables, all these could not simultaneously have values that are well-defined and equal to eigenvalues of corresponding operators.[4] Consider a spin 3/2 particle. Dynamical variable $\boldsymbol{J} \cdot \boldsymbol{J} = J_x^2 + J_y^2 + J_z^2$ has value $j(j+1)\hbar^2 = (15/4)\hbar^2$. We know that the result of measuring J_x, J_y or J_z is necessarily equal to $\pm\hbar/2$ or $\pm 3\hbar/2$. Should hidden variables simultaneously determine J_x, J_y and J_z, values of J_x^2, J_y^2 and J_z^2 would also be determined, and equal to either $\hbar^2/4$ or $9\hbar^2/4$. But this cannot be, since 15/4 is not the sum of three numbers picked among 1/4 and 9/4.

Except in Chap. 21, we shall not mention the possibility of hidden variables in forthcoming chapters. For simplicity, we will most of the time use the terminology of the individual interpretation of the state vector. Most statements would easily translate into the language of the statistical or the epistemic interpretation. In general, the interpretation chosen has no influence on the results of practical calculations. Possible exceptions to this statement will however be pointed out in Chap. 21.

Epistemologists point out two different ways, called realist and instrumentalist, of viewing the status of physical theories. *Realists* assert that all mathematical symbols or abstract entities of a physical theory have their counterpart in the real world. *Instrumentalists*, on the other hand, grant reality to phenomena only (or even sensations), and maintain mathematical formalism is but a tool that allows to predict phenomena observed in given circumstances. For realists a physical theory is true or false. For instrumentalists it is adequate or inadequate. Between the two extremes there is of course a whole spectrum of intermediate attitudes. Bohr and Heisenberg adopted an instrumentalist view of quantum mechanics. Hidden variables, on the other hand, represent a realist attempt to reformulate quantum mechanics.

---

[4] Holland ([114], Sect. 3.5) associates with some dynamical variables values that do not coincide with eigenvalues of corresponding operators.

6.3 State Preparation and Measurement

The notions of measurement and preparation have been encountered several times. They are no doubt the two most important processes for setting up the correspondence between the quantum-mechanical mathematical formalism and the world of phenomena. These processes have points in common, and are sometimes mixed up. Nevertheless they are distinct [152], [153]. We will now investigate them more closely.

The Stern–Gerlach device, introduced in Sect. 4.1, helps to pinpoint their peculiarities. We recall that a beam of spin j particles going through the apparatus is separated into $2j+1$ secondary beams that correspond to different values $m\hbar$ of the spin projection.

Consider a configuration in which all beams can freely go through the apparatus. A detector is located at the end, so that one knows which beam a particle belongs to. In this case we say the apparatus performs a *measurement* of a particle's spin projection.

Suppose now a screen is introduced in the apparatus, that stops all beams except one, say the $\bar{m}$ beam. All particles coming out of the apparatus have then a spin projection equal to $\bar{m}\hbar$. In that case we say the apparatus *prepares* state $|\bar{m}\rangle$.

Hence measurement and state preparation are distinct. They are nevertheless linked by the following consistency condition. If a state $|\bar{m}\rangle$ is prepared, then an immediately following mesurement of dynamical variable J_z yields value $\bar{m}\hbar$. It is important to point out that this is not an experimental statement. It is a logical statement, part of a definition. That condition must be satisfied for processes to be called state preparation and measurement. Assume two identically oriented Stern–Gerlach devices are located one next to the other and that all beams in the first apparatus, except one, are stopped. Then one observes experimentally that the second apparatus only produces the corresponding beam. This, among other things, allows one to say that the first apparatus prepares a state while the second one measures a dynamical variable.

The notions of state preparation and measurement are very general and are met in many contexts other than Stern–Gerlach devices. In general, the process of state preparation can be characterized in the following way. Suppose that after some interactions, a device produces identical microobjets. We say this constitutes a *state preparation* if no subsequent procedure can separate the set S of microobjects in two subsets S_1 and S_2 such that (i) S_1 and S_2 have distinct statistical properties and (ii) $S_1 \cup S_2$ has the same statistical properties as S.[5]

---

[5] Subsets S_1 and S_2 have *distinct statistical properties* if there exists a procedure which, applied to microobjects of S_1 and S_2, yields in the limit of large numbers mean results that are different. The definition of state preparation will be clearer once we have introduced, in Chap. 12, the notion of statistical mixture.

We should point out that the process of state preparation is defined here in a purely operational way, whether quantum mechanics is valid or not. Assume, however, that quantum mechanics is valid and that it is the most complete theory of nature possible. Then one can see that a state preparation procedure yields an ensemble of microobjects described by a state vector $|\phi\rangle$. Such is the filtering, by means of a Stern–Gerlach device, of a beam associated with a specific value $\bar{m}$. A second differently-oriented Stern–Gerlach device would separate beam $\bar{m}$ in two or more statistically distinct beams. But the set of all particles coming out of the second apparatus would have statistical properties very different from the ones of the original beam.

By contrast, filtering by means of a Stern–Gerlach device two distinct beams associated with $\bar{m}_1$ and $\bar{m}_2$, and then merging the two beams in one, does not constitute a state preparation. Indeed a second identically-oriented device would separate the set S of particles in the full beam into two subsets S_1 and S_2 that would have distinct statistical properties. But subsequent merging of S_1 and S_2 would produce a beam with statistical properties identical with those of S.

Assume now that quantum mechanics is not the most complete theory of nature possible. In other words, suppose there exist hidden variables that are physically significant, that is, they can show up in experimental procedures. It is clear then that a state preparation yields an ensemble of microobjects described not only by a state vector $|\phi\rangle$, but also by a precise value of all physically significant hidden variables.

If quantum mechanics is a valid and complete theory, the state preparation procedure can be used to define the notion of state vector in an operational way [181]: a state vector $|\phi\rangle$ is an abbreviation for the description of a state preparation procedure. Note however that the correspondence is not one-to-one: two distinct procedures can prepare the same state.

Let us now turn to the notion of measurement. Let A be a dynamical variable associated with a microobject. We say an interaction between the microobject and a macroscopic apparatus is a measurement of A if the following three conditions are met:

i) To each eigenvalue a_i of operator A there correspond one or many values $\alpha(a_i)$ of the apparatus's pointer, so that $\alpha(a_i) \neq \alpha(a_j)$ if $a_i \neq a_j$.
ii) No matter what the microobject's initial state, the pointer shows one of the $\alpha(a_j)$ once the interaction is over.
iii) Whenever the microobject's initial state coincides with an eigenvector $|a_i\rangle$ of A, the pointer, following the interaction, necessarily shows $\alpha(a_i)$.

As an example, one easily checks that the interaction of a spin particle with a Stern–Gerlach device with detector is indeed a measurement in the sense of the above definition. The pointer of the apparatus corresponds to the height at which a particle is detected upon leaving the magnetic field.

In Chap. 21 we will investigate more closely the nature of the interaction between microobject and apparatus. At any rate, the experimenter has direct access to the pointer's values only. To what extent do they correspond to an objective property of the microobject?

To try to answer this question, let us again consider an apparatus built for the measurement of a dynamical variable A, whose eigenvectors are denoted by $|a_j\rangle$. For simplicity, we assume that eigenvalues of A are not degenerate. If the initial state $|\phi\rangle$ of the microobject coincides with a vector $|a_i\rangle$, then following interaction, the pointer of the apparatus shows value $\alpha(a_i)$. One can say uncontroversially that the apparatus has simply measured the eigenvalue of A that was corresponding to the initial state $|\phi\rangle$.

Suppose now, as is generally the case, that the microobject's initial state $|\phi\rangle$ is a linear combination of vectors $|a_j\rangle$, that is,

$$|\phi\rangle = \sum_j c_j |a_j\rangle \, .$$

Following interaction, the pointer shows $\alpha(a_i)$. What does this result mean?

If we admit the existence of hidden variables we can say, for instance, that the vector $|\phi\rangle$ represents the state of an ensemble of microobjects and that each microobject has, immediately before mesurement, a precise value of dynamical variable A. The apparatus simply indicates the eigenvalue of A that corresponds to the microobject it specifically detects.

Although this answer cannot be ruled out on strictly logical terms, it nevertheless raises a problem. The example involving spin 3/2 showed that all dynamical variables cannot simultaneously have well-defined values. If we want to maintain that the measured value corresponds to a value that existed before measurement, we must admit that the microobjects "knew", in a sense, what dynamical variable would be measured. To speak more impersonally, the value of the measured dynamical variable, rather than any other, should be made real by the hidden variables. It is difficult to see how the mere adjustment of an apparatus to measure this or that dynamical variable could influence a microobject before measurement. The more so since, as we shall see in Chap. 21, the microobject can be arbitrarily far.

Granted thus that the value shown by an apparatus does not in general correspond to an eigenvalue associated with the microobject before interaction, could it be that it corresponds to an eigenvalue associated with the microobject after measurement? Following Dirac ([69], Sect. 10) and von Neumann ([230], Chaps. 5 and 6), many answer in the affirmative.

Specifically, von Neumann postulates that if the apparatus shows value $\alpha(a_i)$, then the state of the microobject, immediately after interaction, coincides with $|a_i\rangle$. The transition from $|\phi\rangle = \sum_j c_j|a_j\rangle$ to $|a_i\rangle$ through the measurement process is called the *state vector collapse*.

Von Neumann's postulate (called the *projection postulate*) was introduced in the context of discussing the interaction process going on during measurement, to which we will come back in Chap. 21. It should be clear, however, that the projection postulate is not a logical consequence of the definition we gave of the measurement process. The postulate is even obviously incorrect if a measurement is made by destruction of a microobject, for instance when a photon's position is revealed through the absorption of a photon at a given point.[6]

But there are also many situations where a measurement of a dynamical variable is at the same time a state preparation, specifically, the state represented by the eigenvector $|a_i\rangle$ associated with the value $\alpha(a_i)$ shown by the apparatus. Such is the case with a Stern–Gerlach device where particles that come out are detected without being destroyed and without their spin being perturbed. To these situations von Neumann's postulate correctly applies.

To sum up, one cannot in general assert that the value shown by a measurement apparatus corresponds to an objective property of the microobject, either before or after interaction.

6.4 The Heisenberg Uncertainty Principle

Whenever two Hermitian operators do not commute, it is impossible to find a basis of eigenvectors of both. This, however, does not forbid that the two operators have several common eigenvectors.

Such is not the case, however, for the position and momentum operators of a particle in one dimension. Indeed we saw that in the coordinate representation, eigenvectors of X are proportional to $\delta(x - x_0)$, while eigenvectors of P are proportional to $\mathrm{e}^{\mathrm{i}px/\hbar}$. Operators X and P have no common eigenvectors, which means there are no states wherein dynamical variables position and momentum are both well-defined.

Could there be, nonetheless, states wherein X and P are defined as well as one wishes? Heisenberg answered this question negatively, in a precise sense. Let $|\psi\rangle$ be the normalized state vector of a particle in one dimension. Variances $(\Delta X)^2$ and $(\Delta P)^2$ of position and momentum are given by

$$\begin{aligned}(\Delta X)^2 &= \langle\psi|\left\{X - \langle X\rangle\right\}^2|\psi\rangle \\ &= \int_{-\infty}^{\infty} \mathrm{d}x\, \psi^*(x) \left\{X - \langle X\rangle\right\}^2 \psi(x)\,, \qquad (6.8)\\ (\Delta P)^2 &= \langle\psi|\left\{P - \langle P\rangle\right\}^2|\psi\rangle \\ &= \int_{-\infty}^{\infty} \mathrm{d}x\, \psi^*(x) \left\{-\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}x} - \langle P\rangle\right\}^2 \psi(x)\,. \qquad (6.9)\end{aligned}$$

Heisenberg showed that for any $|\psi\rangle$, one always has

$$(\Delta X)(\Delta P) \geq \frac{\hbar}{2}\,. \qquad (6.10)$$

---

[6] For a critique of the projection postulate see [13], [152] and [153].

Inequality (6.10) expresses the *Heisenberg uncertainty principle.*[7] Its operational meaning is clear. ΔX represents the standard deviation of a series of position measurements, made on a large number of identical and similarly prepared particles, all described by the state vector $|\psi\rangle$. Likewise, ΔP represents the standard deviation of a series of momentum measurements, made on another sample of particles also prepared in state $|\psi\rangle$. The Heisenberg uncertainty principle asserts that the product $(\Delta X)(\Delta P)$ has a lower bound equal to $\hbar/2$.

To establish inequality (6.10) we first prove the following lemma.

Lemma Let O_1 and O_2 be two Hermitian operators. Then

$$\langle O_1^2\rangle\langle O_2^2\rangle \geq \frac{1}{4}\left|\langle[O_1,O_2]\rangle\right|^2, \tag{6.11}$$

the mean values being computed in any normalized vector $|\psi\rangle$.

Proof Let us apply the Schwartz inequality[8] (2.6) to vectors $O_1|\psi\rangle$ and $O_2|\psi\rangle$. Because O_1 and O_2 are Hermitian, one has

$$\langle O_1^2\rangle\langle O_2^2\rangle = \langle\psi|O_1^\dagger O_1|\psi\rangle\langle\psi|O_2^\dagger O_2|\psi\rangle \geq \left|\langle\psi|O_1O_2|\psi\rangle\right|^2.$$

Letting

$$\begin{aligned} M &= \frac{1}{2}\,[O_1,O_2] = \frac{1}{2}(O_1O_2 - O_2O_1)\,, \\ N &= \frac{1}{2}(O_1O_2 + O_2O_1)\,, \end{aligned}$$

one sees that $M^\dagger = -M$ and $N^\dagger = N$, so that

$$\begin{aligned} \langle O_1^2\rangle\langle O_2^2\rangle &\geq \left|\langle\psi|O_1O_2|\psi\rangle\right|^2 \\ &\geq \left|\langle\psi|M+N|\psi\rangle\right|^2 \\ &\geq \left|\langle\psi|M|\psi\rangle + \langle\psi|N|\psi\rangle\right|^2 \\ &\geq \left|\langle\psi|M|\psi\rangle\right|^2 + \left|\langle\psi|N|\psi\rangle\right|^2 \\ &\qquad + \langle\psi|M|\psi\rangle\langle\psi|N|\psi\rangle^* + \langle\psi|M|\psi\rangle^*\langle\psi|N|\psi\rangle\,. \end{aligned}$$

The third and fourth terms in the last inequality cancel each other. The first and second terms being nonnegative, one finds that

$$\langle O_1^2\rangle\langle O_2^2\rangle \geq \left|\langle\psi|M|\psi\rangle\right|^2. \qquad ♣$$

---

[7] A detailed discussion of the Heisenberg uncertainty principle can be found in [124], Chap. 3.

[8] The proof of the Schwartz inequality, given in Chap. 2, remains valid in infinite-dimensional spaces, as long as the vectors are normalizable.

Let us now apply the lemma to the case where $O_1 = X - \langle X \rangle$ and $O_2 = P - \langle P \rangle$. One has

$$[O_1, O_2] = \mathrm{i}\hbar I, \qquad \langle O_1^2 \rangle = (\Delta X)^2, \qquad \langle O_2^2 \rangle = (\Delta P)^2.$$

Thus one gets

$$(\Delta X)^2 (\Delta P)^2 \geq \frac{1}{4} \left| \langle \psi | \mathrm{i}\hbar I | \psi \rangle \right|^2 = \frac{\hbar^2}{4},$$

from which the Heisenberg inequality immediately follows.

We note that the Heisenberg inequality makes no restrictions on ΔX or ΔP individually (except that $\Delta X > 0$, $\Delta P > 0$). One can, for example, find normalized vectors wherein ΔX is arbitrarily small. But then ΔP is necessarily very large, so as to satisfy (6.10).

The Heisenberg inequality can be illustrated by the phenomenon of photon diffraction (Fig. 6.2). Consider a source emitting monochromatic photons with wavelength λ, toward a screen with a slit of width a. Photons going through the slit are later absorbed by a photographic plate. Development of the plate reveals a diffraction pattern.

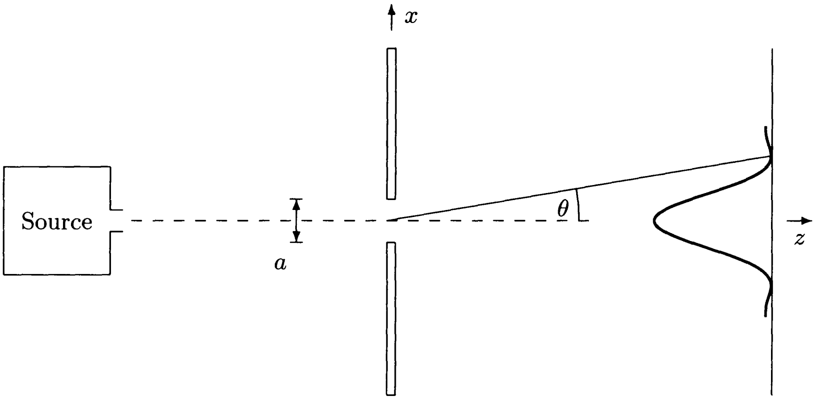

Fig. 6.2. Photon diffraction

One can say that the source and slit prepare photons in a state wherein the standard deviation of position ΔX is of order $a/2$. ΔX can thus be made arbitrarily small. But the smaller a is, the wider is the diffraction pattern. We know the momentum of a photon with wavelength λ is equal to h/λ. One can say the extent of the diffraction pattern is an assessment of the standard deviation ΔP of the x component of momentum, as

$$\frac{\Delta P}{(h/\lambda)} \approx \sin\theta \approx \theta. \tag{6.12}$$

But the angle θ of the first minimum of the diffraction pattern is equal to λ/a ([163], Sect. 12.8). Therefore

$$(\Delta X)(\Delta P) \approx \frac{a}{2}\frac{h}{\lambda}\frac{\lambda}{a} = \frac{h}{2}, \tag{6.13}$$

in agreement with Heisenberg's inequality.

As we pointed out, the operational meaning of the Heisenberg inequality refers to the standard deviations of a large number of measurements of X and P, made on similarly prepared quantum systems. Nevertheless the phrase "Heisenberg uncertainty principle" is often meant to include considerations about the impossibility of measuring the position of a microobject without disturbing its momentum, or vice versa. In the device shown in Fig. 6.2, for instance, photons coming from the source can be prepared with a ΔP (with respect to x) arbitrarily small. The passage of a photon through the slit and its subsequent detection are a measurement of the photon's x position. This measurement, whose uncertainty is equal to $a/2$, gives the photon a momentum of order h/a.

So is it when photons are used to measure the position of another kind of particles, say electrons. Suppose an electron is prepared in a state where its momentum is very well defined and close to zero. It is well-known in optics ([137], Sect. 58) that with photons with wavelength λ, one can determine an object's position only up to $\pm\lambda$. Of course, λ can be arbitrarily small. But photons with wavelength λ have momentum equal to h/λ. The photons' interaction with the electron perturbs the latter's momentum by a quantity of order h/λ. The product of the electron's position and momentum uncertainties thus satisfies the Heisenberg inequality.

It is important to realize that the empirical validity of the Heisenberg uncertainty principle ultimately depends on the universality of Planck's constant. Suppose a photon's energy were given by the relation $E = h'\nu$, with $h' \ll h$. It would then be possible, in a state where the electron's momentum is sufficiently well-defined, to measure the electron's position with an uncertainty of order λ, while perturbing its momentum by h'/λ only. In so doing a state would have been prepared that violates Heisenberg's inequality.

In closing we should point out that some investigators refuse to apply the Heisenberg uncertainty principle to individual particle. In the diffraction experiment, for instance, they maintain that a photon can have an arbitrarily well-defined momentum before and when going through the slit. Likewise, accurate determination of the point where a photon is absorbed allows reconstructing its momentum immediately after going through the slit with arbitrary precision [184]. A reply to that is that surely, a photon can be emitted by the source with an arbitrarily well-defined momentum. The process of going through the slit, however, involves a complex interaction between the photon and atoms of the screen, which perturbs the photon's momentum. Likewise determination, through the absorption point, of the momentum the photon supposedly had immediately after going through the slit, is problematic. This in fact assumes that the measurement of a dynamical variable reveals the value the variable had before the measurement process. Problems with this hypothesis were noted in Sect. 6.3 and will be examined again in Sect. 21.7.

6.5 Complementarity

The problem of the nature of microobjects was raised before the discovery of quantum mechanics, and contributed to its development. In Chap. 1 we pointed out the riddle of light's apparently contradictory properties. Like a wave, it gives rise to interference and diffraction phenomena. Like a corpuscle, it is absorbed discretely in the photoelectric effect. Eventually, this wave–particle duality was generalized to other microobjects, such as neutrons whose wave properties were revealed in Bragg scattering.

The quantum description of microobjects reproduces this duality, while making it more precise. In a state with well-defined position a microobject resembles a corpuscle. On the other hand, in a state with well-defined momentum (hence with ill-defined position) a microobject resembles a wave.

But in the end, is the photon a wave or a corpuscle? The correct answer is that it is neither. It is a microobject governed by the equations of quantum mechanics, with properties sometimes wavelike, sometimes particle-like.

From an operational point of view, we saw that the state vector can be considered as the symbol of a state preparation process. Ultimately, this process is carried out by means of macroscopic devices. That observation led Bohr to believe that the attribution of a property to a microobject is inseparably linked with the apparatus or macroscopic device used to prepare the microobject's state or to measure a dynamical variable pertaining to it. Whenever two dynamical variables do not commute, one cannot set at the same time the two experimental arrangements required to measure them. Both arrangements are nonetheless equally essential to a complete description of the microobject and are in this sense *complementary.*

Microobject and macroscopic apparatus are inseparable. That inseparability, together with the existence of the quantum of action h, is according to Bohr at the source of the statistical character of quantum-mechanical predictions.

To describe experimental results objectively and without ambiguity, one must use classical language. The descriptive inseparability of microobject and macroscopic apparatus comes from the requirement to communicate in plain language. According to Bohr, the description of phenomena provided by quantum mechanics is nevertheless complete. The indivisibility of the quantum of action and the requirements of objectivity and communicability entail that one cannot logically ask for more.

The interpretation of quantum mechanics proposed by Bohr and Heisenberg is called the *Copenhagen interpretation.*[9] Most contemporary physicists side with this interpretation, although by that they do not all mean the same thing. But a good source of information can be found in [217]. After a close look at Bohr's writings and after consulting with Heigenberg, H. P. Stapp gave what according to him are the two fundamental characteristics of the Copenhagen interpretation:

i) It is pragmatic [i.e. instrumentalist], in the sense that the formalism only predicts phenomena which, ultimately, are described in classical language.
ii) It states that quantum mechanics provides a complete description of atomic phenomena. Simultaneous specification of noncommuting dynamical variables is impossible because it can only be conceived in relation with macroscopic arrangements, and these are mutually exclusive.

6.6 Quantum Mechanics and Classical Mechanics

In the Copenhagen interpretation, the quantum description of phenomena necessarily involves classical concepts. State preparation and measurement processes, essential to the interpretation of the quantum-mechanical formalism, require macroscopic devices whose behavior is classical. But this is not the only meeting place between quantum mechanics and classical mechanics. The relation between the two theories is complex, profound and far from completely elucidated. We shall look at three different aspects of it: (i) formal analogies, (ii) the classical behavior of microobjects and (iii) the reduction of classical to quantum properties.

Formal analogies are particularly obvious and have had great importance in the historical development of quantum mechanics. The quantum Hamiltonian of a particle in one dimension, given in (5.15), has exactly the same form as the classical Hamiltonian. This is also the case with the Hamiltonian of a particle carrying magnetic moment, in a constant magnetic field (4.40). We will repeatedly observe the heuristic value of these analogies [68]. It is important, however, to stress their limits. Symbols appearing in quantum and classical Hamiltonians have different meanings. They stand for commuting real numbers in the classical case, and for noncommuting Hermitian operators in the quantum case. That being said, there is an important analogy between commutators in quantum mechanics and Poisson brackets in classical mechanics. It is developed, for example, in [93], Sect. 9.6 and [160], Sect. 8.11.

In several circumstances, especially when quantum numbers are large, the behavior of microobjects resembles classical behavior. For example let us look at a quantum harmonic oscillator of mass m and angular frequency ω. Suppose that at $t = t_0$ the oscillator's wave function is given by

---

[9] On the Copenhagen interpretation one can read [37] and [124], Chaps. 4 and 6.

$$\psi(x_0,t_0)=\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\exp\left\{-\frac{m\omega}{2\hbar}(x_0-\bar{x})^2\right\}. \tag{6.14}$$

This is a Gaussian function centered on $\bar{x}$, whose square is normalized and whose width is proportional to the square root of $\hbar/m\omega$. In Sect. 5.8 we saw that the wave packet evolves so that at t its absolute square is given by

$$|\psi(x,t)|^2=\left(\frac{m\omega}{\pi\hbar}\right)^{1/2}\exp\left\{-\frac{m\omega}{\hbar}\left[x-\bar{x}\cos[\omega(t-t_0)]\right]^2\right\}. \tag{6.15}$$

At time t the wave packet is therefore a Gaussian function centered on $\bar{x}\cos[\omega(t-t_0)]$. This means that the center of the wave packet oscillates like the center of mass of a classical harmonic oscillator. In the case where $\bar{\xi}=(m\omega/\hbar)^{1/2}\bar{x}\gg 1$, one can show (Exercise 6.2) that energies most involved in the wave packet gather around the quantum number $\bar{n}=\bar{\xi}^2/2$, which corresponds to the energy $\bar{E}=m\omega^2\bar{x}^2/2$ of a classical oscillator of mass m, angular frequency ω and amplitude $\bar{x}$.

The quantum problems of potential wells and barriers also show certain similarities with the corresponding classical problems. A quantum particle can cross a barrier higher than its energy, but with a probability that gets smaller as the barrier gets wider and higher. In the potential well problem shown in Fig. 5.1, a negative-energy quantum particle is not strictly confined inside the well, but the probability that it is found outside at a distance higher than l, say, goes to zero as l increases. If, on the other hand, the particle has positive energy, the probability that it is found near the well goes to zero, as in the case of a classical particle which then excapes to infinity.

One sometimes reads that classical mechanics is obtained from quantum mechanics by a limiting process in which Planck's constant h goes to zero. Commutators like $[X,P]$, $[X,H]$ and $[P,H]$ then vanish, which may suggest that symbols X, P and H no longer represent operators, but rather ordinary numbers. We recall, however, that in the coordinate representation P is essentially given by $-\mathrm{i}\hbar\partial/\partial x$, which implies that $P\to 0$ as $\hbar\to 0$. Hence the limiting process is not very meaningful. But let us look at the problem differently, and consider a quantum system where all natural variables with the dimensions of action are much larger than Planck's constant. Is the behavior of such a system close to classical?

Part of the answer to this question may come from the following argument. Consider a particle of mass m whose Hamiltonian is given by

$$H=\frac{1}{2m}P^2+V(X). \tag{6.16}$$

Let us compute the time derivative of the mean value, in state $|\psi(t)\rangle$, of operators X and P. One easily finds that

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle X\rangle=\frac{1}{\mathrm{i}\hbar}\langle[X,H]\rangle=\frac{1}{m}\langle P\rangle, \tag{6.17}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle P\rangle=\frac{1}{\mathrm{i}\hbar}\langle[P,H]\rangle=-\left\langle\frac{\mathrm{d}V}{\mathrm{d}X}\right\rangle. \tag{6.18}$$

These equations are known as *Ehrenfest's relations.* They entail that mean values of X, P and $\mathrm{d}V/\mathrm{d}X$ obey the equations of Newtonian dynamics. Ehrenfest's relations are valid for any value of m. If m is large and the potential does not change much on small or intermediate distances, one can make up a wave packet that remains spatially narrow and focussed in momentum over fairly large time intervals. Such a wave packet is well characterized by mean values $\langle X\rangle$ and $\langle P\rangle$, and therefore its behavior is close to classical.[10]

In spite of this, a classical object cannot genuinely be modeled by a particle in one, or even in three, dimensions. Such an object is made up of a very large number of particles. To appropriately establish the correspondence between quantum mechanics and classical mechanics, the real challenge is to apply the quantum-mechanical formalism to macroscopic objects, and to show that the formalism then predicts a behavior that virtually coincides with the classical behavior. This program has recently undergone important developments, which we will examine in Chap. 16.

Exercises

6.1. Derive (6.6).

6.2. Show that in the case where $\bar{\xi} = (m\omega/\hbar)^{1/2}\bar{x} \gg 1$, the energies most involved in the harmonic oscillator wave packet (6.14) gather around the quantum number $\bar{n} = \bar{\xi}^2/2$.

---

[10] A good discussion of Ehrenfest's relations can be found in [208], Chap. 6. We should note that, strictly speaking, Ehrenfest's relations coincide with Hamilton's equations only if $\langle \mathrm{d}V/\mathrm{d}X\rangle = (\mathrm{d}V/\mathrm{d}x)_{\langle X\rangle}$. This equation is true in the classical limit.

7 Particle in Three Dimensions

Many concepts introduced in the study of a particle in one dimension can be adapted directly to the particle in three dimensions. Angular momentum operators, however, are new. Closely linked with the particle in three dimensions, they are much like the spin operators we examined in Chap. 4. Angular momentun operators are particularly useful where the potential is spherically symmetric. They help in partially solving the problem of the Hamiltonian's eigenvalues and eigenfunctions. For the r^{-1} and the spherical well potential, a complete solution is obtained. Electrically charged, a particle in three dimensions will interact with an external electromagnetic field. That interaction is responsible for the existence of atoms and molecules.

7.1 State Space and Schrödinger's Equation

As a first approximation, the state space of a particle in three dimensions has a basis made up of kets $|\boldsymbol{r}\rangle$. They represent states wherein the particle has a well-defined position $\boldsymbol{r} = (x, y, z)$. Three Hermitian operators X, Y and Z correspond to dynamical variables x *position*, y *position* and z *position* and are collectively denoted by $\boldsymbol{R}$. Kets $|\boldsymbol{r}\rangle$ are common eigenvectors of X, Y and Z, so that

$$X|\boldsymbol{r}\rangle = x|\boldsymbol{r}\rangle , \qquad Y|\boldsymbol{r}\rangle = y|\boldsymbol{r}\rangle , \qquad Z|\boldsymbol{r}\rangle = z|\boldsymbol{r}\rangle .$$

Concisely

$$\boldsymbol{R}|\boldsymbol{r}\rangle = \boldsymbol{r}|\boldsymbol{r}\rangle . \tag{7.1}$$

Operators X, Y and Z pairwise commute, since they have a basis of common eigenvectors.

The normalization of kets $|\boldsymbol{r}\rangle$ is chosen so that

$$\langle \boldsymbol{r}|\boldsymbol{r}'\rangle = \delta(\boldsymbol{r}-\boldsymbol{r}') = \delta(x-x')\delta(y-y')\delta(z-z') . \tag{7.2}$$

The resolution of the identity is written as

$$\int \mathrm{d}\boldsymbol{r}\, |\boldsymbol{r}\rangle\langle \boldsymbol{r}| = I . \tag{7.3}$$

L. Marchildon, *Quantum Mechanics*

Unless otherwise indicated, integrals over d$\boldsymbol{r}$ will always be carried out on the whole space.

Three Hermitian operators P_x, P_y and P_z, collectively denoted by $\boldsymbol{P}$, correspond to dynamical variables x, y and z *momentum*. Their matrix elements in the $|\boldsymbol{r}\rangle$ basis are given by

$$\langle \boldsymbol{r}|\boldsymbol{P}|\boldsymbol{r}'\rangle = -\mathrm{i}\hbar\boldsymbol{\nabla}_{\boldsymbol{r}}\delta(\boldsymbol{r}-\boldsymbol{r}'). \tag{7.4}$$

Operators P_x, P_y and P_z pairwise commute. A given component of $\boldsymbol{P}$ commutes with a component of $\boldsymbol{R}$ along an orthogonal direction. For components along the same direction we have

$$[X, P_x] = [Y, P_y] = [Z, P_z] = \mathrm{i}\hbar I.$$

These commutation relations can be written concisely if one lets X_i ($i = 1$, 2, 3) denote components of $\boldsymbol{R}$ and P_i components of $\boldsymbol{P}$. Then

$$[X_i, X_j] = 0\,, \tag{7.5}$$
$$[P_i, P_j] = 0\,, \tag{7.6}$$
$$[X_i, P_j] = \mathrm{i}\hbar\delta_{ij} I\,. \tag{7.7}$$

Eigenvectors of momentum operators are denoted by $|\boldsymbol{p}\rangle$. One has

$$\boldsymbol{P}|\boldsymbol{p}\rangle = \boldsymbol{p}|\boldsymbol{p}\rangle\,, \tag{7.8}$$
$$\langle \boldsymbol{p}|\boldsymbol{p}'\rangle = \delta(\boldsymbol{p}-\boldsymbol{p}')\,, \tag{7.9}$$
$$\int \mathrm{d}\boldsymbol{p}\, |\boldsymbol{p}\rangle\langle \boldsymbol{p}| = I\,. \tag{7.10}$$

The relation between kets $|\boldsymbol{p}\rangle$ and $|\boldsymbol{r}\rangle$ is such that

$$\langle \boldsymbol{r}|\boldsymbol{p}\rangle = (2\pi\hbar)^{-3/2}\exp\left\{\frac{\mathrm{i}}{\hbar}\boldsymbol{p}\cdot\boldsymbol{r}\right\}. \tag{7.11}$$

The *Hamiltonian*, which corresponds to the dynamical variable energy, is given in terms of operators $\boldsymbol{P}$ and $\boldsymbol{R}$ as in classical mechanics. Thus for simple systems

$$H = \frac{1}{2m}\boldsymbol{P}\cdot\boldsymbol{P} + V(\boldsymbol{R}, t) = \frac{1}{2m}\left(P_x^2 + P_y^2 + P_z^2\right) + V(\boldsymbol{R}, t). \tag{7.12}$$

Here m is the mass of the particle (in the present case spinless and with no internal degrees of freedom) and V is a real function of the three operators X, Y, Z and of time.

At any time t, the particle's state is represented by a vector in the state space, denoted by $|\psi(t)\rangle$. The *wave function* $\psi(\boldsymbol{r}, t)$ is the component of $|\psi(t)\rangle$ on $|\boldsymbol{r}\rangle$, that is,

$$\psi(\boldsymbol{r},t) = \langle \boldsymbol{r}|\psi(t)\rangle . \tag{7.13}$$

For any $|\psi\rangle$ and $|\phi\rangle$

$$\langle\phi|\psi\rangle = \int \mathrm{d}\boldsymbol{r}\, \langle\phi|\boldsymbol{r}\rangle\langle\boldsymbol{r}|\psi\rangle = \int \mathrm{d}\boldsymbol{r}\, \phi^*(\boldsymbol{r})\psi(\boldsymbol{r}). \tag{7.14}$$

The action of operators $\boldsymbol{R}$, $\boldsymbol{P}$ and H on $|\psi(t)\rangle$ is conveniently expressed in terms of the wave function:

$$\langle\boldsymbol{r}|\boldsymbol{R}|\psi(t)\rangle = \boldsymbol{r}\psi(\boldsymbol{r},t)\,, \tag{7.15}$$

$$\langle\boldsymbol{r}|\boldsymbol{P}|\psi(t)\rangle = -\mathrm{i}\hbar\boldsymbol{\nabla}\psi(\boldsymbol{r},t)\,, \tag{7.16}$$

$$\langle\boldsymbol{r}|H|\psi(t)\rangle = -\frac{\hbar^2}{2m}\nabla^2\psi(\boldsymbol{r},t) + V(\boldsymbol{r},t)\psi(\boldsymbol{r},t)\,. \tag{7.17}$$

The absolute square $|\psi(\boldsymbol{r},t)|^2$ of a normalized wave function represents the probability density to find, at time t, the particle at point $\boldsymbol{r}$.

The time evolution of the wave function is given by the *Schrödinger equation*

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(\boldsymbol{r},t) = -\frac{\hbar^2}{2m}\nabla^2\psi(\boldsymbol{r},t)+V(\boldsymbol{r},t)\psi(\boldsymbol{r},t). \tag{7.18}$$

This equation determines the wave function at any time t from its value at t_0. The wave function $\psi(\boldsymbol{r},t)$ is a continuous function of $\boldsymbol{r}$ and t. Unless the potential has an infinite discontinuity, the gradient $\boldsymbol{\nabla}\psi(\boldsymbol{r},t)$ is also a continuous function of $\boldsymbol{r}$ and t.

If V does not depend on time, the Schrödinger equation has particular solutions of the form

$$\psi(\boldsymbol{r},t) = \psi_E(\boldsymbol{r})\exp\left\{-\frac{\mathrm{i}}{\hbar}Et\right\}, \tag{7.19}$$

where $\psi_E(\boldsymbol{r})$ satisfies the eigenvalue equation of the Hamiltonian, that is,

$$-\frac{\hbar^2}{2m}\nabla^2\psi_E(\boldsymbol{r})+V(\boldsymbol{r})\psi_E(\boldsymbol{r}) = E\psi_E(\boldsymbol{r}). \tag{7.20}$$

This equation has solutions for any value of E. From a physical point of view, however, solutions must respect the hermiticity condition on operators $\boldsymbol{R}$, $\boldsymbol{P}$ and H. This condition is satisfied if wave functions are restricted to periodic functions in x, y and z, with periods $2L$ (Exercise 7.1).[1] The limit $L \to \infty$ is taken in the end. Let

$$V_\infty = \lim_{|\boldsymbol{r}|\to\infty} V(\boldsymbol{r}). \tag{7.21}$$

Admissible solutions then fall into two categories:

---

[1] Wave functions that vanish on the faces of a cubic box can also be used.

i) For $E < V_\infty$ normalizable solutions only are admissible, for which

$$\int \mathrm{d}\boldsymbol{r}\, |\psi_E(\boldsymbol{r})|^2 < \infty . \tag{7.22}$$

Such solutions exist for certain discrete values of E only, denoted by E_i. One writes

$$\psi_{E_i}(\boldsymbol{r}) = \langle \boldsymbol{r} | E_i \rangle , \tag{7.23}$$

and one can pick kets $|E_i\rangle$ that are orthonormal, that is,

$$\langle E_i | E_j \rangle = \delta_{ij} . \tag{7.24}$$

A particle with state described by $|E_i\rangle$ has a well-defined energy equal to E_i. It is essentially confined to a finite region of space. We say it is in a *bound state.*

ii) For $E > V_\infty$ every solution is admissible (provided it satisfies the continuity conditions) and none is normalizable. Solutions can be labeled by a continuous index α, which can be multidimensional. One writes

$$\psi_{E_\alpha}(\boldsymbol{r}) = \langle \boldsymbol{r} | E_\alpha \rangle . \tag{7.25}$$

Kets $|E_\alpha\rangle$ can be made orthogonal and normalized by means of the delta function, defined in the present context by (5.121). Thus

$$\langle E_\alpha | E_\beta \rangle = \delta(\alpha - \beta) . \tag{7.26}$$

A particle with state described by $|E_\alpha\rangle$ has a vanishing probability to be found inside any bounded region.

As with a particle in one dimension (Sect. 5.3), the state space of the particle in three dimensions is built with kets $|E_i\rangle$ and $|E_\alpha\rangle$. One has

$$\sum_i |E_i\rangle\langle E_i| + \int \mathrm{d}\alpha\, |E_\alpha\rangle\langle E_\alpha| \equiv \mathsf{S}_E |E\rangle\langle E| = I , \tag{7.27}$$

$$\sum_i E_i |E_i\rangle\langle E_i| + \int \mathrm{d}\alpha\, E_\alpha |E_\alpha\rangle\langle E_\alpha| \equiv \mathsf{S}_E E |E\rangle\langle E| = H . \tag{7.28}$$

The general solution of Schrödinger's equation is given by[2]

$$\begin{aligned}\psi(\boldsymbol{r},t) &= \mathsf{S}_E C_E \psi_E(\boldsymbol{r}) \mathrm{e}^{-\mathrm{i}Et/\hbar} \\ &= \sum_i C_i \psi_{E_i}(\boldsymbol{r}) \mathrm{e}^{-\mathrm{i}E_i t/\hbar} + \int \mathrm{d}\alpha\, C_\alpha \psi_{E_\alpha}(\boldsymbol{r}) \mathrm{e}^{-\mathrm{i}E_\alpha t/\hbar} .\end{aligned} \tag{7.29}$$

---

[2] Recall that S_E denotes a sum (and an integral) on eigenvectors of H, and not only on eigenvalues. In other words, there is in the summation a term for every orthonormal eigenvector of H.

The state vector at time t is obtained from the state vector at t_0 by means of the *evolution operator*:

$$|\psi(t)\rangle = U(t,t_0)|\psi(t_0)\rangle = \left\{ \mathsf{S}_E \mathrm{e}^{-\mathrm{i}E(t-t_0)/\hbar}|E\rangle\langle E| \right\} |\psi(t_0)\rangle. \tag{7.30}$$

Matrix elements of the evolution operator in the $|\boldsymbol{r}\rangle$ basis, that is, the propagator, can be written as

$$U(\boldsymbol{r},t;\boldsymbol{r}_0,t_0) = \langle \boldsymbol{r}|U(t,t_0)|\boldsymbol{r}_0\rangle. \tag{7.31}$$

As an example, consider a free particle, with $V(\boldsymbol{r}) = 0$. The eigenvalue equation of H becomes

$$\nabla^2\psi_E(\boldsymbol{r}) + \frac{2mE}{\hbar^2}\psi_E(\boldsymbol{r}) = 0. \tag{7.32}$$

There are no normalizable solutions for $E < V_\infty = 0$. For $E > 0$ one has

$$\psi_{E_{\boldsymbol{k}}}(\boldsymbol{r}) \equiv \psi_{\boldsymbol{k}}(\boldsymbol{r}) = (2\pi)^{-3/2}\exp\{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}\}, \tag{7.33}$$

where

$$E_{\boldsymbol{k}} = \frac{\hbar^2\boldsymbol{k}\cdot\boldsymbol{k}}{2m}. \tag{7.34}$$

Here the index α is a vector $\boldsymbol{k}$ whose three real components go from $-\infty$ to $+\infty$. We note that

$$\langle \boldsymbol{k}|\boldsymbol{k}'\rangle = \delta(\boldsymbol{k}-\boldsymbol{k}'). \tag{7.35}$$

Each nonzero value of E is infinitely degenerate. Kets $|\boldsymbol{k}\rangle$ are, in fact, eigenvectors of $\boldsymbol{P}$ with eigenvalues $\hbar\boldsymbol{k}$. As with a particle in one dimension, one shows that the propagator is given by

$$U_{\text{free}}(\boldsymbol{r},t;\boldsymbol{r}_0,t_0) = \left[\frac{m}{2\pi\mathrm{i}\hbar(t-t_0)}\right]^{3/2} \exp\left[\frac{\mathrm{i}m|\boldsymbol{r}-\boldsymbol{r}_0|^2}{2\hbar(t-t_0)}\right]. \tag{7.36}$$

7.2 Probability Current

The absolute square $|\psi(\boldsymbol{r},t)|^2$ of a normalized wave function represents the probability density to find, at time t, the particle at point $\boldsymbol{r}$. If N independent particles are all described by the same normalizable wave function, one can set the norm so that

$$\int \mathrm{d}\boldsymbol{r}\,|\psi(\boldsymbol{r},t)|^2 = N. \tag{7.37}$$

The quantity $|\psi(\boldsymbol{r},t)|^2 \mathrm{d}\boldsymbol{r}$ then represents the average number of particles found, at time t, in the volume element $d\boldsymbol{r}$ around point $\boldsymbol{r}$. This interpretation also applies to a nonnormalizable wave function.

In many circumstances, in particular in the context of scattering processes, it is important to know not only the density, but also the flux of particles. To come to this, let us compute the time derivative of the absolute square of the wave function. From Schrödinger's equation,

$$\begin{aligned}\frac{\partial}{\partial t}|\psi(\boldsymbol{r},t)|^2 &= \left(\frac{\partial}{\partial t}\psi^*\right)\psi + \psi^*\frac{\partial}{\partial t}\psi \\ &= \frac{1}{-\mathrm{i}\hbar}\left\{-\frac{\hbar^2}{2m}\nabla^2\psi^* + V\psi^*\right\}\psi + \psi^*\frac{1}{\mathrm{i}\hbar}\left\{-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi\right\} \\ &= \frac{\mathrm{i}\hbar}{2m}\left\{\psi^*\nabla^2\psi - \left(\nabla^2\psi^*\right)\psi\right\} \\ &= \frac{\mathrm{i}\hbar}{2m}\boldsymbol{\nabla}\cdot\left\{\psi^*\boldsymbol{\nabla}\psi - \psi\boldsymbol{\nabla}\psi^*\right\}.\end{aligned}$$

Defining

$$\varrho(\boldsymbol{r},t) = |\psi(\boldsymbol{r},t)|^2\,, \tag{7.38}$$

$$\boldsymbol{j}(\boldsymbol{r},t) = \frac{\mathrm{i}\hbar}{2m}\left\{\psi\boldsymbol{\nabla}\psi^* - \psi^*\boldsymbol{\nabla}\psi\right\} = -\frac{\hbar}{m}\mathrm{Im}\left\{\psi\boldsymbol{\nabla}\psi^*\right\}, \tag{7.39}$$

we can write

$$\frac{\partial}{\partial t}\varrho(\boldsymbol{r},t) + \boldsymbol{\nabla}\cdot\boldsymbol{j}(\boldsymbol{r},t) = 0\,. \tag{7.40}$$

This is the so-called *continuity equation.* It shows up, in general, in situations where a quantity (here the number of particles) is conserved.

To interpret (7.40), let us integrate it over a volume V in three-dimensional space, bounded by a surface S, and let us apply the divergence theorem. We obtain

$$\frac{\partial}{\partial t}\int_V \mathrm{d}\boldsymbol{r}\,\varrho(\boldsymbol{r},t) = -\int_S \mathrm{d}\boldsymbol{S}\cdot\boldsymbol{j}(\boldsymbol{r},t)\,. \tag{7.41}$$

If the norm of $\psi(\boldsymbol{r},t)$ is equal to 1, the left-hand side represents the time variation of the probability of finding the particle in V. If the norm of $\psi(\boldsymbol{r},t)$ is determined by (7.37), the left-hand side represents the time variation of the average number of particles in V.

Assume first that V represents the whole space, understood as the limit of a cube of size $2L$ as $L \to \infty$. Then the surface integral of $\boldsymbol{j}$ vanishes, owing to the periodicity of $\psi(\boldsymbol{r},t)$ on the cube. Hence the time variation of the total number of particles (or of the total probability) vanishes, in other words the total number of particles is conserved.

Assume now that V represents part of space only. The variation of the number of particles in V is equal to minus the surface integral of $\boldsymbol{j}(\boldsymbol{r},t)$ on

S. The vector $\boldsymbol{j}(\boldsymbol{r},t)$ can therefore be interpreted as the flux of particles at point $\boldsymbol{r}$ at time t. In other words $\boldsymbol{j}\cdot \mathrm{d}\boldsymbol{S}$ represents the number of particles that cross, around point $\boldsymbol{r}$ and per unit time, the surface element $\mathrm{d}\boldsymbol{S}$. The minus sign indicates that $\boldsymbol{j}\cdot\mathrm{d}\boldsymbol{S}$ represents the flux of particles coming out of the volume. There results a decrease in the number of particles inside.[3]

The vector $\boldsymbol{j}(\boldsymbol{r},t)$ is called the *probability current.* It is used especially in scattering processes. For example, consider the one-dimensional step potential illustrated in Fig. 5.4 (p. 106):

$$V(x)=\begin{cases}0 & \text{if } x<0\,,\\ V_0>0 & \text{if } x>0\,.\end{cases} \tag{7.42}$$

For $E>V_0$, the eigenvalue equation of H has solutions of the form

$$\psi(x)=\begin{cases}A_-\mathrm{e}^{\mathrm{i}kx}+B_-\mathrm{e}^{-\mathrm{i}kx} & \text{if } x<0\,,\\ A_+\mathrm{e}^{\mathrm{i}k'x} & \text{if } x>0\,,\end{cases} \tag{7.43}$$

where

$$k=\left\{\frac{2mE}{\hbar^2}\right\}^{1/2}, \qquad k'=\left\{\frac{2m}{\hbar^2}(E-V_0)\right\}^{1/2}, \tag{7.44}$$

and where (Exercise 5.12)

$$A_-=\left(\frac{k+k'}{2k}\right)A_+, \qquad B_-=\left(\frac{k-k'}{2k}\right)A_+. \tag{7.45}$$

For the probability current one finds (Exercise 7.2):

$$\boldsymbol{j}(x)=\begin{cases}\frac{\hbar k}{m}\left\{|A_-|^2-|B_-|^2\right\}\hat{x} & \text{if } x<0\,,\\ \frac{\hbar k'}{m}|A_+|^2\hat{x} & \text{if } x>0\,.\end{cases} \tag{7.46}$$

On the left the current can be interpreted as the sum of an incoming flux, proportional to $|A_-|^2$, and of a reflected flux, proportional to $|B_-|^2$. On the right the current represents a transmitted flux.[4] One defines the *reflection coefficient* $\mathcal{R}$ as the ratio of reflected flux to incoming flux, that is,

$$\mathcal{R}=\frac{|\boldsymbol{j}_{\text{ref}}|}{|\boldsymbol{j}_{\text{inc}}|}=\frac{|B_-|^2}{|A_-|^2}=\frac{(k-k')^2}{(k+k')^2}. \tag{7.47}$$

---

[3] We should point out that the definition of $\boldsymbol{j}$ given in (7.39) is not compulsory. Indeed one could add to $\boldsymbol{j}$ the curl of an arbitrary vector, without the continuity equation being affected. The integral of $\boldsymbol{j}$ on an open surface would have a different value, while the integral on a close surface would remain unchanged. The association, with the probability current, of the right-hand side of (7.39) is an interpretation postulate.

[4] The present analysis justifies the discussion in Sect. 5.6.

One also defines the *transmission coefficient* $\mathcal{T}$ as the ratio of transmitted flux to incoming flux, that is,

$$\mathcal{T} = \frac{|\boldsymbol{j}_{\text{tr}}|}{|\boldsymbol{j}_{\text{inc}}|} = \frac{k'|A_+|^2}{k|A_-|^2} = \frac{k'}{k}\frac{(2k)^2}{(k+k')^2}. \tag{7.48}$$

Note that

$$\mathcal{T} + \mathcal{R} = 1, \tag{7.49}$$

which really follows from the continuity of the probability current. This equation means that the number of transmitted particles, together with the number of reflected particles, is equal to the number of incoming particles.

7.3 Angular Momentum

With the position and momentum operators, one can define three operators L_x, L_y and L_z as

$$L_x = YP_z - ZP_y, \tag{7.50}$$
$$L_y = ZP_x - XP_z, \tag{7.51}$$
$$L_z = XP_y - YP_x. \tag{7.52}$$

More concisely,

$$\boldsymbol{L} = \boldsymbol{R} \times \boldsymbol{P}. \tag{7.53}$$

One easily checks that L_x, L_y and L_z are Hermitian. They correspond to dynamical variables that are the components of *orbital angular momentum.*

To compute commutation relations of components of $\boldsymbol{L}$ with each other, or with components of $\boldsymbol{R}$ and $\boldsymbol{P}$, it is very useful to introduce the *Levi-Civita symbol* ε_{ijk} defined as

$$\varepsilon_{ijk} = \begin{cases} +1 & \text{if } (ijk) \text{ is an even permutation of } (123), \\ -1 & \text{if } (ijk) \text{ is an odd permutation of } (123), \\ 0 & \text{otherwise}. \end{cases} \tag{7.54}$$

Clearly ε_{ijk} is antisymmetric with respect to any permutation of two indices, that is,

$$\varepsilon_{ijk} = -\varepsilon_{jik} = \varepsilon_{jki} = -\varepsilon_{kji}. \tag{7.55}$$

One is easily convinced that

$$\sum_{k=1}^{3} \varepsilon_{ijk}\varepsilon_{lmk} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}. \tag{7.56}$$

Denote by L_i ($i = 1, 2, 3$) the three components of $\boldsymbol{L}$. Equations (7.50)–(7.52) can then be written in terms of the symbol ε_{ijk}. Thus

$$L_i = \sum_{j,k} \varepsilon_{ijk} X_j P_k \,, \tag{7.57}$$

where the summation goes from 1 to 3.

Let us now compute the commutator of two components of $\boldsymbol{L}$. We have

$$\begin{aligned}
[L_i, L_j] &= \left[\sum_{k,l} \varepsilon_{ikl} X_k P_l, \sum_{m,n} \varepsilon_{jmn} X_m P_n \right] \\
&= \sum_{k,l,m,n} \varepsilon_{ikl} \varepsilon_{jmn} \left\{ -\mathrm{i}\hbar \delta_{lm} X_k P_n + \mathrm{i}\hbar \delta_{kn} X_m P_l \right\} \\
&= \mathrm{i}\hbar \left\{ \sum_{l,m} \left(\sum_k \varepsilon_{ikl} \varepsilon_{jmk} \right) X_m P_l - \sum_{k,n} \left(\sum_l \varepsilon_{ikl} \varepsilon_{jln} \right) X_k P_n \right\} \\
&= \mathrm{i}\hbar \left\{ \sum_{l,m} \left(\delta_{lj} \delta_{im} - \delta_{lm} \delta_{ij} \right) X_m P_l - \sum_{k,n} \left(\delta_{in} \delta_{kj} - \delta_{ij} \delta_{kn} \right) X_k P_n \right\} \\
&= \mathrm{i}\hbar \left(X_i P_j - X_j P_i \right) \\
&= \mathrm{i}\hbar \sum_k \varepsilon_{ijk} L_k \,.
\end{aligned} \tag{7.58}$$

Explicitly, this means that

$$[L_x, L_y] = \mathrm{i}\hbar L_z \,, \tag{7.59}$$

$$[L_y, L_z] = \mathrm{i}\hbar L_x \,, \tag{7.60}$$

$$[L_z, L_x] = \mathrm{i}\hbar L_y \,. \tag{7.61}$$

These relations are familiar. We have just shown that components of orbital angular momentum satisfy the same commutation relations as spin components.

Commutation relations of $\boldsymbol{L}$ with $\boldsymbol{R}$ and $\boldsymbol{P}$ are easily calculated with the help of the Levi-Civita symbol. One finds

$$[L_i, X_j] = \mathrm{i}\hbar \sum_k \varepsilon_{ijk} X_k \,, \tag{7.62}$$

$$[L_i, P_j] = \mathrm{i}\hbar \sum_k \varepsilon_{ijk} P_k \,. \tag{7.63}$$

Making use of the antisymmetry of ε_{ijk}, one also obtains that

$$[L_i, \boldsymbol{R} \cdot \boldsymbol{R}] = \left[L_i, \sum_j X_j X_j \right] = \mathrm{i}\hbar \sum_{j,k} \varepsilon_{ijk} \left(X_j X_k + X_k X_j \right) = 0 \,, \tag{7.64}$$

$$[L_i, \boldsymbol{P} \cdot \boldsymbol{P}] = 0 \,. \tag{7.65}$$

This means that all components of $\boldsymbol{L}$ commute with $|\boldsymbol{R}|^2$ and $|\boldsymbol{P}|^2$, and therefore with any function of R and of P.

The action of components of $\boldsymbol{L}$ on a ket $|\psi\rangle$ can be written in terms of the wave function as

$$\langle \boldsymbol{r}|L_x|\psi\rangle = -\mathrm{i}\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)\psi(\boldsymbol{r})\,, \tag{7.66}$$

$$\langle \boldsymbol{r}|L_y|\psi\rangle = -\mathrm{i}\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)\psi(\boldsymbol{r})\,, \tag{7.67}$$

$$\langle \boldsymbol{r}|L_z|\psi\rangle = -\mathrm{i}\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\psi(\boldsymbol{r})\,. \tag{7.68}$$

In later calculations it will be very useful to have these matrix elements in spherical coordinates. Letting

$$x = r\sin\theta\cos\phi\,, \tag{7.69}$$

$$y = r\sin\theta\sin\phi\,, \tag{7.70}$$

$$z = r\cos\theta\,, \tag{7.71}$$

one finds that

$$\langle \boldsymbol{r}|L_x|\psi\rangle = \mathrm{i}\hbar\left(\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right)\psi(\boldsymbol{r})\,, \tag{7.72}$$

$$\langle \boldsymbol{r}|L_y|\psi\rangle = \mathrm{i}\hbar\left(-\cos\phi\frac{\partial}{\partial\theta} + \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right)\psi(\boldsymbol{r})\,, \tag{7.73}$$

$$\langle \boldsymbol{r}|L_z|\psi\rangle = -\mathrm{i}\hbar\frac{\partial}{\partial\phi}\psi(\boldsymbol{r})\,. \tag{7.74}$$

One also obtains

$$\langle \boldsymbol{r}|\boldsymbol{L}\cdot\boldsymbol{L}|\psi\rangle = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]\psi(\boldsymbol{r}). \tag{7.75}$$

In Chap. 4 we have built all finite-dimensional vector spaces where operators can be defined that obey the commutation relations of components of $\boldsymbol{L}$. In the present case the state space is infinite-dimensional. Of course $\boldsymbol{L}\cdot\boldsymbol{L}$ and L_z commute. Let us look for their common eigenvectors, and their eigenvalues:

$$L_z|\psi\rangle = \mu|\psi\rangle\,,$$

$$\boldsymbol{L}\cdot\boldsymbol{L}|\psi\rangle = \lambda|\psi\rangle\,.$$

In terms of wave functions and spherical coordinates, these equations become

$$-\mathrm{i}\hbar\frac{\partial}{\partial\phi}\psi(\boldsymbol{r}) = \mu\psi(\boldsymbol{r})\,, \tag{7.76}$$

$$-\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]\psi(\boldsymbol{r}) = \lambda\psi(\boldsymbol{r})\,. \tag{7.77}$$

Equation (7.76) is easily integrated to yield the ϕ dependence of the function ψ:

$$\psi(\boldsymbol{r}) = f(r,\theta)\exp\left\{\frac{\mathrm{i}\mu}{\hbar}\phi\right\}. \tag{7.78}$$

But $\psi(\boldsymbol{r})$ must be well-defined at every point $\boldsymbol{r}$. This implies that, as a function of the azimuthal angle ϕ, $\psi(\boldsymbol{r})$ is periodic with period 2π. Hence one must have

$$\frac{\psi(r,\theta,\phi+2\pi)}{\psi(r,\theta,\phi)} = \exp\left\{\frac{\mathrm{i}\mu}{\hbar}2\pi\right\} = 1. \tag{7.79}$$

This equation implies that μ is an integral multiple of $\hbar$. Thus eigenvalues of L_z are equal to $m\hbar$, where m is an integer.

Let us now substitute in (7.77) the expression we just found for $\psi(\boldsymbol{r})$. We easily find that

$$\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) - \frac{m^2}{\sin^2\theta} + \frac{\lambda}{\hbar^2}\right] f(r,\theta) = 0.$$

This equation simplifies if we let $w = \cos\theta$. Dropping for the moment the coordinate r, one gets

$$\frac{\mathrm{d}}{\mathrm{d}w}\left[(1-w^2)\frac{\mathrm{d}}{\mathrm{d}w}f(w)\right] + \left[\frac{\lambda}{\hbar^2} - \frac{m^2}{1-w^2}\right] f(w) = 0. \tag{7.80}$$

For any value of λ, this equation has two linearly independent solutions. But we require solutions that are everywhere finite and continuous in the interval $0 \le \theta \le \pi$, that is, for $-1 \le w \le 1$. One can show ([65], p. 371) that such is the case only if $\lambda/\hbar^2 = l(l+1)$, where l is an integer and $l \ge |m|$. Then one (and only one) solution is finite and continuous everywhere in the interval, hence acceptable. It is denoted by $P_l^m(w)$ and is called the *associated Legendre function.* Eigenvalues of $\boldsymbol{L}\cdot\boldsymbol{L}$ are thus equal to $l(l+1)\hbar^2$.

Let us sum up what we have found. Wave functions corresponding to eigenvectors of $\boldsymbol{L}\cdot\boldsymbol{L}$ and L_z have an r, θ and ϕ dependence given by a product like $u(r)P_l^m(\cos\theta)\mathrm{e}^{\mathrm{i}m\phi}$, where $u(r)$ is an arbitrary function and P_l^m is an associated Legendre function. It is useful to define the *spherical harmonic* $Y_{lm}(\theta,\phi)$ as a multiple of $P_l^m(\cos\theta)\mathrm{e}^{\mathrm{i}m\phi}$ chosen so that

$$\int_0^\pi \sin\theta\,\mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi\, Y_{lm}^*(\theta,\phi)Y_{lm}(\theta,\phi) = 1.$$

Precise definitions and properties of associated Legendre functions and spherical harmonics are summarized in Sect. 7.9.1.

Wave functions corresponding to eigenvectors of $\boldsymbol{L}\cdot\boldsymbol{L}$ and L_z can thus be written as $u(r)Y_{lm}(\theta,\phi)$, where $u(r)$ is an arbitrary function of r. The symbols $\boldsymbol{L} \cdot \boldsymbol{L}$ and L_z are often used to denote these same operators in the coordinate representation, that is,

$$\boldsymbol{L} \cdot \boldsymbol{L} = -\hbar^2 \left[\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial}{\partial\theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2} \right] , \tag{7.81}$$

$$L_z = -\mathrm{i}\hbar \frac{\partial}{\partial\phi} . \tag{7.82}$$

One can then write

$$\boldsymbol{L} \cdot \boldsymbol{L} Y_{lm}(\theta, \phi) = l(l+1)\hbar^2 Y_{lm}(\theta, \phi) , \tag{7.83}$$

$$L_z Y_{lm}(\theta, \phi) = m\hbar Y_{lm}(\theta, \phi) . \tag{7.84}$$

For any nonnegative integer l, spherical harmonics Y_{lm} (with m going from $-l$ to l) make up a basis of a $(2l+1)$-dimensional vector space. This space is identical with those we investigated in Chap. 4. The matrix representation of operators L_x and L_y, for instance, coincides with the matrix representation of operators S_x and S_y that we then defined. We should point out, however, that only spaces with integral l (by contrast with half-integral) can be realized in the state space of a particle in three dimensions.

7.4 Spherically Symmetric Potential

Let $V(\boldsymbol{R})$ be a time-independent potential. In general V can depend on operators X, Y and Z in an essentially arbitrary way. We say the potential is *spherically symmetric* if it depends on X, Y and Z only through the combination $R = (X^2 + Y^2 + Z^2)^{1/2}$. In this case the Hamiltonian is given by

$$H = \frac{1}{2m} \boldsymbol{P} \cdot \boldsymbol{P} + V(R) . \tag{7.85}$$

We saw in Sect. 7.3 that components of $\boldsymbol{L}$ commute with every function of R and of P. Thus for a spherically symmetric potential

$$[H, L_i] = 0 , \qquad i = 1, 2, 3 ; \tag{7.86}$$

$$[H, \boldsymbol{L} \cdot \boldsymbol{L}] = 0 . \tag{7.87}$$

Hence one can find a basis of common eigenvectors of H, L_z and $\boldsymbol{L} \cdot \boldsymbol{L}$.

Wave functions corresponding to common eigenvectors of L_z and $\boldsymbol{L} \cdot \boldsymbol{L}$ have the form $u(r)Y_{lm}(\theta, \phi)$, where $u(r)$ is an arbitrary function and Y_{lm} is a spherical harmonic. Such a wave function will satisfy the eigenvalue equation of H if

$$\left\{ -\frac{\hbar^2}{2m} \nabla^2 + V(r) \right\} u(r) Y_{lm}(\theta, \phi) = E u(r) Y_{lm}(\theta, \phi) . \tag{7.88}$$

In spherical coordinates the Laplacian is written as

$$\begin{aligned}\nabla^2 &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \\ &= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{1}{r^2}\frac{\boldsymbol{L}\cdot\boldsymbol{L}}{\hbar^2}\,. \end{aligned} \tag{7.89}$$

Substituting this in (7.88) and making use of (7.83), one easily obtains that

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left\{r^2\frac{\mathrm{d}}{\mathrm{d}r}u(r)\right\} - \left\{\frac{l(l+1)}{r^2} + \frac{2m}{\hbar^2}\left[V(r) - E\right]\right\}u(r) = 0. \tag{7.90}$$

This equation for the function $u(r)$ is called the *radial equation* and $u(r)$ is called the *radial function.* With no loss of generality one can take $u(r)$ real.

For any value of E, the radial equation has two linearly independent solutions. Suppose that, for small r, the potential $V(r)$ is no more singular than r^{-1}. Near $r = 0$ the two linearly independent solutions of the radial equation then behave as r^l and $r^{-(l+1)}$, respectively.[5] The solution involving $r^{-(l+1)}$ diverges at $r = 0$ and is therefore not physically acceptable. Hence physically acceptable solutions of the radial equation behave as r^l near $r = 0$.

In the radial equation the term proportional to $l(l+1)$ vanishes if $l = 0$. If $l \neq 0$, it can be interpreted as an additional term in the potential $V(r)$. Define an *effective potential* as

$$V_{\text{eff}}(r) = V(r) + \frac{\hbar^2}{2m}\frac{l(l+1)}{r^2}\,. \tag{7.91}$$

We see at once that a potential $V(r)$ associated with a nonvanishing value of l corresponds to the same eigenvalues E as a potential $V_{\text{eff}}(r)$ associated with a null value of the spherical harmonic's index l.

For $E < V_\infty$, the radial equation has normalized solutions for certain values of E only. These eigenvalues and the corresponding eigenfunctions depend on l. For each value of l there are in general many solutions, usually denoted by $u_{nl}(r)$. Corresponding eigenvalues are denoted by E_{nl}.[6] Thus in the discrete spectrum common eigenvectors of H, L_z and $\boldsymbol{L}\cdot\boldsymbol{L}$ can be written as $|nlm\rangle$, where

$$\langle \boldsymbol{r}|nlm\rangle \equiv \psi_{nlm}(\boldsymbol{r}) = u_{nl}(r)Y_{lm}(\theta,\phi) \tag{7.92}$$

and

---

[5] This is obvious if $l \neq 0$ or if V is regular at the origin. If $l = 0$ and if $V(r)$ has a term going like r^{-1}, the result follows from differential equations theory. See, for example, [65], Sect. 4.14.

[6] The index n is related to the number of nodes (i.e. zeros) of the radial function. If energies are ordered increasingly, the function $u_{n+1,l}(r)$ has one more node than the function $u_{nl}(r)$.

$$\begin{aligned} H|nlm\rangle &= E_{nl}|nlm\rangle\,, && (7.93)\\ \boldsymbol{L}\cdot\boldsymbol{L}|nlm\rangle &= \hbar^2 l(l+1)|nlm\rangle\,, && (7.94)\\ L_z|nlm\rangle &= m\hbar|nlm\rangle\,. && (7.95)\end{aligned}$$

The kets $|nlm\rangle$ can always be taken orthonormal, so that

$$\langle n'l'm'|nlm\rangle = \delta_{nn'}\delta_{mm'}\delta_{ll'}\,. \tag{7.96}$$

Making use of orthogonality properties of spherical harmonics, one easily checks that this condition entails that

$$\int_0^\infty \mathrm{d}r\, r^2 |u_{nl}(r)|^2 = 1\,. \tag{7.97}$$

For a spherically symmetric potential, operators H, $\boldsymbol{L}\cdot\boldsymbol{L}$ and L_z make up a complete set of commuting operators. We stress that eigenvalues E_{nl} of H depend on n and l, but not on m. The origin of this degeneracy will be elucidated in Chaps. 13 and 14.

7.5 $r^{\ 1}$ Potential

The radial equation (7.90) does not, in general, have solutions in terms of simple functions. In Chap. 8 we will see how, with the help of a computer, one can find approximations to eigenvalues E_{nl} and radial functions $u_{nl}(r)$.

For now we will investigate the radial equation in the case where $V(r)$ is proportional to r^{-1}. This is a situation where one can find bound-state eigenvalues E_{nl} exactly, as well as corresponding functions $u_{nl}(r)$.[7] The r^{-1} potential represents the electrostatic interaction of two point charges. In Chap. 9 we will see how the solutions we will presently develop describe a hydrogen atom, as well as any one-electron atomic ion.

Let us write

$$V(r) = -\frac{Ze^2}{r}\,, \tag{7.98}$$

where Z and e^2 are positive constants.[8] This potential is shown in Fig. 7.1. Since $V_\infty = 0$, bound-state energies must be negative. Equation (7.90) becomes

---

[7] Eigenvalues and eigenfunctions associated with the r^{-1} potential can also be obtained with the use of parabolic coordinates. See, for example, [138], Sect. 37 and [200], Sect. 16. Furthermore, eigenfunctions associated with the Hamiltonian's continuous spectrum can also be obtained exactly. See [138], Sect. 135, [160], Sect. 11.7 and [200], Sect. 21.

[8] This notation anticipates Chap. 9, where Z will denote atomic number and e^2 will stand for $q_\mathrm{e}^2/4\pi\varepsilon_0$.

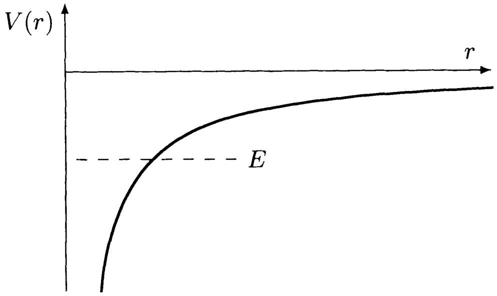

Fig. 7.1. r^{-1} potential

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left\{r^2\frac{\mathrm{d}u}{\mathrm{d}r}\right\}-\left[\frac{l(l+1)}{r^2}-\frac{2m}{\hbar^2}\left\{\frac{Ze^2}{r}+E\right\}\right]u=0. \tag{7.99}$$

It is useful to replace r and E by dimensionless variables ζ and ε defined as

$$\zeta=\sqrt{\frac{-8mE}{\hbar^2}}r, \qquad \varepsilon=\frac{Ze^2}{\hbar}\sqrt{\frac{-m}{2E}}. \tag{7.100}$$

One gets

$$\frac{1}{\zeta^2}\frac{\mathrm{d}}{\mathrm{d}\zeta}\left\{\zeta^2\frac{\mathrm{d}u}{\mathrm{d}\zeta}\right\}+\left\{\frac{\varepsilon}{\zeta}-\frac{1}{4}-\frac{l(l+1)}{\zeta^2}\right\}u=0. \tag{7.101}$$

Let us also write

$$u(\zeta)=f(\zeta)\mathrm{e}^{-\zeta/2}. \tag{7.102}$$

One then obtains the following equation for $f(\zeta)$ (Exercise 7.8):

$$\frac{\mathrm{d}^2f}{\mathrm{d}\zeta^2}+\left(\frac{2}{\zeta}-1\right)\frac{\mathrm{d}f}{\mathrm{d}\zeta}+\left\{\frac{\varepsilon-1}{\zeta}-\frac{l(l+1)}{\zeta^2}\right\}f=0. \tag{7.103}$$

Our problem is to find solutions of this equation that yield normalizable wave functions, i.e. such that

$$\int_0^\infty \mathrm{d}r\, r^2|u(r)|^2=\left(\frac{-8mE}{\hbar^2}\right)^{-3/2}\int_0^\infty \mathrm{d}\zeta\,\zeta^2\mathrm{e}^{-\zeta}|f(\zeta)|^2<\infty. \tag{7.104}$$

Let us look for a power series solution. We write

$$f(\zeta)=\zeta^s\sum_{k=0}^{\infty}c_k\zeta^k, \tag{7.105}$$

where $c_0 \neq 0$ and $s \geq 0$. Substituting this expression in (7.103) and setting to zero the coefficients of each power of ζ, one finds

$$c_0 \left\{ s(s+1) - l(l+1) \right\} = 0 \,,$$
$$c_{k+1} \left\{ (s+k+1)(s+k+2) - l(l+1) \right\} = c_k (s+k+1-\varepsilon) \,.$$

Since $c_0 \neq 0$, either $s = l$ or $s = -l - 1$. This result agrees with what we found in Sect. 7.4, about the behavior of the radial function around $r = 0$. As we pointed out then, the second option produces a divergence of $f(\zeta)$ at $\zeta = 0$, and is therefore excluded. Hence $s = l$. The recurrence equation thus becomes

$$c_{k+1} = \left\{ \frac{l+k+1-\varepsilon}{(k+1)(k+2l+2)} \right\} c_k \,. \tag{7.106}$$

For large k we have $c_{k+1}/c_k \approx 1/k$, which corresponds to the series of $\zeta^\nu e^\zeta$ for finite ν. The normalizability condition is then violated. This argument does not work, however, if for a value of k (say $k = n'$)

$$l + n' + 1 - \varepsilon = 0 \,. \tag{7.107}$$

In this case the series stops with the term $c_{n'}\zeta^{n'}$, and one finds

$$f(\zeta) = \zeta^l \sum_{k=0}^{n'} c_k \zeta^k \,, \tag{7.108}$$

where

$$c_{k+1} = \left\{ \frac{k-n'}{(k+1)(k+2l+2)} \right\} c_k \,. \tag{7.109}$$

Letting $n = n' + l + 1$ one has $n = \varepsilon$, which implies the following energy quantization condition:

$$E = -\frac{mZ^2e^4}{2\hbar^2 n^2} \,. \tag{7.110}$$

We recall that $0 \leq l < \infty$. Moreover it is clear that $0 \leq n' < \infty$. Therefore one has $l + 1 \leq n < \infty$.

For a specific choice of c_0, the polynomial that multiplies ζ^l in the expression of $f(\zeta)$ is denoted by $L_{n-l-1}^{2l+1}(\zeta)$ and is called a *Laguerre polynomial.* These polynomials are defined in Sect. 7.9.2. Orthonormal solutions for the r^{-1} potential can be written as[9]

$$\psi_{nlm}(\boldsymbol{r}) = \langle \boldsymbol{r} | nlm \rangle = u_{nl}(r) Y_{lm}(\theta, \phi) \,, \tag{7.111}$$

where

---

[9] The functions $\psi_{nlm}(\boldsymbol{r})$ are orthogonal with respect to indices l and m due to spherical harmonics. They are orthogonal with respect to n because different values of n correspond to different eigenvalues of the Hamiltonian.

$$u_{nl}(r) = \left[\left(\frac{2Z}{na_0}\right)^3 \frac{(n-l-1)!}{2n[(n+l)!]}\right]^{1/2} \mathrm{e}^{-\zeta/2}\zeta^l L_{n-l-1}^{2l+1}(\zeta), \tag{7.112}$$

$$a_0 = \frac{\hbar^2}{me^2}, \qquad \zeta = \frac{2Z}{na_0}r. \tag{7.113}$$

The value of the normalization constant is obtained in Sect. 7.9.2. Note that E depends only on index n, not on l or m. We call n the *total* (or *principal*) *quantum number*, l the *orbital* (or *azimuthal*) *quantum number* and m the *magnetic quantum number*. Radial functions u_{10}, u_{20} and u_{21} (multiplied by r) are shown in Fig. 7.4 (p. 156).

7.6 Spherical Square Well

The *spherical square well*, shown in Fig. 7.2, is the spherically symmetric potential defined by the following equation:

$$V(r) = \begin{cases} -V_0 & \text{if } r < a\,, \\ 0 & \text{if } r > a\,. \end{cases} \tag{7.114}$$

Here V_0 and a are positive constants. As with the r^{-1} potential, this problem has exact solutions. We shall obtain the discrete spectrum energies and corresponding wave functions.

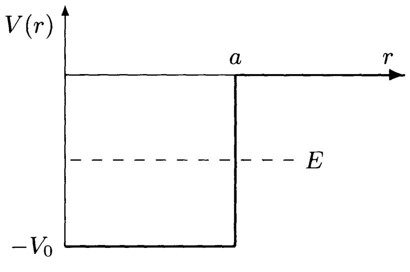

Fig. 7.2. Spherical square well

The radial equation (7.90) can be written as

$$\frac{\mathrm{d}^2u}{\mathrm{d}r^2} + \frac{2}{r}\frac{\mathrm{d}u}{\mathrm{d}r} + \left\{\frac{2m}{\hbar^2}\left[E - V(r)\right] - \frac{l(l+1)}{r^2}\right\}u(r) = 0. \tag{7.115}$$

In any region where the potential is constant, the solutions of this equation are *spherical Bessel* and *Neumann functions.* Properties of these functions are summarized in Sect. 7.9.3.

Let us first examine the region where $r < a$. Here $E - V(r) = E + V_0 > 0$. Letting

$$k = \left\{ \frac{2m}{\hbar^2}(E + V_0) \right\}^{1/2} , \tag{7.116}$$

one easily sees, from the results of Sect. 7.9.3, that the general solution of (7.115) is a linear combination of functions $j_l(kr)$ and $n_l(kr)$. For small r, $n_l(kr)$ behaves as r^{-l-1}. Hence the radial function around the origin will only involve $j_l(kr)$, that is,

$$u(r) = Aj_l(kr) \quad \text{if} \quad r < a . \tag{7.117}$$

In the region where $r > a$ we have $E - V(r) = E < 0$. Here it is useful to define

$$\kappa = \left\{ -\frac{2mE}{\hbar^2} \right\}^{1/2} . \tag{7.118}$$

The general solution of (7.115) is again a linear combination of spherical Bessel and Neumann functions, this time with argument $\mathrm{i}\kappa r$. In this case it is useful to pick as linearly independent solutions the functions $h_l^+(\mathrm{i}\kappa r)$ and $h_l^-(\mathrm{i}\kappa r)$. But $h_l^-(\mathrm{i}\kappa r)$ diverges as $r \to \infty$. For the wave function to be normalizable, $u(r)$ must involve $h_l^+(\mathrm{i}\kappa r)$ only where r is large, that is,

$$u(r) = A' h_l^+(\mathrm{i}\kappa r) \quad \text{if} \quad r > a . \tag{7.119}$$

To obtain bound-state energies, we have to apply continuity conditions on $u(r)$ and its derivative at $r = a$. Denoting by j_l' and $(h_l^+)'$ the derivatives of j_l and h_l^+ with respect to their argument, one easily sees that these conditions translate as

$$Aj_l(ka) = A' h_l^+(\mathrm{i}\kappa a) , \tag{7.120}$$

$$k h_l^+(\mathrm{i}\kappa a) j_l'(ka) = \mathrm{i}\kappa j_l(ka)(h_l^+)'(\mathrm{i}\kappa a) . \tag{7.121}$$

From (7.120) the constant A' can be expressed in terms of A. Equation (7.121) is the energy quantization relation. It can be solved numerically or graphically. The solutions (for the unknown E) are discrete-spectrum energies.

One can show that, for a given l, the number of bound states in the spherical square well increases as V_0 and a get larger. For V_0 and a fixed, on the other hand, the number of bound states decreases as l increases. This comes from the repulsive characters of the r^{-2} term in the effective potential.

We point out that the method used in this section can in principle be applied to any sectionally constant spherically symmetric potential. As the number of sections becomes appreciable, however, it is more convenient to use the numerical methods that will be developed in Chap. 8.

7.7 Electromagnetic Field

A classical electromagnetic field is entirely specified by an *electric field* vector $\boldsymbol{E}$ and a *magnetic induction* vector $\boldsymbol{B}$. Sources of the field are a *charge density* $\varrho(\boldsymbol{r},t)$ and a *current density* $\boldsymbol{j}(\boldsymbol{r},t)$. The relation between sources and fields is given by *Maxwell's equations*, which in SI units are written as[10]

$$\nabla\cdot\boldsymbol{E}=\frac{\varrho}{\varepsilon_0}\,,\qquad \nabla\times\boldsymbol{E}+\frac{\partial\boldsymbol{B}}{\partial t}=0\,, \tag{7.122}$$

$$\nabla\cdot\boldsymbol{B}=0\,,\qquad \nabla\times\boldsymbol{B}-\frac{1}{c^2}\frac{\partial\boldsymbol{E}}{\partial t}=\mu_0\boldsymbol{j}\,. \tag{7.123}$$

The force that an electromagnetic field exerts on a classical particle of charge q that moves with velocity $\boldsymbol{v}$ is given by

$$\boldsymbol{F}=q\,(\boldsymbol{E}+\boldsymbol{v}\times\boldsymbol{B})\,. \tag{7.124}$$

It is always possible to find a scalar function ϕ and a vector function $\boldsymbol{A}$, called the *scalar potential* and the *vector potential*, such that

$$\boldsymbol{B}=\nabla\times\boldsymbol{A},\qquad \boldsymbol{E}=-\nabla\phi-\frac{\partial\boldsymbol{A}}{\partial t}. \tag{7.125}$$

In terms of potentials Maxwell's equations are written as

$$\nabla^2\phi+\frac{\partial}{\partial t}(\nabla\cdot\boldsymbol{A})=-\frac{\varrho}{\varepsilon_0}\,, \tag{7.126}$$

$$\nabla^2\boldsymbol{A}-\frac{1}{c^2}\frac{\partial^2\boldsymbol{A}}{\partial t^2}-\nabla(\nabla\cdot\boldsymbol{A})-\frac{1}{c^2}\nabla\left(\frac{\partial\phi}{\partial t}\right)=-\mu_0\boldsymbol{j}\,. \tag{7.127}$$

Knowledge of potentials ϕ and $\boldsymbol{A}$ completely determines the fields $\boldsymbol{E}$ and $\boldsymbol{B}$. But the converse is not true. Indeed let $\chi(\boldsymbol{r},t)$ be an arbitrary function and let

$$\phi'=\phi-\frac{\partial\chi}{\partial t},\qquad \boldsymbol{A}'=\boldsymbol{A}+\nabla\chi. \tag{7.128}$$

One easily checks that $(\phi,\boldsymbol{A})$ and $(\phi',\boldsymbol{A}')$ determine the same fields $\boldsymbol{E}$ and $\boldsymbol{B}$. Equations (7.128) are called a *gauge transformation.* By *gauge invariance* we mean the fact that $\boldsymbol{E}$ and $\boldsymbol{B}$ do not depend on the function χ.

Let $\boldsymbol{A}_0$ be a constant vector, let ω and η be two constants and let $\boldsymbol{q}$ be a constant vector such that $c^2\boldsymbol{q}\cdot\boldsymbol{q}=\omega^2$ and $\boldsymbol{q}\cdot\boldsymbol{A}_0=0$. If $\varrho=0=\boldsymbol{j}$, (7.126) and (7.127) have the following solutions:

---

[10] These are the microscopic Maxwell's equations. They are true in all circumstances, provided that ϱ and $\boldsymbol{j}$ represent the total charge and current densities. Properties of Maxwell's equations are developed in electromagnetic treatises, for example [123] and [137].

$$\phi = 0, \qquad \boldsymbol{A} = 2\boldsymbol{A}_0 \cos(\boldsymbol{q}\cdot\boldsymbol{r} - \omega t + \eta). \tag{7.129}$$

These solutions represent linearly polarized electromagnetic waves with angular frequency ω.

We now examine the problem of a quantum particle in three dimensions, with mass m and charge q, in an electromagnetic field. First we must write down the Hamiltonian. If the particle were classical, its Hamiltonian would be given by ([93], p. 346; [137], p. 46)

$$H_{\text{cl}} = \frac{1}{2m}\left[\boldsymbol{p} - q\boldsymbol{A}\right]^2 + q\phi. \tag{7.130}$$

It is understood that $\boldsymbol{A}$ and ϕ are functions of $\boldsymbol{r}$ and t, and that $\boldsymbol{p}$ is the canonical momentum corresponding to $\boldsymbol{r}$. The canonical momentum is related to the velocity vector by

$$\boldsymbol{p} = m\boldsymbol{v} + q\boldsymbol{A}. \tag{7.131}$$

The quantum Hamiltonian of a particle in an electromagnetic field is given by the same expression as the classical Hamiltonian, provided the canonical momentum is replaced by the momentum operator $\boldsymbol{P}$ and the position vector is replaced by the position operator $\boldsymbol{R}$. Thus[11]

$$H = \frac{1}{2m}\left[\boldsymbol{P} - q\boldsymbol{A}(\boldsymbol{R}, t)\right]^2 + q\phi(\boldsymbol{R}, t). \tag{7.132}$$

The *Schrödinger equation*, given by

$$\mathrm{i}\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle,$$

is written in the coordinate representation as

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi(\boldsymbol{r}, t) = \frac{1}{2m}\left[-\mathrm{i}\hbar\boldsymbol{\nabla} - q\boldsymbol{A}(\boldsymbol{r}, t)\right]^2\psi(\boldsymbol{r}, t) + q\phi(\boldsymbol{r}, t)\psi(\boldsymbol{r}, t). \tag{7.133}$$

Equation (7.133) involves the potentials ϕ and $\boldsymbol{A}$ rather than the fields $\boldsymbol{E}$ and $\boldsymbol{B}$. At first sight one could conclude that the dynamics of a quantum particle, unlike the one of a classical particle, depends on the gauge. In Sect. 7.8, however, we shall see that such is not the case.

In the Hamiltonian (7.132) let us expand the square of the expression in square brackets. We find

$$H = \frac{1}{2m}\boldsymbol{P}\cdot\boldsymbol{P} - \frac{q}{2m}\left\{\boldsymbol{P}\cdot\boldsymbol{A} + \boldsymbol{A}\cdot\boldsymbol{P}\right\} + \frac{q^2}{2m}\boldsymbol{A}\cdot\boldsymbol{A} + q\phi.$$

But

---

[11] The electromagnetic field remains classical, because all its components pairwise commute.

$$\boldsymbol{P}\cdot\boldsymbol{A}=\boldsymbol{A}\cdot\boldsymbol{P}+\sum_i\left[P_i,A_i\right],$$

so that

$$H=\frac{1}{2m}\boldsymbol{P}\cdot\boldsymbol{P}-\frac{q}{m}\boldsymbol{A}\cdot\boldsymbol{P}-\frac{q}{2m}\sum_i\left[P_i,A_i\right]+\frac{q^2}{2m}\boldsymbol{A}\cdot\boldsymbol{A}+q\phi. \tag{7.134}$$

In Chap. 19 this expression will be used to compute the emission and absorption of electromagnetic radiation by atoms.

Here we will examine the form of the Hamiltonian in two particular cases: (i) an electrostatic field and (ii) a constant (time and space-independent) magnetic field. For an electrostatic field one sets $\boldsymbol{A}=0$ and $\phi=\phi(\boldsymbol{r})$. The Hamiltonian is thus simply given by

$$H=\frac{1}{2m}\boldsymbol{P}\cdot\boldsymbol{P}+q\phi\left(\boldsymbol{R}\right). \tag{7.135}$$

For a constant magnetic field one can set $\phi=0$ and $\boldsymbol{A}=-(\boldsymbol{R}\times\boldsymbol{B})/2$. Hence

$$\begin{aligned}H=&\frac{1}{2m}\boldsymbol{P}\cdot\boldsymbol{P}+\frac{q}{2m}\left(\boldsymbol{R}\times\boldsymbol{B}\right)\cdot\boldsymbol{P}\\&+\frac{q}{4m}\sum_i\left[P_i,\left(\boldsymbol{R}\times\boldsymbol{B}\right)_i\right]+\frac{q^2}{8m}\left(\boldsymbol{R}\times\boldsymbol{B}\right)^2.\end{aligned}$$

It is easy to see that the third term on the right-hand side vanishes. Furthermore,

$$(\boldsymbol{R}\times\boldsymbol{B})\cdot\boldsymbol{P}=-\left(\boldsymbol{B}\times\boldsymbol{R}\right)\cdot\boldsymbol{P}=-\left(\boldsymbol{R}\times\boldsymbol{P}\right)\cdot\boldsymbol{B}=-\boldsymbol{L}\cdot\boldsymbol{B}.$$

Thus

$$H=\frac{1}{2m}\boldsymbol{P}\cdot\boldsymbol{P}-\frac{q}{2m}\boldsymbol{L}\cdot\boldsymbol{B}+\frac{q^2}{8m}\left(\boldsymbol{R}\times\boldsymbol{B}\right)^2. \tag{7.136}$$

The second term on the right-hand side of (7.136) means that a particle with charge q, mass m and angular momentum $\boldsymbol{L}$ has a magnetic moment given by

$$\boldsymbol{\mu}=\frac{q}{2m}\boldsymbol{L}. \tag{7.137}$$

Here again, one sees the similarity with the classical formula ([40], p. 660; [123], p. 187).

In the Hamiltonian (7.136) let us compute the relative importance of the linear and quadratic terms in B. Let q be the charge of an electron and let $|\psi\rangle$ represent a state with spatial extension of order 1 Å and angular momentum of order $\hbar$. Then approximately

$$\frac{\langle\psi|\frac{q^2}{8m}(\boldsymbol{R}\times\boldsymbol{B})^2|\psi\rangle}{\langle\psi|\frac{q}{2m}(\boldsymbol{L}\cdot\boldsymbol{B})|\psi\rangle}\approx 10^{-5}B/\mathrm{T}.$$

This means that unless the field is very strong, the quadratic term is much smaller than the linear term.

7.8 Gauge Transformation

The scalar and the vector potential directly appear in the Schrödinger equation associated with a particle in an electromagnetic field. In this section we will investigate the question of their physical meaning.

One finds experimentally that the behavior of a microscopic system immersed in an electromagnetic field does not depend on the choice of gauge. Let us first see how quantum mechanics accounts for this observation.

It is not difficult to show that the Schrödinger equation displays a gauge invariance analogous to the one of Maxwell's equations. Let $\chi(\boldsymbol{r},t)$ be an arbitrary function and let us transform ϕ and $\boldsymbol{A}$ by means of (7.128). One then checks (Exercise 7.13) that (7.133) does not change, provided the wave function $\psi(\boldsymbol{r},t)$ transforms into a function $\psi'(\boldsymbol{r},t)$ such that

$$\psi'(\boldsymbol{r},t) = \exp\left\{\frac{\mathrm{i}q}{\hbar}\chi(\boldsymbol{r},t)\right\}\psi(\boldsymbol{r},t). \tag{7.138}$$

The wave function thus changes under a gauge transformation. This is not disturbing, since the wave function is not directly observable. We saw in Sect. 3.2 that the quantities which can be measured are mean values $\langle\psi|O|\psi\rangle$, where O is the Hermitian operator associated with a dynamical variable. Therefore, it is necessary that these mean values do not depend on the gauge, that is,

$$\langle\psi'|O'|\psi'\rangle = \langle\psi|O|\psi\rangle\,, \tag{7.139}$$

where $|\psi'\rangle$ and O' are obtained from $|\psi\rangle$ and O by a gauge transformation. Clearly, not all Hermitian operators satisfy (7.139). Hence in the presence of an electromagnetic field, not all Hermitian operators correspond to dynamical variables.

In the coordinate representation, it is not difficult to check that operator $\boldsymbol{R}$ satisfies (7.139) while operators $\boldsymbol{P}$ and $\boldsymbol{L}$ do not. However, operators

$$\boldsymbol{\mathcal{P}} = \boldsymbol{P} - q\boldsymbol{A} \quad \text{and} \quad \boldsymbol{\mathcal{L}} = \boldsymbol{R}\times\boldsymbol{\mathcal{P}} \tag{7.140}$$

do obey (7.139). Hence $\boldsymbol{R}$, $\boldsymbol{\mathcal{P}}$ and $\boldsymbol{\mathcal{L}}$ correspond to dynamical variables whereas $\boldsymbol{P}$ and $\boldsymbol{L}$, strictly speaking, do not. Whenever the electromagnetic field vanishes, though, one can pick a gauge such that $\boldsymbol{A} = 0 = \phi$. In that gauge, $\boldsymbol{\mathcal{P}}$ coincides with $\boldsymbol{P}$ and $\boldsymbol{\mathcal{L}}$ coincides with $\boldsymbol{L}$.

One easily checks that the commutation relations of X_i and P_j coincide with those of X_i and $\mathcal{P}_j$. Furthermore, one shows in Exercise (7.14) that components of $\mathcal{P}_i$ pairwise commute if and only if the magnetic induction vanishes. Commutation relations of components $\mathcal{L}_i$ coincide with those of components L_i if and only if $\boldsymbol{B} = 0$.

One can show (Exercise 7.15) that in the presence of an electromagnetic field, mean values $\langle\boldsymbol{R}\rangle$ and $\langle\boldsymbol{\mathcal{P}}\rangle$ satisfy

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \boldsymbol{R}\rangle = \frac{1}{m}\langle \boldsymbol{\mathcal{P}}\rangle\,, \tag{7.141}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \boldsymbol{\mathcal{P}}\rangle = q\left\langle \boldsymbol{E} + \frac{1}{2m}\left\{\boldsymbol{\mathcal{P}}\times\boldsymbol{B} - \boldsymbol{B}\times\boldsymbol{\mathcal{P}}\right\}\right\rangle. \tag{7.142}$$

Ehrenfest's relations are thus properly written in terms of operators $\boldsymbol{\mathcal{P}}$ rather than in terms of operators $\boldsymbol{P}$.

At first sight, it may be surprizing to realize that the Hamiltonian does not satisfy (7.139). Does this mean that a quantum system's energy is not a genuine dynamical variable? To answer this question, let us consider a general Hamiltonian, which explicitly depends on time. Such dependence means that the system's dynamics is controlled by time-dependent agents. This in turn entails continuous energy exchange between the system and its environment, so that no clear distinction can be established between the energy of the system and that of the environment. In this sense, the system's energy is not a dynamical variable if the hamiltonian is time dependent. By contrast, consider a quantum system immersed in a time-independent electromagnetic field. Such a field can always be obtained from time-independent potentials. In this case the Hamiltonian does not depend on time, and one can define the system's energy. The most general gauge transformation that transforms time-independent potentials into time-independent potentials is a transformation generated by a function $\chi(\boldsymbol{r},t) = f(\boldsymbol{r}) + Ct$, where C depends neither on $\boldsymbol{r}$ nor on t. Up to an additive constant (which, for energy, has no physical meaning), the Hamiltonian satisfies (7.139) under restriction to such gauge transformations. In this sense, we shall say that a time-independent Hamiltonian corresponds to a dynamical variable.

Quantum-mechanical predictions, like the ones in classical mechanics, do not depend on the gauge [134]. Nevertheless, quantum-mechanical electromagnetic potentials have a significance that they do not have in classical mechanics.

Consider a situation where, in some region of space, the fields $\boldsymbol{E}$ and $\boldsymbol{B}$ vanish. One can pick $\phi = 0$, and then $\boldsymbol{A} = \boldsymbol{A}(\boldsymbol{r})$. Since $\boldsymbol{B}$ vanishes, $\boldsymbol{\nabla}\times\boldsymbol{A} = 0$. Hence there exists a function $\chi(\boldsymbol{r})$ such that $\boldsymbol{A} = \boldsymbol{\nabla}\chi$. The Schrödinger equation can be written as

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\psi = \frac{1}{2m}\left[-\mathrm{i}\hbar\boldsymbol{\nabla} - q\boldsymbol{\nabla}\chi\right]^2\psi. \tag{7.143}$$

If $\psi_0(\boldsymbol{r},t)$ is a solution of (7.143) for $\boldsymbol{A} = 0$, then

$$\psi(\boldsymbol{r},t) = \exp\left\{\frac{\mathrm{i}q}{\hbar}\chi(\boldsymbol{r})\right\}\psi_0(\boldsymbol{r},t) \tag{7.144}$$

is a solution of (7.143) for $\boldsymbol{A} \neq 0$.

Let us now turn to the experimental device shown in Fig. 7.3 [2], [24]. A beam of electrons splits at point P. Electrons can reach P' through two

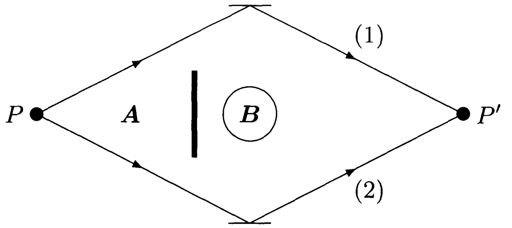

Fig. 7.3. Electron interference and vector potential

trajectories (1) and (2). In the middle, a thin long solenoid creates a magnetic field essentially confined within the coil.[12] A screen prevents electrons to come close to the solenoid. Thus electrons are always in a region where magnetic induction vanishes. The vector potential, however, does not vanish and outside the solenoid it is given by ([40], p. 457)

$$\boldsymbol{A} = \frac{\Phi}{2\pi\varrho}\hat{\phi}, \tag{7.145}$$

where Φ is the magnetic flux through the solenoid, ϱ is the distance from the axis of the solenoid and $\hat{\phi}$ is a unit vector going counterclockwise around the axis.

The condition of destructive interference at P' is given by

$$\psi^{(1)}(P') + \psi^{(2)}(P') = 0. \tag{7.146}$$

Here $\psi^{(1)}(P')$ and $\psi^{(2)}(P')$ are wave functions, evaluated at P', of an electron that follows trajectory (1) or (2), respectively. With the field turned off, the condition of destructive interference is given by

$$\psi_0^{(1)}(P') + \psi_0^{(2)}(P') = 0. \tag{7.147}$$

Along trajectory (1) the vector potential is the gradient of a function $\chi_1(\boldsymbol{r})$ given by

$$\chi_1(\boldsymbol{r}) = \int_{(1)} \boldsymbol{A}\cdot \mathrm{d}\boldsymbol{r}. \tag{7.148}$$

Similarly, along trajectory (2) $\boldsymbol{A}$ is the gradient of

$$\chi_2(\boldsymbol{r}) = \int_{(2)} \boldsymbol{A}\cdot \mathrm{d}\boldsymbol{r}. \tag{7.149}$$

These relations do not hold in the middle, where $\boldsymbol{B}$ does not vanish. At any rate, one easily obtains from (7.145) that

---

[12] We assume the length of the solenoid is much larger than the distance between P and P', so that field lines close on themselves far from electron trajectories.

$$\chi_1(P') = -\frac{\Phi}{2}, \qquad \chi_2(P') = \frac{\Phi}{2}. \tag{7.150}$$

Wave functions $\psi^{(1)}(P')$ and $\psi^{(2)}(P')$ are related to functions $\psi_0^{(1)}(P')$ and $\psi_0^{(2)}(P')$ through (7.144), where χ is equal to χ_1 or χ_2, respectively. Thus,

$$\psi^{(1)}(P') = \exp\left\{-\frac{\mathrm{i}q\Phi}{2\hbar}\right\}\psi_0^{(1)}(P'), \qquad \psi^{(2)}(P') = \exp\left\{\frac{\mathrm{i}q\Phi}{2\hbar}\right\}\psi_0^{(2)}(P'). \tag{7.151}$$

Condition (7.146) is therefore equivalent to

$$\psi_0^{(1)}(P') + \exp\left\{\frac{\mathrm{i}q\Phi}{\hbar}\right\}\psi_0^{(2)}(P') = 0, \tag{7.152}$$

and it does not, in general, coincide with condition (7.147). Hence the presence of a field inside the solenoid affects the electron's interference pattern.

The phenomenon just described is called the *Aharonov–Bohm effect* (or the *Ehrenberg–Siday–Aharonov–Bohm effect*). The vector potential influences an electron even in a region where the electromagnetic field vanishes. This is a purely quantum-mechanical effect, with no classical analogue.

7.9 Appendix

7.9.1 Legendre Polynomials and Spherical Harmonics

For any s and any w such that $|s^2 - 2sw| < 1$, the function

$$F(s, w) = \left(1 - 2sw + s^2\right)^{-1/2} \tag{7.153}$$

is analytic in s and in w. Expanded in a series of powers of s, this function can be written as

$$\left(1 - 2sw + s^2\right)^{-1/2} = \sum_{l=0}^{\infty} P_l(w) s^l, \tag{7.154}$$

where $P_l(w)$ is a polynomial of degree l in w, called a *Legendre polynomial*.[13] $F(s, w)$ is the *generating function* of Legendre polynomials. They can also be obtained from *Rodrigues's formula*:

$$P_l(w) = \frac{1}{2^l l!}\frac{\mathrm{d}^l}{\mathrm{d}w^l}(w^2 - 1)^l. \tag{7.155}$$

---

[13] The various functions introduced in this subsection are discussed in more detail in [65], Sects. 3.10 and 4.31, [143], Chaps. 4 and 7 and [236], Chap. 15.

Letting $w = 1$ in (7.154) one immediately sees that

$$P_l(1) = 1\,. \tag{7.156}$$

The first four Legendre polynomials are given by

$$\begin{aligned} P_0 &= 1\,, & P_1 &= w\,, \\ P_2 &= \frac{1}{2}(3w^2 - 1)\,, & P_3 &= \frac{1}{2}(5w^3 - 3w)\,. \end{aligned} \tag{7.157}$$

Upon differentiation of (7.154) with respect to s or with respect to w, one gets with some manipulations the following relations:

$$(l+1)P_{l+1} = (2l+1)wP_l - lP_{l-1}\,, \tag{7.158}$$

$$\left(1 - w^2\right)\frac{\mathrm{d}P_l}{\mathrm{d}w} = -lwP_l + lP_{l-1}\,. \tag{7.159}$$

One easily finds that

$$\frac{\mathrm{d}}{\mathrm{d}w}\left[\left(1 - w^2\right)\frac{\mathrm{d}P_l}{\mathrm{d}w}\right] + l(l+1)P_l = 0. \tag{7.160}$$

From that equation one can prove the following orthogonality relation obeyed by Legendre polynomials:

$$\int_{-1}^{1} \mathrm{d}w\, P_l(w)P_{l'}(w) = \frac{2\delta_{ll'}}{2l+1}\,. \tag{7.161}$$

For any integer m such that $0 \le m \le l$, the *associated Legendre function*, denoted by $P_l^m(w)$, is defined from $P_l(w)$ as

$$P_l^m(w) = \left(1 - w^2\right)^{m/2}\frac{\mathrm{d}^m}{\mathrm{d}w^m}P_l(w). \tag{7.162}$$

Clearly $P_l^m(w)$ is finite for any finite w. One can show that $P_l^m(w)$ satisfies the equation

$$\frac{\mathrm{d}}{\mathrm{d}w}\left[\left(1 - w^2\right)\frac{\mathrm{d}}{\mathrm{d}w}P_l^m\right] + \left[l(l+1) - \frac{m^2}{1 - w^2}\right]P_l^m = 0. \tag{7.163}$$

Furthermore, one can define $P_l^m(w)$ for negative m by means of the following formula:

$$P_l^{-m}(w) = (-1)^m\frac{(l-m)!}{(l+m)!}P_l^m(w). \tag{7.164}$$

Associated Legendre functions obey orthogonality relations with respect to the index l:

$$\int_{-1}^{1} \mathrm{d}w\, P_l^m(w) P_{l'}^m(w) = \frac{2\delta_{ll'}}{2l+1} \frac{(l+m)!}{(l-m)!} . \tag{7.165}$$

One defines *spherical harmonics* $Y_{lm}(\theta, \phi)$ in the following way:[14]

$$Y_{lm}(\theta, \phi) = (-1)^m \left[\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!} \right]^{1/2} P_l^m(\cos\theta) \mathrm{e}^{\mathrm{i}m\phi} . \tag{7.166}$$

From the orthogonality relation of associated functions one easily shows that

$$\int_0^{\pi} \sin\theta \, \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\phi \, Y_{lm}^*(\theta, \phi) Y_{l'm'}(\theta, \phi) = \delta_{ll'} \delta_{mm'} . \tag{7.167}$$

Furthermore

$$Y_{lm}^*(\theta, \phi) = (-1)^m Y_{l-m}(\theta, \phi) , \tag{7.168}$$

$$Y_{lm}(\pi - \theta, \phi + \pi) = (-1)^l Y_{lm}(\theta, \phi) , \tag{7.169}$$

$$Y_{l0}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi}} P_l(\cos\theta) . \tag{7.170}$$

One also checks ([73], Sect. 2.5) that the following relations are satisfied:

$$L_{\pm} Y_{lm}(\theta, \phi) = \hbar \{(l \mp m)(l + 1 \pm m)\}^{1/2} Y_{l\, m\pm 1}(\theta, \phi) . \tag{7.171}$$

Spherical harmonics for $l \leq 2$ are listed in Table 7.1.

Table 7.1. Spherical harmonics for $l \leq 2$

l, m	$Y_{lm}(\theta, \phi)$	l, m	$Y_{lm}(\theta, \phi)$
0, 0	$\sqrt{\frac{1}{4\pi}}$	2, 0	$\sqrt{\frac{5}{16\pi}} \left(3\cos^2\theta - 1\right)$
1, 0	$\sqrt{\frac{3}{4\pi}} \cos\theta$	2, ±1	$\mp\sqrt{\frac{15}{8\pi}} \sin\theta \cos\theta \mathrm{e}^{\pm \mathrm{i}\phi}$
1, ±1	$\mp\sqrt{\frac{3}{8\pi}} \sin\theta \mathrm{e}^{\pm \mathrm{i}\phi}$	2, ±2	$\sqrt{\frac{15}{32\pi}} \sin^2\theta \mathrm{e}^{\pm 2\mathrm{i}\phi}$

Let $\boldsymbol{r}_1 = (r_1, \theta_1, \phi_1)$, $\boldsymbol{r}_2 = (r_2, \theta_2, \phi_2)$ and let γ be the angle between $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$, so that

$$\cos\gamma = \cos\theta_1 \cos\theta_2 + \sin\theta_1 \sin\theta_2 \cos(\phi_1 - \phi_2) . \tag{7.172}$$

---

[14] This is the definition of Cowan [59], Edmonds [73] and Schiff [200]. Talman [223] adds a factor $[4\pi/(2l+1)]^{1/2}$ while Dennery and Krzywicki [65] add a factor $(-1)^m$ to their definition of associated polynomials (and spherical harmonics) but suppress it in their list of spherical harmonics. In general, different authors agree on the definition of Legendre polynomials but often disagree on associated functions.

From the generating function of Legendre polynomials one checks that

$$\frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} = \begin{cases} \frac{1}{r_1} \sum_{l=0}^{\infty} \left(\frac{r_2}{r_1}\right)^l P_l(\cos\gamma) & \text{if } r_1 > r_2 , \\ \frac{1}{r_2} \sum_{l=0}^{\infty} \left(\frac{r_1}{r_2}\right)^l P_l(\cos\gamma) & \text{if } r_2 > r_1 . \end{cases} \tag{7.173}$$

Moreover, we will show in Exercise (14.6) that

$$P_l(\cos\gamma) = \frac{4\pi}{2l+1} \sum_{m=-l}^{l} Y_{lm}^*(\theta_1, \phi_1) Y_{lm}(\theta_2, \phi_2). \tag{7.174}$$

In the case where $\theta_2 = \theta_1$ and $\phi_2 = \phi_1$, this relation and (7.156) imply that

$$\sum_{m=-l}^{l} Y_{lm}^*(\theta_1, \phi_1) Y_{lm}(\theta_1, \phi_1) = \frac{2l+1}{4\pi}. \tag{7.175}$$

Let n be a nonnegative integer and let $\alpha > -1$ and $\beta > -1$. The *Jacobi polynomial* $P_n^{(\alpha,\beta)}(w)$ is defined as[15]

$$\begin{aligned} P_n^{(\alpha,\beta)}(w) &= \frac{(-1)^n}{2^n n!}(1-w)^{-\alpha}(1+w)^{-\beta} \frac{\mathrm{d}^n}{\mathrm{d}w^n} \left\{(1-w)^{\alpha+n}(1+w)^{\beta+n}\right\} \\ &= \frac{1}{2^n} \sum_{l=0}^{n} \frac{\Gamma(n+\alpha+1)}{l!\,\Gamma(n+\alpha+1-l)} \\ &\quad \times \frac{\Gamma(n+\beta+1)}{(n-l)!\,\Gamma(\beta+l+1)} (w-1)^{n-l}(w+1)^l , \end{aligned} \tag{7.176}$$

where Γ stands for Euler's gamma function. Legendre polynomials and associated Legendre functions can be expressed in terms of Jacobi polynomials as

$$P_l(w) = P_l^{(0,0)}(w) , \tag{7.177}$$

$$P_l^m(w) = \frac{(l+m)!}{2^m l!} \left(1-w^2\right)^{m/2} P_{l-m}^{(m,m)}(w) , \qquad 0 \le m \le l . \tag{7.178}$$

7.9.2 Laguerre Polynomials

Let $\alpha > 1$. For any ζ and any complex s such that $|s| < 1$, the function

$$w(s,\zeta) = (1-s)^{-\alpha-1} \exp\left\{-s\zeta(1-s)^{-1}\right\} \tag{7.179}$$

is analytic in ζ and in s. Expanded in a series of powers of s, this function can be written as

---

[15] In Maple V the polynomial $P_n^{(a,b)}(w)$ is obtained through the command `P(n,a,b,w)` in the `orthopoly` package. In Mathematica, it is obtained through the command `JacobiP[n,a,b,w]`.

$$w(s,\zeta) = \sum_{\nu=0}^{\infty} L_\nu^\alpha(\zeta) s^\nu . \tag{7.180}$$

One can show that the coefficient $L_\nu^\alpha(\zeta)$ is a polynomial of degree ν in ζ, called a *Laguerre polynomial.*[16] The function $w(s,\zeta)$ is the *generating function* of Laguerre polynomials. These are given explicitly by

$$L_\nu^\alpha(\zeta) = \sum_{k=0}^{\nu} \frac{\Gamma(\nu+\alpha+1)(-\zeta)^k}{\Gamma(k+\alpha+1)k!(\nu-k)!} . \tag{7.181}$$

The ratio of the coefficient of ζ^{k+1} to the one of ζ^k in the previous expression is equal to

$$\frac{k-\nu}{(k+\alpha+1)(k+1)} .$$

This coincides with the quotient c_{k+1}/c_k of coefficients in (7.106), provided that $\nu = n - l - 1$ and $\alpha = 2l + 1$.

With the help of the generating function it is not difficult to show that

$$\int_0^\infty \mathrm{d}\zeta\, \zeta^{\alpha+1} \mathrm{e}^{-\zeta} L_\nu^\alpha(\zeta) L_\nu^\alpha(\zeta) = \frac{1}{\nu!}\Gamma(\nu+\alpha+1)[2\nu+\alpha+1], \tag{7.182}$$

from which one obtains the normalization constant in (7.112).

Radial functions $u_{nl}(r)$ associated with the r^{-1} potential, found in Sect. 7.5, can be written in terms of Laguerre polynomials. We recall, however, that the relation between r and ζ depends on n. We find that

$$u_{10}(r) = \left(\frac{Z}{a_0}\right)^{3/2} 2\mathrm{e}^{-Zr/a_0} , \tag{7.183}$$

$$u_{20}(r) = \left(\frac{Z}{2a_0}\right)^{3/2} \left(2 - \frac{Zr}{a_0}\right) \mathrm{e}^{-Zr/2a_0} , \tag{7.184}$$

$$u_{21}(r) = \left(\frac{Z}{2a_0}\right)^{3/2} \frac{Zr}{\sqrt{3}a_0} \mathrm{e}^{-Zr/2a_0} . \tag{7.185}$$

These functions (multiplied by r) are plotted in Fig. 7.4.

7.9.3 Bessel Functions

Consider the differential equation

---

[16] See [65], Sect. 3.10 and [143], Chap. 4. The polynomial $L_n^a(x)$ is obtained in Maple V through the command `L(n,a,x)` in the `orthopoly` package; and, in Mathematica, through the command `LaguerreL[n,a,x]`.

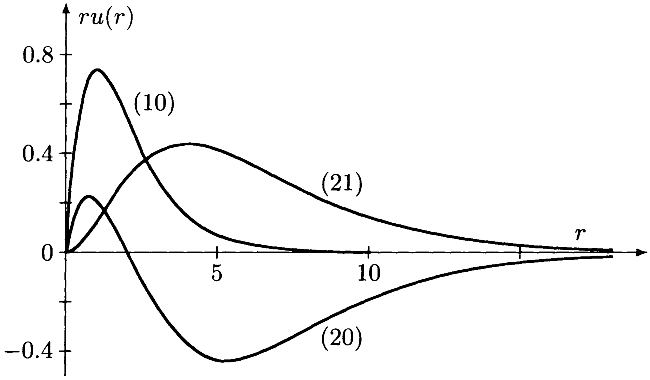

Fig. 7.4. Functions ru_{10}, ru_{20} and ru_{21} associated with the r^{-1} potential; the unit of length is a_0/Z

$$\frac{\mathrm{d}^2 f}{\mathrm{d}z^2} + \frac{1}{z}\frac{\mathrm{d}f}{\mathrm{d}z} + \left(1 - \frac{\nu^2}{z^2}\right) f(z) = 0, \tag{7.186}$$

where z and ν are, in general, complex variables. One easily shows that this equation has a series solution of the form

$$f(z) = J_\nu(z) = \sum_{k=0}^{\infty} \frac{(-1)^k (z/2)^{\nu+2k}}{\Gamma(k+1)\Gamma(k+\nu+1)}. \tag{7.187}$$

The function $J_\nu(z)$ is called a *Bessel function of the first kind* of order ν.[17]

If ν is not an integer, $J_\nu(z)$ and $J_{-\nu}(z)$ are two linearly independent solutions of (7.186). If on the other hand $\nu = n$ is an integer, one sees that $J_n(z) = (-1)^n J_{-n}(z)$. In that case ont can still find a second linearly independent solution of (7.186), but we will not need it here.

The function $J_\nu(z)$ is analytic for any complex ν and any z defined in the complex plane cut on the interval $(-\infty, 0]$. In fact $J_\nu(z)$ becomes analytic for any $z \neq 0$ if the variable z is defined on an appropriate Riemann surface.

From the series expansion one shows that Bessel functions satisfy the following recurrence relations:

$$J_{\nu-1}(z) + J_{\nu+1}(z) = \frac{2\nu}{z} J_\nu(z)\,, \tag{7.188}$$

$$J_{\nu-1}(z) - J_{\nu+1}(z) = 2\frac{\mathrm{d}}{\mathrm{d}z} J_\nu(z)\,. \tag{7.189}$$

One can also show that for $|z|$ large and $|\arg z| < \pi$, $J_\nu(z)$ has the following asymptotic expansion

---

[17] See [143], Chap. 5 and [236], Chap. 17. The function $J_n(x)$ is obtained through the commands `BesselJ(n,x)` and `BesselJ[n,x]` in Maple V and Mathematica, respectively.

$$J_\nu(z) = \left(\frac{2}{\pi z}\right)^{1/2} \left\{f_\nu(z)\cos\left[z - \frac{\nu\pi}{2} - \frac{\pi}{4}\right] + g_\nu(z)\sin\left[z - \frac{\nu\pi}{2} - \frac{\pi}{4}\right]\right\}, \tag{7.190}$$

where

$$f_\nu(z) = 1 - \frac{1}{8z^2}\left\{\left(\nu^2 - 1/4\right)\left(\nu^2 - 9/4\right)\right\} + O(z^{-4}), \tag{7.191}$$

$$g_\nu(z) = -\frac{1}{2z}\left(\nu^2 - 1/4\right) + O(z^{-3}). \tag{7.192}$$

In quantum mechanics one mostly deals with Bessel functions of half-integral order. Let l be a nonnegative integer.[18] One defines the *spherical Bessel function* $j_l(z)$ and the *spherical Neumann function* $n_l(z)$ as

$$j_l(z) = \left(\frac{\pi}{2z}\right)^{1/2} J_{l+1/2}(z), \tag{7.193}$$

$$n_l(z) = (-1)^l \left(\frac{\pi}{2z}\right)^{1/2} J_{-l-1/2}(z). \tag{7.194}$$

From (7.186) it is easy to see that functions $j_l(z)$ and $n_l(z)$ satisfy the following differential equation:

$$\frac{\mathrm{d}^2F}{\mathrm{d}z^2} + \frac{2}{z}\frac{\mathrm{d}F}{\mathrm{d}z} + \left\{1 - \frac{l(l+1)}{z^2}\right\}F(z) = 0. \tag{7.195}$$

By means of the series expansion of $J_\nu(z)$ one can see that

$$j_0(z) = \frac{1}{z}\sin z, \qquad n_0(z) = \frac{1}{z}\cos z. \tag{7.196}$$

The recurrence relations of $J_\nu(z)$ yield all functions j_l and n_l (for any positive integer l). One finds for example

$$\begin{aligned}
j_1(z) &= \frac{1}{z^2}(\sin z - z\cos z), \\
n_1(z) &= \frac{1}{z^2}(\cos z + z\sin z), \\
j_2(z) &= \frac{1}{z^3}\left\{(3 - z^2)\sin z - 3z\cos z\right\}, \\
n_2(z) &= \frac{1}{z^3}\left\{(3 - z^2)\cos z + 3z\sin z\right\}.
\end{aligned}$$

For small values of z the series expansion of $J_\nu(z)$ implies that

---

[18] In fact that restriction can be waived and the following definitions adapted to complex values of l.

$$j_l(z) = \frac{z^l}{(2l+l)!!}\left\{1 - \frac{z^2}{2(2l+3)} + \cdots\right\}, \tag{7.197}$$

$$n_l(z) = \frac{(2l-1)!!}{z^{l+1}}\left\{1 + \frac{z^2}{2(2l-1)} + \cdots\right\}. \tag{7.198}$$

The symbol $(2l+1)!!$ stands for $(2l+1) \times (2l-1) \times (\cdots) \times 5 \times 3 \times 1$.

For large values of $|z|$ the asymptotic expansion of $J_\nu(z)$ entails that ($|\arg z| < \pi$)

$$j_l(z) = \frac{1}{z}\left\{f_\nu(z)\sin\left[z - \frac{l\pi}{2}\right] - g_\nu(z)\cos\left[z - \frac{l\pi}{2}\right]\right\}, \tag{7.199}$$

$$n_l(z) = \frac{1}{z}\left\{f_\nu(z)\cos\left[z - \frac{l\pi}{2}\right] + g_\nu(z)\sin\left[z - \frac{l\pi}{2}\right]\right\}. \tag{7.200}$$

Here $\nu = l + 1/2$.

One can show ([223], p. 226) that

$$\exp(\mathrm{i}kr\cos\theta) = \sum_{l=0}^{\infty}(2l+1)\mathrm{i}^l j_l(kr)P_l(\cos\theta). \tag{7.201}$$

It is useful to define two linear combinations h_l^+ and h_l^- of j_l and n_l that, of course, are also solutions of (7.195):

$$h_l^+(z) = j_l(z) - \mathrm{i}n_l(z)\,, \tag{7.202}$$

$$h_l^-(z) = j_l(z) + \mathrm{i}n_l(z)\,. \tag{7.203}$$

Asymptotic expansions of these functions are given by

$$h_l^+(z) = -\frac{\mathrm{i}}{z}\left\{f_\nu(z) - \mathrm{i}g_\nu(z)\right\}\exp\left\{\mathrm{i}\left[z - \frac{l\pi}{2}\right]\right\}, \tag{7.204}$$

$$h_l^-(z) = \frac{\mathrm{i}}{z}\left\{f_\nu(z) + \mathrm{i}g_\nu(z)\right\}\exp\left\{-\mathrm{i}\left[z - \frac{l\pi}{2}\right]\right\}. \tag{7.205}$$

Exercises

7.1. A particle in three dimensions is restricted to a cubic box of length $2L$. The Hamiltonian is given by

$$H = \frac{1}{2m}\boldsymbol{P}\cdot\boldsymbol{P} + V(\boldsymbol{R}).$$

a) Evaluate $\langle\phi|H|\psi\rangle$ and $\langle\psi|H|\phi\rangle$ in terms of wave functions $\phi(\boldsymbol{r})$ and $\psi(\boldsymbol{r})$.
b) Check that the Hamiltonian is Hermitian if wave functions and their derivatives are equal on opposite faces of the cube [$\psi(L,y,z) = \psi(-L,y,z)$, etc.].

7.2. Show that the probability current associated with the wave function

$$\psi(x) = A_{-}\mathrm{e}^{\mathrm{i}kx} + B_{-}\mathrm{e}^{-\mathrm{i}kx},$$

where k is real and A_{-} and B_{-} are complex, is given by

$$\boldsymbol{j}(x) = \frac{\hbar k}{m}\left\{|A_{-}|^2 - |B_{-}|^2\right\}\hat{x}.$$

7.3. Show that the probability current associated with the wave function

$$\psi(\boldsymbol{r},t) = |\psi(\boldsymbol{r},t)|\exp\left\{\mathrm{i}\Theta(\boldsymbol{r},t)\right\},$$

where $\Theta(\boldsymbol{r},t)$ is real, is given by

$$\boldsymbol{j}(\boldsymbol{r},t) = \frac{\hbar}{m}|\psi(\boldsymbol{r},t)|^2\,\boldsymbol{\nabla}\Theta(\boldsymbol{r},t).$$

7.4. Spherical coordinates are defined in (7.69)–(7.71). Compute $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ in terms of $\frac{\partial}{\partial r}$, $\frac{\partial}{\partial \theta}$ and $\frac{\partial}{\partial \phi}$ and show that

$$x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} = \frac{\partial}{\partial \phi}.$$

7.5. Check that

$$\int_0^{\pi}\sin\theta\,\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\,Y_{10}^*(\theta,\phi)Y_{10}(\theta,\phi) = 1,$$
$$\int_0^{\pi}\sin\theta\,\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\phi\,Y_{20}^*(\theta,\phi)Y_{10}(\theta,\phi) = 0.$$

7.6. Let $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ be three real three-dimensional vectors. Use properties of the Levi-Civita symbol ε_{ijk} to show that

$$\boldsymbol{a}\times(\boldsymbol{b}\times\boldsymbol{c}) = \boldsymbol{b}(\boldsymbol{a}\cdot\boldsymbol{c}) - \boldsymbol{c}(\boldsymbol{a}\cdot\boldsymbol{b}),$$
$$\boldsymbol{a}\cdot(\boldsymbol{b}\times\boldsymbol{c}) = \boldsymbol{c}\cdot(\boldsymbol{a}\times\boldsymbol{b}).$$

7.7. In spherical coordinates and for a spherically symmetric potential, the eigenvalue equation of H can be written as

$$-\frac{\hbar^2}{2m}\left\{\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi(\boldsymbol{r})}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi(\boldsymbol{r})}{\partial\theta}\right)\right.$$
$$\left. + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\psi(\boldsymbol{r})\right\} + [V(r) - E]\,\psi(\boldsymbol{r}) = 0.$$

Suppose one looks for a solution of the form $\psi(\boldsymbol{r}) = u(r)f(\theta,\phi)$ without a priori knowing what are $u(r)$ and $f(\theta,\phi)$. Show that these functions necessarily satisfy the following equations, where λ is a constant:

$$\frac{1}{f(\theta,\phi)\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right)+\frac{1}{f(\theta,\phi)\sin^2\theta}\frac{\partial^2 f}{\partial\phi^2}=\lambda\,,$$

$$\frac{1}{u(r)}\frac{\partial}{\partial r}\left(r^2\frac{\partial u}{\partial r}\right)+\lambda-\frac{2mr^2}{\hbar^2}\left[V(r)-E\right]=0\,.$$

7.8. Check that substitution of (7.102) in (7.101) indeed yields (7.103).

7.9. Check that for the r^{-1} potential

$$\int \mathrm{d}\boldsymbol{r}\,|\psi_{100}(\boldsymbol{r})|^2=1\,,$$

$$\int \mathrm{d}\boldsymbol{r}\,\psi^*_{200}(\boldsymbol{r})\psi_{100}(\boldsymbol{r})=0\,.$$

7.10. Compute $\langle R\rangle$ in the ground state of a particle subject to an r^{-1} potential and find the value of $\{\langle R^2\rangle-\langle R\rangle^2\}^{1/2}\equiv\Delta R$.

7.11.

a) Show that in the ground state $\psi_{100}(\boldsymbol{r})$ of the r^{-1} potential,

$$\langle X\rangle=0=\langle P_x\rangle,\qquad \langle X^2\rangle=\frac{1}{3}\langle R^2\rangle,\qquad \langle P_x^2\rangle=\frac{1}{3}\langle \boldsymbol{P}\cdot\boldsymbol{P}\rangle.$$

b) Find the value of the product $\Delta X\Delta P_x$ and check that the Heisenberg inequality is satisfied.

7.12. The eigenvalue equation of the three-dimensional harmonic oscillator Hamiltonian is given by

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)\psi(\boldsymbol{r})+\frac{1}{2}m\omega^2r^2\psi(\boldsymbol{r})=E\psi(\boldsymbol{r}).$$

Look for a solution of the form

$$\psi(\boldsymbol{r})=f(x)g(y)h(z)\,,$$

and show that the eigenvalue equation separates in three independent equations, each one describing a one-dimensional harmonic oscillator. What values of E correspond to normalizable wave functions?

7.13. Show that the Schrödinger equation for a particle in an electromagnetic field is invariant under a gauge transformation given by

$$\phi'=\phi-\frac{\partial\chi}{\partial t},\qquad \boldsymbol{A}'=\boldsymbol{A}+\boldsymbol{\nabla}\chi,\qquad \psi'=\exp\left\{\frac{\mathrm{i}q\chi}{\hbar}\right\}\psi.$$

Here, χ is an arbitrary function of $\boldsymbol{r}$ and t.

7.14. Let $\mathcal{P}$ and $\mathcal{L}$ be the operators defined in (7.140).
a) Show that

$$[\mathcal{P}_i, \mathcal{P}_j] = \mathrm{i}\hbar q \sum_{k=1}^{3} \varepsilon_{ijk} B_k \,,$$

$$[\mathcal{L}_i, \mathcal{L}_j] = \mathrm{i}\hbar \sum_{k=1}^{3} \varepsilon_{ijk} \left\{ \mathcal{L}_k + q\,(\boldsymbol{R}\cdot\boldsymbol{B})\, X_k \right\}.$$

This means that the components of $\mathcal{P}$ and $\mathcal{L}$ have the same commutation relations as those of $\boldsymbol{P}$ and $\boldsymbol{L}$ if and only if the magnetic induction vanishes.
b) In the commutation relations of $\mathcal{P}$ and $\mathcal{L}$, what is the importance of terms linear in B if B is of order 1 T?

7.15. Obtain Ehrenfest's relations (7.141) and (7.142) for a particle in an electromagnetic field.

7.16. Show that the probability current associated with a particle in an electromagnetic field is given by

$$\boldsymbol{j}(\boldsymbol{r},t) = -\frac{\hbar}{m} \mathrm{Im} \left\{ \psi \left[\left(\nabla - \frac{\mathrm{i}q}{\hbar} \boldsymbol{A} \right) \psi \right]^* \right\}.$$

7.17.
a) Show that

$$\frac{\mathrm{d}}{\mathrm{d}z} \left\{ z n_l(z) \frac{\mathrm{d}}{\mathrm{d}z} \left[z j_l(z) \right] - z j_l(z) \frac{\mathrm{d}}{\mathrm{d}z} \left[z n_l(z) \right] \right\} = 0.$$

b) Compute the expression between curly brackets for small z and deduce that

$$n_l(z) \frac{\mathrm{d}}{\mathrm{d}z} \left[z j_l(z) \right] - j_l(z) \frac{\mathrm{d}}{\mathrm{d}z} \left[z n_l(z) \right] = \frac{1}{z}.$$

8 Numerical Solution

For a particle either in one or in three dimensions, there are few situations where the eigenvalue equation for the Hamiltonian has closed-form solutions. In most cases one must turn to approximate methods. These, we will see, are very different from one another. Not all methods are adapted to any specific problem, each method having its own domain of application.

This chapter deals with simple numerical methods. The differential equation for the wave function $\psi(x)$ or the radial function $u(r)$ is transformed into a finite-difference equation, a discrete approximation of the former. The finite-difference equation can be solved recursively, with the help of a computer. This technique yields the Hamiltonian's eigenvalues for a one-dimensional or a three-dimensional spherically symmetric potential. It also provides the corresponding eigenfunctions, in the continuous as well as in the discrete spectrum. Numerical methods are primarily interesting due to their generality. But they also bring a deeper understanding of the Hamiltonian's eigenvalue equation.[1]

8.1 Finite-Difference Equation

Consider the differential equation

$$\frac{\mathrm{d}^2 F}{\mathrm{d}\xi^2} + h(\xi) F = 0 \,, \tag{8.1}$$

where ξ is a real variable, $h(\xi)$ a known function and $F(\xi)$ the dependent variable we are trying to solve for. In the case of a one-dimensional potential it is clear that the eigenvalue equation of the Hamiltonian can be written in this form. Later we will show that the radial equation associated with a three-dimensional spherically symmetric potential can also be written in the form (8.1).

The variable ξ is continuous. We define a discrete variable ξ_n, associated with ξ, in the following way:

---

[1] For a pedagogical application of numerical methods to several quantum-mechanical problems, see [41], [42] and [112]. References [82] and [118] introduce a systematic use of symbolic softwares.

L. Marchildon, *Quantum Mechanics*

$$\xi_n = n\varDelta\,. \tag{8.2}$$

Here n is an integer between two integers N_G and N_D.[2] In general $N_D \gg N_G$. The constant $\varDelta$ is chosen so that the function $h(\xi)$ does not change much in an interval of length $\varDelta$. Integers N_D and N_G are chosen so that the interval $(N_G\varDelta, N_D\varDelta)$ covers the whole region where the dependent variable F is to be determined.

Having replaced the continuous variable ξ by the discrete variable ξ_n, we similarly associate, with functions $h(\xi)$ and $F(\xi)$, discrete variables h_n and F_n defined as

$$h_n = h(\xi_n)\,, \qquad F_n = F(\xi_n)\,. \tag{8.3}$$

We can find a discrete approximation of the second derivative of F, evaluated at the point ξ_n. From the Taylor series expansion one has

$$\begin{aligned} F_{n+1} &= F(\xi_{n+1}) = F(\xi_n + \varDelta) \\ &= F_n + \varDelta\left(\frac{\mathrm{d}F}{\mathrm{d}\xi}\right)_{\xi_n} + \frac{\varDelta^2}{2}\left(\frac{\mathrm{d}^2F}{\mathrm{d}\xi^2}\right)_{\xi_n} + \frac{\varDelta^3}{6}\left(\frac{\mathrm{d}^3F}{\mathrm{d}\xi^3}\right)_{\xi_n} + O(\varDelta^4)\,, \end{aligned} \tag{8.4}$$

$$F_{n-1} = F_n - \varDelta\left(\frac{\mathrm{d}F}{\mathrm{d}\xi}\right)_{\xi_n} + \frac{\varDelta^2}{2}\left(\frac{\mathrm{d}^2F}{\mathrm{d}\xi^2}\right)_{\xi_n} - \frac{\varDelta^3}{6}\left(\frac{\mathrm{d}^3F}{\mathrm{d}\xi^3}\right)_{\xi_n} + O(\varDelta^4)\,. \tag{8.5}$$

Here $O(\varDelta^4)$ stands for all terms of order $\varDelta^4$ and more that are neglected. Adding together (8.4) and (8.5) and dividing the result by $\varDelta^2$, one easily obtains that

$$\left(\frac{\mathrm{d}^2F}{\mathrm{d}\xi^2}\right)_{\xi_n} = \frac{F_{n+1} - 2F_n + F_{n-1}}{\varDelta^2} + O(\varDelta^2). \tag{8.6}$$

Let us now evaluate (8.1) at the point ξ_n and let us substitute (8.6) in it. We find that

$$F_{n+1} + \left\{-2 + \varDelta^2 h_n\right\} F_n + F_{n-1} = O(\varDelta^4).$$

This equation is exact if $O(\varDelta^4)$ represents all terms of order $\varDelta^4$ and more that were omitted in the Taylor series. If, however, these terms are neglected, the following approximate equation can be written down:

$$F_{n+1} = \left\{2 - \varDelta^2 h_n\right\} F_n - F_{n-1}\,. \tag{8.7}$$

From a differential equation we have gone to a finite-difference equation, which is a discrete approximation of the former. Clearly, the smaller $\varDelta$ is, the better the approximation.

---

[2] Indices G and D stand for “gauche” and “droit”, i.e. left and right.

In (8.7) the discrete variable h_n is assumed known while F_n is the unknown one is looking for. The variable F_n can be determined for any n from two consecutive initial values. Indeed suppose one knows F_{n_0} and F_{n_0-1}. Equation (8.7) then yields F_{n_0+1}. From F_{n_0} and F_{n_0+1} one next finds F_{n_0+2}, and so on. Similarly, one obtains F_{n_0-2}, F_{n_0-3} and so on. In this way one can get values of F_n for any n between N_G and N_D, no matter what the values of N_G and N_D are.

In practice this recursive way of solving (8.7) is carried out with the help of a computer. It should be kept in mind that at each step of the recursive computation, an error of order Δ^4 is made. This, for instance, means that F_{n_0+1}, computed from F_{n_0} and F_{n_0-1} by means of (8.7), differs from the value $F(\xi_{n_0+1})$ obtained from solving the differential equation by a quantity of order Δ^4. After a large number of iterations, small errors can add up to an eventually significant overall error. The overall error can in general be reduced by decreasing the value of Δ. The smaller Δ is, however, the larger the number of steps necessary to cover a given interval (ξ_G, ξ_D) of the independent variable ξ.

It is also possible to reduce the error at each step of the recursive process without decreasing the value of Δ. For this purpose, one must develop a more precise recurrence formula than (8.7). For a differential equation like (8.1), such a formula can be obtained easily. Carrying out the Taylor expansion up to terms of order Δ^6, one finds that (Exercise 8.1)

$$\left(\frac{\mathrm{d}^2F}{\mathrm{d}\xi^2}\right)_{\xi_n} = \frac{F_{n+1} - 2F_n + F_{n-1}}{\Delta^2} - \frac{\Delta^2}{12}\left(\frac{\mathrm{d}^4F}{\mathrm{d}\xi^4}\right)_{\xi_n} + O(\Delta^4). \tag{8.8}$$

But

$$\left(\frac{\mathrm{d}^4F}{\mathrm{d}\xi^4}\right)_{\xi_n} = \frac{1}{\Delta^2}\left\{\left(\frac{\mathrm{d}^2F}{\mathrm{d}\xi^2}\right)_{\xi_{n+1}} - 2\left(\frac{\mathrm{d}^2F}{\mathrm{d}\xi^2}\right)_{\xi_n} + \left(\frac{\mathrm{d}^2F}{\mathrm{d}\xi^2}\right)_{\xi_{n-1}}\right\} + O(\Delta^2),$$

so that

$$\begin{aligned}\frac{1}{12}&\left\{\left(\frac{\mathrm{d}^2F}{\mathrm{d}\xi^2}\right)_{\xi_{n+1}} + 10\left(\frac{\mathrm{d}^2F}{\mathrm{d}\xi^2}\right)_{\xi_n} + \left(\frac{\mathrm{d}^2F}{\mathrm{d}\xi^2}\right)_{\xi_{n-1}}\right\}\\ &= \frac{F_{n+1} - 2F_n + F_{n-1}}{\Delta^2} + O(\Delta^4)\,.\end{aligned}$$

Substituting (8.1) evaluated at ξ_{n+1}, ξ_n and ξ_{n-1} in the previous equation, multiplying by Δ^2 and rearranging, we obtain

$$\begin{aligned}&\left\{1 + \frac{\Delta^2}{12}h_{n+1}\right\}F_{n+1}\\ &\quad = 2\left\{1 - \frac{5\Delta^2}{12}h_n\right\}F_n - \left\{1 + \frac{\Delta^2}{12}h_{n-1}\right\}F_{n-1} + O(\Delta^6)\,.\end{aligned}$$

Finally, neglecting terms of order Δ^6 we find the following approximate equation:

$$F_{n+1} = \left\{1 + \frac{\Delta^2}{12} h_{n+1}\right\}^{-1} \left\{2\left[1 - \frac{5\Delta^2}{12} h_n\right] F_n - \left[1 + \frac{\Delta^2}{12} h_{n-1}\right] F_{n-1}\right\}. \tag{8.9}$$

Like (8.7), the recurrence equation (8.9) yields all values of F_n from two initial values, say F_{n_0} and F_{n_0-1}. At this stage these two values are arbitrary. They correspond to the well-known result that the solution of a second-order differential equation is unique if two initial conditions are specified. The use of (8.9) to solve the differential equation (8.1) approximately is called the *Numerov method.*

At each step, (8.9) makes an error of order Δ^6, in contrast with an error of order Δ^4 made by (8.7). For a given value of Δ, (8.9) is therefore much more precise than (8.7). From another point of view, suppose one would like to solve (8.1) in a specified interval (ξ_G, ξ_D), with a given precision. One can then pick a larger value of Δ (hence make fewer iterations) if one uses (8.9) rather than (8.7).

8.2 Discrete Spectrum in One Dimension

Consider a particle of mass m in one dimension. The Hamiltonian's eigenvalue equation is given by

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\psi(x) + V(x)\psi(x) = E\psi(x).$$

We will use results of Sect. 8.1 to obtain energies and wave functions associated with this equation.[3]

Compared with meters and joules, lengths and energies typical of atomic processes are very small. So it is convenient to make some changes of variables. Let us introduce a dimensionless variable

$$\xi = x/a, \tag{8.10}$$

where a is a constant characteristic of the potential's range. Let us also introduce a constant ε and a variable $v(\xi)$, both dimensionless, as

$$\varepsilon = \frac{2ma^2E}{\hbar^2}, \qquad v(\xi) = \frac{2ma^2}{\hbar^2}V(x). \tag{8.11}$$

The eigenvalue equation becomes

---

[3] Numerical solution of the Schrödinger equation and time evolution of the wave function are discussed in [92] and [229].

$$\frac{\mathrm{d}^2}{\mathrm{d}\xi^2}\psi(\xi)+\{\varepsilon-v(\xi)\}\,\psi(\xi)=0\,. \tag{8.12}$$

Equation (8.12) has the form of (8.1) with

$$h(\xi)=\varepsilon-v(\xi)\,. \tag{8.13}$$

Let v_∞ be the limit, as $\xi \to \pm\infty$, of the function $v(\xi)$. In general (8.12) has a continuous spectrum for $\varepsilon > v_\infty$ and a discrete spectrum for $\varepsilon < v_\infty$. Here we are interested in the discrete spectrum. The continuous spectrum will be examined in Sect. 8.3. We look for values of ε for wich (8.12) has a normalizable solution, that is, a solution that goes to zero as $|\xi| \to \infty$. For these values of ε we wish to obtain the corresponding solution $\psi(\xi)$ and normalize it.

To be specific, assume that $v_\infty = 0$ and that, as $|\xi| \to \infty$, the potential $v(\xi)$ goes to zero faster than $|\xi|^{-1}$. Solutions of (8.12) then have the following asymptotic form (Exercise 8.2):

$$\psi(\xi) \sim \exp\left\{\pm\sqrt{-\varepsilon}|\xi|\right\} \quad \text{if} \quad |\xi| \to \infty\,. \tag{8.14}$$

The constant $-\varepsilon$ is positive in the discrete spectrum. If ε is an eigenvalue, the asymptotic form only involves the decreasing exponential.

Let $\varepsilon < 0$. For $\psi(\xi)$ specify the form $\exp\{-\sqrt{-\varepsilon}|\xi|\}$ if $|\xi|$ is large, and numerically integrate (8.12) up to the point $\xi = 0$, starting from negative as well as positive values of ξ. It is clear that values of $\psi(0)$ and $(\mathrm{d}\psi/\mathrm{d}\xi)_{\xi=0}$, obtained both ways, will then coincide (up to a multiplicative factor) if and only if ε is an eigenvalue. This observation forms the basis of the algorithm we will presently develop.

For $\xi \ll 0$ let

$$\psi^-(\xi) = A_-\exp\left\{\sqrt{-\varepsilon}\xi\right\}\,. \tag{8.15}$$

The exponent $(-)$ indicates that $\psi(\xi)$ will here be calculated starting from negative values of ξ. Let N_G be the farthest negative value of the index n used for the discretization of the variable ξ. The two initial values of ψ_n can be chosen as

$$\psi^-_{N_G} = A_-\exp\left\{\sqrt{-\varepsilon}N_G\Delta\right\}, \tag{8.16}$$

$$\psi^-_{N_G+1} = A_-\exp\left\{\sqrt{-\varepsilon}(N_G+1)\Delta\right\}. \tag{8.17}$$

Numerov's recurrence formula, found in Sect. 8.1, can then be used starting from these two values. In the present case, (8.9) is given by

$$\psi^-_{n+1} = \left\{1+\frac{\Delta^2}{12}(\varepsilon - v_{n+1})\right\}^{-1} \times\left\{2\left[1-\frac{5\Delta^2}{12}(\varepsilon-v_n)\right]\psi^-_n - \left[1+\frac{\Delta^2}{12}(\varepsilon-v_{n-1})\right]\psi^-_{n-1}\right\}. \tag{8.18}$$

This equation is solved recursively up to values of n near zero. One thus obtains ψ_0^-, the value of the function ψ at $\xi = 0$. The derivative of ψ at $\xi = 0$ can be calculated appropriately by the following formula (Exercise 8.1):[4]

$$\left(\frac{\mathrm{d}\psi^-}{\mathrm{d}\xi}\right)_{\xi=0} = \frac{-\psi_2^- + 8\psi_1^- - 8\psi_{-1}^- + \psi_{-2}^-}{12\Delta} + O(\Delta^4). \tag{8.19}$$

Define $R_-(\varepsilon)$, the ratio of the derivative of ψ^- to ψ^-, as

$$R_-(\varepsilon) = \frac{1}{\psi^-(0)}\left(\frac{\mathrm{d}\psi^-}{\mathrm{d}\xi}\right)_{\xi=0}. \tag{8.20}$$

As the notation indicates, R_- depends on the value of ε picked at the beginning of the recursive computation. It is important to point out, however, that $R_-(\varepsilon)$ is independent of the multiplicative factor A_- in (8.16) and (8.17).

Let us now turn to positive asymptotic values of ξ. For $\xi \gg 0$ one can take

$$\psi^+(\xi) = A_+ \exp\left\{-\sqrt{-\varepsilon}\xi\right\}. \tag{8.21}$$

Let N_D be the largest positive value of the index n. One has

$$\psi_{N_D}^+ = A_+ \exp\left\{-\sqrt{-\varepsilon}N_D\Delta\right\}, \tag{8.22}$$

$$\psi_{N_D-1}^+ = A_+ \exp\left\{-\sqrt{-\varepsilon}(N_D - 1)\Delta\right\}. \tag{8.23}$$

Constants $\psi_{N_D}^+$ and $\psi_{N_D-1}^+$ can be taken as initial values in Numerov's recurrence formula. This formula is then used to compute the ψ_n^+ for n decreasing. Once the ψ_n^+ for n near zero are obtained, the ratio $R_+(\varepsilon)$ can be defined as

$$R_+(\varepsilon) = \frac{1}{\psi^+(0)}\left(\frac{\mathrm{d}\psi^+}{\mathrm{d}\xi}\right)_{\xi=0}. \tag{8.24}$$

The derivative of ψ^+ is calculated by a formula similar to (8.19).

We know that ε is an eigenvalue just in case there exist constants A_+ and A_- such that ψ^+ and ψ^-, as well as their derivatives, coincide at $\xi = 0$. Clearly this implies that $R_+(\varepsilon) = R_-(\varepsilon)$. Conversely, it is not difficult to show that unless ψ vanishes at $\xi = 0$, the equality of $R_+(\varepsilon)$ and $R_-(\varepsilon)$ implies the equality of ψ^+ and ψ^- and the equality of their derivatives, at $\xi = 0$, for an appropriate choice of A_+ and A_-. Hence if $R_+(\varepsilon)$ and $R_-(\varepsilon)$ exist, a necessary and sufficient condition for ε to be an eigenvalue is that

$$R_+(\varepsilon) = R_-(\varepsilon). \tag{8.25}$$

---

[4] It is not necessary to use, for the derivative, so precise a formula as Numerov's recurrence formula. The formula for the derivative is used at one point only, while Numerov's is applied a large number of times. The cumulative error of Numerov's method is of order Δ^4 [215].

The search for eigenvalues ε_i associated with (8.12) thus proceeds as follows. In a given interval of values of ε, one systematically computes ratios $R_-(\varepsilon)$ and $R_+(\varepsilon)$, following the method just described. Eigenvalues are then obtained as zeros of the function $R_+(\varepsilon) - R_-(\varepsilon)$.[5]

The initial values used in the recursive calculation ($\psi^-_{N_G}$, $\psi^-_{N_G+1}$; $\psi^+_{N_D}$, $\psi^+_{N_D-1}$) are not really crucial. They represent the value of ψ and of its derivative at the starting point, where $|\xi|$ is large. The common multiplicative factor has no effect on the search for eigenvalues. As far as the ratio of derivative to function is concerned, it corresponds to the relative weight of the solutions $\exp\{+\sqrt{-\varepsilon}|\xi|\}$ and $\exp\{-\sqrt{-\varepsilon}|\xi|\}$ in the function $\psi(\xi)$. But any contribution coming from the solution $\exp\{+\sqrt{-\varepsilon}|\xi|\}$ becomes rapidly negligible as $|\xi|$ decreases. Only the correct asymptotic form contributes significantly to $\psi(\xi)$, after a certain number of iterations. This observation entails that the algorithm we developed also applies to potentials $v(\xi)$ with an asymptotic form much more general than the one we specified at the beginning.

When eigenvalues ε_i have been obtained in a given interval of values of ε, there remains to find corresponding eigenfunctions $\psi_{\varepsilon_i}(\xi)$ as well as their norm. Suppose that for an eigenvalue ε_i, values ψ^-_n and ψ^+_n have been obtained. Define ψ_n in the following way:

$$\psi_n = \begin{cases} \psi^-_n & \text{if } n \leq 0 \,, \\ A\psi^+_n & \text{if } n \geq 0 \,. \end{cases} \tag{8.26}$$

Here A is picked so that $\psi^-_0 = A\psi^+_0$. So defined, ψ_n is the discrete approximation of a function that is continuous and has a continuous derivative everywhere, that solves (8.12) and that goes to zero as $|\xi| = |n\Delta|$ gets large. The variable ψ_n thus represents the eigenfunction we are looking for. Its norm can be computed by means of the relation

$$\int_{-\infty}^{\infty} \mathrm{d}x\, |\psi(x)|^2 = a \int_{-\infty}^{\infty} \mathrm{d}\xi\, |\psi(\xi)|^2 = a\Delta \sum_{n=N_G}^{N_D} |\psi_n|^2 \,, \tag{8.27}$$

where the trapezoidal rule was used for the numerical evaluation of the integral. More precise formulas can also be used if needed ([187], Chap. 4).

The method we just presented for the search for eigenvalues and eigenfunctions of the Hamiltonian of a particle in one dimension essentially applies

---

[5] The search for zeros of a function is treated in detail in [187], Chap. 9. Note that if one looks for zeros of the function $R_+(\varepsilon) - R_-(\varepsilon)$ through a change of sign, one must eliminate most changes of sign that occur during a transition through infinity. This is what happens for a value of ε such that $\psi(\xi = 0) = 0$. It may happen, however, that an eigenfunction has a zero precisely at $\xi = 0$. This occurs, for instance, with odd eigenfunctions of an even potential. In that case it is useful to move the potential by a small $\delta\xi$. No eigenfunction will then have a zero at $\xi = 0$, and all eigenvalues will be obtained by (8.25).

to all types of one-dimensional potentials. Some more restrictive methods can apply to specific potentials.[6] Exercise (8.4) develops a method that applies to potentials that are even functions of the independent variable.

8.3 Box Normalization

We have developed an algorithm to find eigenvalues and eigenfunctions of (8.12) in the case where $\varepsilon < v_\infty$, that is, for discrete spectrum. We will now examine the case where $\varepsilon > v_\infty$. To make things precise we assume once more that $v_\infty = 0$ and that $v(\xi)$ goes to zero faster than $|\xi|^{-1}$ as $|\xi| \to \infty$. Solutions of (8.12) then have the following asymptotic form:

$$\psi(\xi) \sim \cos\left\{\sqrt{\varepsilon}\xi\right\}, \quad \sin\left\{\sqrt{\varepsilon}\xi\right\}. \tag{8.28}$$

For any $\varepsilon > v_\infty$, (8.12) has two linearly-independent solutions. They are obtained by applying Numerov's recursive procedure to two pairs of distinct initial conditions, for instance

$$\text{i) } \psi_0 = 1,\ \psi_1 = 1; \qquad \text{ii) } \psi_0 = 1,\ \psi_1 = 2. \tag{8.29}$$

In the continuous spectrum, solutions of (8.12) are not normalizable. Strictly speaking, only normalizable functions belong to the state space. One way to obtain normalizable functions from the continuous spectrum consists in taking not a single solution of (8.12) corresponding to just one value of ε, but rather a linear combination of solutions corresponding to several values of ε (in fact, to a continuous infinity of values). The coefficients of the linear combination can be chosen so that the wave function is normalizable. This procedure is correct, but it may be difficult to implement in concrete problems.

The most convenient way to normalize wave functions in the continuous spectrum consists in enclosing the quantum system in a box. This means that allowed values of position are restricted to a given interval, say $(-L, L)$. Here L is a length large enough for the potential $v(\xi)$ to reach its asymptotic form. The effect of the box can be represented by a change in the potential, so that

$$v(\xi) = \infty \quad \text{if} \quad |\xi| > L. \tag{8.30}$$

All solutions of (8.12) then satisfy the following boundary conditions:

$$\psi(L) = 0 = \psi(-L). \tag{8.31}$$

Denote here by v_∞ the value of the potential just before the points $\xi = \pm L$, that is,

---

[6] In [31] one can find many observations on the solution of the eigenvalue equation by Numerov's method. Other methods are developed in [128].

$$v_\infty = \lim_{|\xi| \to L^-} v(\xi)\,. \tag{8.32}$$

With conditions (8.31), (8.12) only has a discrete spectrum left. The nature of the spectrum is, however, very different according as ε is smaller or larger than v_∞. In the case where $\varepsilon < v_\infty$, values of ε for which (8.12) has a solution that satisfies the boundary conditions come close to eigenvalues ε_i found in Sect. 8.2. There is rapid convergence toward the ε_i as L, the length of the box, goes to infinity.

In the case where $\varepsilon > v_\infty$, values of ε for which conditions (8.31) are satisfied make up a denser and denser spectrum as L increases. In fact the number of eigenvalues in a given interval $\Delta\varepsilon$ is proportional to L. The amplitude of normalized wave functions is proportional to $L^{-1/2}$.

The numerical methods developed in Sect. 8.2 are well adapted to box normalization. For a given value of L one looks for values of ε for which boundary conditions (8.31) are satisfied. One must first compute ratios $R_-(\varepsilon)$ and $R_+(\varepsilon)$ defined in Sect. 8.2. To do this, initial values of ψ for numerical integration can be picked as

$$\begin{aligned} \psi^-_{N_G} &= 0\,, & \psi^+_{N_D} &= 0\,, \\ \psi^-_{N_G+1} &= 1\,, & \psi^+_{N_D-1} &= 1\,. \end{aligned} \tag{8.33}$$

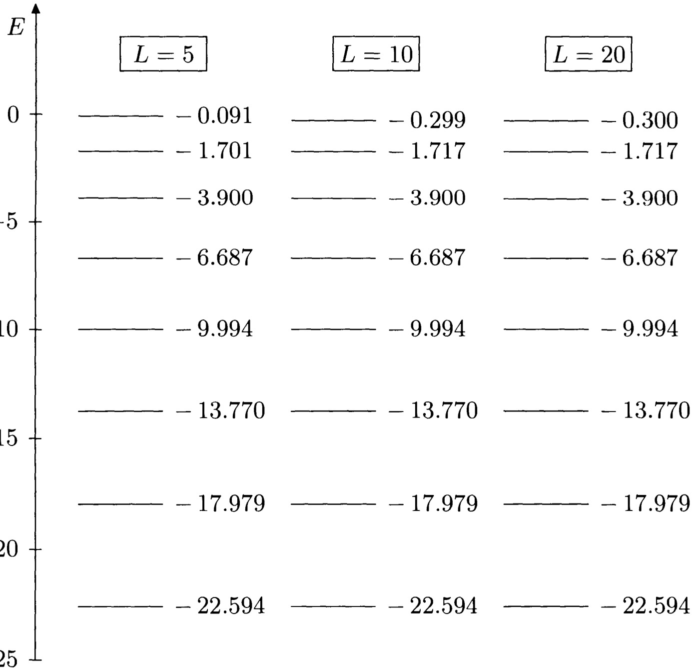

Fig. 8.1. Negative-energy spectrum of Gaussian potential in boxes of different lengths

Values of ε one is looking for are given by zeros of the function $R_+(\varepsilon) - R_-(\varepsilon)$. The solution $\psi(\xi)$ corresponding to each eigenvalue can be normalized.

We illustrate the energy spectrum of a quantum system in a box by the example of a truncated Gaussian potential given by

$$v(\xi) = \begin{cases} -25\exp\left\{-\frac{(\xi-0.2)^2}{4}\right\} & \text{if } |\xi| < L\,, \\ \infty & \text{if } |\xi| > L\,. \end{cases} \tag{8.34}$$

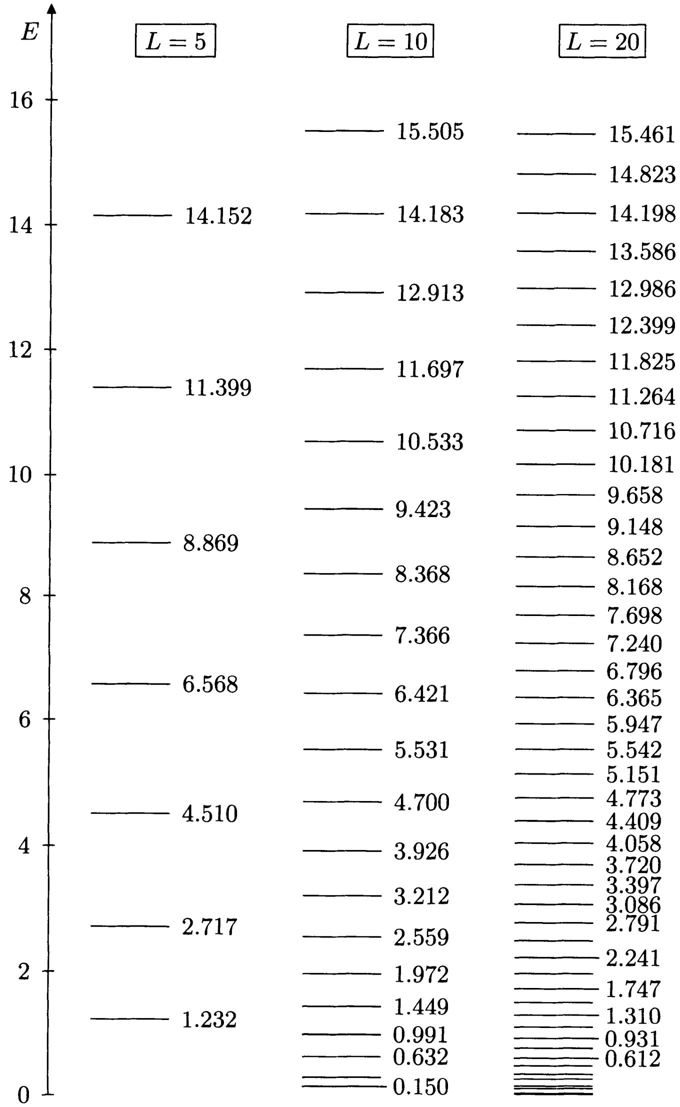

Fig. 8.2. Positive-energy spectrum of Gaussian potential in boxes of different lengths

Figure 8.1 shows the spectrum of negative values of ε for three choices of the parameter L, i.e. $L = 5$, 10 and 20. Eigenvalues quickly stabilize and converge toward the untruncated potential's discrete-spectrum values.

Figure 8.2 shows the spectrum of positive values, up to $\varepsilon = 16$, for the same values of the parameter L. As L increases, the density of eigenvalues also increases. Ultimately ($L \to \infty$) the spectrum becomes continuous.

Several times in the forthcoming we shall treat the continuous spectrum by enclosing the quantum system in a box and taking the limit $L \to \infty$ at the end of computations only. We should point out that box normalization can also be implemented starting from periodic boundary conditions, that is,

$$\psi(L) = \psi(-L), \qquad \left(\frac{\mathrm{d}\psi}{\mathrm{d}\xi}\right)_{\xi=L} = \left(\frac{\mathrm{d}\psi}{\mathrm{d}\xi}\right)_{\xi=-L}. \tag{8.35}$$

One checks that, as $L \to \infty$, negative eigenvalues of ε converge toward those of the discrete spectrum, while the number of positive eigenvalues in a given interval $\Delta\varepsilon$ is proportional to L.

8.4 Discrete Spectrum in Three Dimensions

For a spherically symmetric potential, the eigenvalue equation of the Hamiltonian of a particle in three dimensions separates into an angular and a radial equation.[7] The latter is given by (7.90):

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left\{r^2\frac{\mathrm{d}}{\mathrm{d}r}u(r)\right\} - \left\{\frac{l(l+1)}{r^2} + \frac{2m}{\hbar^2}\left[V(r) - E\right]\right\}u(r) = 0.$$

As in Sect. 8.2, let us introduce the following variables:

$$\zeta = \frac{r}{a}, \qquad \varepsilon = \frac{2ma^2E}{\hbar^2}, \qquad v(\zeta) = \frac{2ma^2}{\hbar^2}V(r). \tag{8.36}$$

Let us also define a function $\bar{u}(\zeta)$ as

$$\bar{u}(\zeta) = \zeta u(r). \tag{8.37}$$

The equation for $\bar{u}(\zeta)$ is then given by

$$\frac{\mathrm{d}^2\bar{u}}{\mathrm{d}\zeta^2} + \left\{\varepsilon - v(\zeta) - \frac{l(l+1)}{\zeta^2}\right\}\bar{u}(\zeta) = 0. \tag{8.38}$$

This has the form of (8.1) with

---

[7] Eigenvalues associated with potentials that are not spherically symmetric can be obtained by the finite-element method. The problem of a hydrogen atom in a magnetic field is treated in this way in [192].

$$h(\zeta) = \varepsilon - v(\zeta) - \frac{l(l+1)}{\zeta^2}. \tag{8.39}$$

Recall that l is a nonnegative integer. The variable ζ is defined in the interval $[0, \infty)$.

We will assume that, as $\zeta \to \infty$, $v(\zeta)$ goes to zero at least as fast as ζ^{-1}. Moreover, for small ζ we will assume that $\zeta v(\zeta)$ is an analytic function of ζ. This means that at $\zeta = 0$, $v(\zeta)$ has at most a pole of order 1.

In Sect. 7.4 we saw that the two linearly-independent solutions of the radial equation behave as r^l and r^{-l-1} respectively. Only the first one is physically acceptable. Hence acceptable solutions for the function $\bar{u}(\zeta)$ behave as ζ^{l+1}. In particular, they all vanish at $\zeta = 0$. Note that eigenfunctions of the Hamiltonian have the form

$$\psi(\boldsymbol{r}) = u(r) Y_{lm}(\theta, \phi) = \frac{\bar{u}(\zeta)}{\zeta} Y_{lm}(\theta, \phi). \tag{8.40}$$

In this section we focus on the discrete spectrum, that is, on negative values of ε. The square of the norm of $\psi(\boldsymbol{r})$ is given by

$$\int \mathrm{d}\boldsymbol{r}\, |\psi(\boldsymbol{r})|^2 = a^3 \int_0^\infty \mathrm{d}\zeta\, |\bar{u}(\zeta)|^2. \tag{8.41}$$

Clearly the wave function is normalizable only if $\bar{u}(\zeta)$ goes to zero as $\zeta \to \infty$.

One can use Numerov's method to numerically integrate (8.38) and obtain its discrete spectrum eigenvalues. Here the index n goes from $N_G = 0$ to N_D, where N_D is large enough that $\zeta_D = N_D \Delta$ is in the asymptotic region of $\bar{u}(\zeta)$. As initial values we use

$$\bar{u}_0 = 0, \qquad \bar{u}_1 = 1, \tag{8.42}$$

the choice of $\bar{u}_1$ being arbitrary and corresponding to a multiplicative factor of the function $\bar{u}(\zeta)$. Values of $\bar{u}_n$ are obtained recursively by means of (8.9), which in the present case is written as

$$\begin{aligned} \bar{u}_{n+1} = {} & \left\{1 + \frac{\Delta^2}{12}\left[\varepsilon - v_{n+1} - \frac{l(l+1)}{\zeta_{n+1}^2}\right]\right\}^{-1} \\ & \times \left\{2\left[1 - \frac{5\Delta^2}{12}\left(\varepsilon - v_n - \frac{l(l+1)}{\zeta_n^2}\right)\right]\bar{u}_n \right. \\ & \left. - \left[1 + \frac{\Delta^2}{12}\left(\varepsilon - v_{n-1} - \frac{l(l+1)}{\zeta_{n-1}^2}\right)\right]\bar{u}_{n-1}\right\}. \end{aligned} \tag{8.43}$$

Denote by $\bar{u}_{N_D}(\varepsilon)$ the final value of $\bar{u}_n$ for a given choice of ε. If ε is not an eigenvalue, one can see from (8.38) that $\bar{u}(\zeta)$ exponentially diverges as $\zeta \to \infty$. If ε is an eigenvalue, $\bar{u}(\zeta) \to 0$ as $\zeta \to \infty$. The eigenvalues' search algorithm will thus consist in determining the zeros of the function $\bar{u}_{N_D}(\varepsilon)$.

Practically, the presence of a zero at $\varepsilon = \varepsilon_i$ can be recognized from the fact that $\bar{u}_{N_D}(\varepsilon)$ changes sign around ε_i. For $\varepsilon > \varepsilon_i$, for instance, $\bar{u}_{N_D}(\varepsilon)$ diverges toward large positive values, whereas $\bar{u}_{N_D}(\varepsilon)$ diverges toward large negative values for $\varepsilon < \varepsilon_i$ (or vice versa). In this way one can find, in a given interval, all eigenvalues ε_i that correspond to a value of the orbital quantum number l. There is, of course, a distinct set of eigenvalues $\{\varepsilon_i\}$ for each value of l.

When an eigenvalue ε_i has been obtained, the radial function $\bar{u}(\zeta)$ can be normalized by numerically computing the integral (8.41). Here it can be necessary to truncate the sum of $|\bar{u}_n|^2$ before the final value $n = N_D$. Indeed even though ε may be very close to an eigenvalue ε_i, values of $\bar{u}_n$ usually start increasing exponentially as n nears N_D. Due to numerical errors, this can occur even if ε coincides with an eigenvalue. The radial function normalization algorithm must be conceived so as not to include an exponentially increasing tail.

So here we have a search procedure for eigenvalues and eigenfunctions of (8.38). It is simple and, in general, adequate. Two additions can nonetheless contribute to make it more precise and more efficient. They come from a close examination of the behavior of eigenfunctions at small and at large values of ζ.

The function $h(\zeta)$, defined in (8.38), is singular at the point $\zeta = 0$. Instead of using, for small values of ζ, numerical methods, we shall expand $\bar{u}(\zeta)$ in a power series in ζ. The solution we are interested in behaves as ζ^{l+1} around $\zeta = 0$. Let us thus write

$$\bar{u}(\zeta) = \zeta^{l+1} \sum_{k=0}^{\infty} a_k \zeta^k . \tag{8.44}$$

Assumptions already made on the potential $v(\zeta)$ entail that one can write

$$v(\zeta) = \sum_{j=-1}^{\infty} b_j \zeta^j . \tag{8.45}$$

Substituting (8.44) and (8.45) in (8.38), one finds (Exercise 8.5) that coefficients a_k must satisfy the following recurrence relation ($a_{-1} = 0$, $k > 0$):

$$a_k = \{k(k + 2l + 1)\}^{-1} \left\{ -\varepsilon a_{k-2} + \sum_{j=1}^{k} a_{k-j} b_{j-2} \right\} . \tag{8.46}$$

The value of a_0 is arbitrary and corresponds to a multiplicative factor of $\bar{u}(\zeta)$.

Knowledge of coefficients a_k for enough values of k ($k = 0, 1, \ldots, K$) allows one to compute $\bar{u}(\zeta)$ around $\zeta = 0$ very precisely. One can then write

$$\bar{u}_n = (n\Delta)^{l+1} \sum_{k=0}^{K} a_k (n\Delta)^k . \tag{8.47}$$

This formula is used to obtain the first few values of $\bar{u}_n$, till an n_0 such that

$$0 \ll n_0 \Delta \ll 1 . \tag{8.48}$$

Numerov's recurrence formula is used thereafter.

Let us now turn to large values of ζ. In a number of applications it so happens that the potential $v(\zeta)$ goes to zero rather slowly, for instance as ζ^{-1}. Eigenfunctions may then take their asymptotic form at very large values of ζ only. For the potential $v(\zeta) = \zeta^{-1}$, for instance, an eigenfunction corresponding to the principal quantum number n sometimes stops oscillating where $\zeta \gg n^2$ only. This means that for a choice of Δ that is appropriate around $\zeta = 1$, a very large number of iterations will be required to reach the asymptotic region.

It is then convenient to pick a Δ that depends on ζ. A simple way to go is to double the value of Δ at a few steps in the recursive process. The final value of Δ is then 2, 4, 8 or 16 times the initial value. One must, of course, always choose $\Delta(\zeta)$ so that $\bar{u}(\zeta)$ does not change much in the interval $\Delta(\zeta)$. Numerov's procedure must be suitably applied at transition points. The radial function's normalization algorithm must also be adapted.[8]

The method described above is certainly not the only one that can yield eigenvalues and eigenfunctions of (8.38). Let us outline a few other possibilities.

When considering potentials in one dimension, we have integrated the differential equation first from $-\infty$ to zero, and then from $+\infty$ to zero. Eigenvalues were obtained by looking for values of ε for which $R_+(\varepsilon) = R_-(\varepsilon)$. Clearly one can proceed similarly here, by integrating first from zero to some intermediate point (say ζ_0), and then from $+\infty$ to ζ_0. Eigenvalues are obtained by equating at $\zeta = \zeta_0$ ratios of the derivative of $\bar{u}$ to $\bar{u}$, computed both ways.[9] Since the initial value of the wave function at infinity (in practice at large ζ) vanishes, this method makes wave function normalization simpler.

A third method of integrating (8.38) proceeds from $+\infty$ to zero. Specifically, one starts with a large value of ζ, where one lets $\bar{u} = 0$, and one applies the recurrence equation up to $\zeta = 0$. Eigenvalues are then values of ε for which $\bar{u}$ vanishes at $\zeta = 0$. As the previous method, this one is appropriate to wave function normalization. It has no problems in dealing with potentials more singular than ζ^{-1} at the origin.

---

[8] Instead of changing the value of Δ at specific points, one can perform a change of the independent variable that results in covering a larger interval where ζ is large. For example one can let $\zeta' = \ln(\zeta)$. Note that Numerov's formula can no longer be used, unless the dependent variable is also changed. See [38], [101] and [179].

[9] For instance ζ_0 can be chosen as the last zero of the first or second derivative of $\bar{u}(\zeta)$.

8.5 Phase Shifts

In the case where $\varepsilon > 0$ (in general $\varepsilon > v_\infty$), (8.38) has a continuous energy spectrum. To any ε there corresponds a function $\bar{u}(\zeta)$ that behaves as ζ^{l+1} around $\zeta = 0$, and that is oscillating as $\zeta \to \infty$. One can get a discrete approximation of $\bar{u}(\zeta)$ by using the first method for the numerical solution of (8.38) described in Sect. 8.4.

The technique of box normalization applies as in the case of discrete spectrum in one dimension. One has to find values of ε for which the function $\bar{u}(\zeta)$ satisfies the following boundary conditions:

$$\bar{u}(0) = 0 = \bar{u}(L) . \tag{8.49}$$

Here L is a length much larger than the region in which $v(\zeta)$ substantially differs from zero. Wave functions can be normalized in a box of radius L.

In the remainder of this section, we shall assume that functions $\bar{u}(\zeta)$ are defined in the whole interval $[0, \infty)$. We will analyze their asymptotic behavior for large ζ.

Let us examine (8.38) in the region where ζ is large. Inasmuch as $v(\zeta)$ goes to zero fast enough, $\bar{u}(\zeta)$ is a linear combination of $\sin(\sqrt{\varepsilon}\zeta)$ and $\cos(\sqrt{\varepsilon}\zeta)$. Such a combination can always be written as

$$\bar{u}(\zeta) = A \sin\left\{\sqrt{\varepsilon}\zeta - \frac{\pi l}{2} + \delta\right\}, \qquad \zeta \gg 1 . \tag{8.50}$$

Here A and δ are constants and $A > 0$. The reason for the introduction of the term $\pi l/2$ will soon be clear. It is important to point out that for a given potential, the value of δ is completely determined by values of ε and l. Indeed the solution of (8.38) that satisfies $\bar{u}(0) = 0$ is unique up to a multiplicative factor. Hence one can write $\delta = \delta_l(\varepsilon)$.

In the case where $v(\zeta)$ vanishes identically, (8.38) has exact solutions. Making use of results of Sect. 7.9.3 one easily shows that

$$u(\zeta) = \zeta^{-1}\bar{u}(\zeta) = \begin{cases} j_l\left(\sqrt{\varepsilon}\zeta\right) \\ n_l\left(\sqrt{\varepsilon}\zeta\right) \end{cases} \text{if } v(\zeta) = 0 . \tag{8.51}$$

Here j_l and n_l are spherical Bessel and Neumann functions. Only the function j_l is finite at $\zeta = 0$. Its asymptotic behavior for large ζ is represented by (7.199). The leading term is such that

$$\bar{u}(\zeta) \approx \sin\left\{\sqrt{\varepsilon}\zeta - \frac{\pi l}{2}\right\} \quad \text{if} \quad \begin{cases} v(\zeta) = 0 , \\ \zeta \gg 1 . \end{cases} \tag{8.52}$$

Thus when $v(\zeta) = 0$, the constants $\delta_l(\varepsilon)$ all vanish. In the general case where $v(\zeta) \neq 0$ they do not vanish. Constants $\delta_l(\varepsilon)$ represent the change in the phase of the asymptotic form of the function $\bar{u}(\zeta)$, and therefore of the radial function $u(\zeta)$, due to the presence of the potential. For this reason they are called *phase shifts.*

In Chap. 11 we will see that phase shifts have a major import in scattering theory. In fact the differential scattering cross section for a spherically symmetric potential is entirely determined by the phase shifts. We will therefore investigate how they can be computed by the numerical methods we developed ([157]; [202], Chaps. 18 and 19).

Let $v(\zeta)$ be given and pick a value of ε and l. Assume a discrete approximation $\bar{u}_n$ of the solution of (8.38) has been obtained, such that $\bar{u}_0 = 0$ and $\bar{u}_1 > 0$. This corresponds to an exact solution such that $\bar{u}(0) = 0$ and $(\mathrm{d}\bar{u}/\mathrm{d}\zeta)_{\zeta=0} > 0$, and completely determines the phase shift $\delta_l(\varepsilon)$. The latter can then be calculated from asymptotic values of $\bar{u}(\zeta)$, that is, from $\bar{u}_n$ for n large enough. From (8.50) one finds that

$$\delta_l(\varepsilon) = \frac{\pi l}{2} - \sqrt{\varepsilon}\zeta + \arctan\left\{ \frac{\sqrt{\varepsilon}\bar{u}(\zeta)}{(\mathrm{d}\bar{u}/\mathrm{d}\zeta)} \right\}. \tag{8.53}$$

The phase shift can be obtained by evaluating the right-hand side of (8.53) at $\zeta = \zeta_n$ and $\bar{u} = \bar{u}_n = \bar{u}(\zeta_n)$.[10]

Equivalently, one can obtain $\delta_l(\varepsilon)$ by writing the equation

$$\bar{u}_n = A \sin\left\{ \sqrt{\varepsilon}\zeta_n - \frac{\pi l}{2} + \delta_l(\varepsilon) \right\} \tag{8.54}$$

for two consecutive values of n and solving the resulting system of two equations for the two unknowns $A > 0$ and $\delta_l(\varepsilon)$.[11]

The formulas obtained for $\delta_l(\varepsilon)$ are, in principle, adequate. In some circumstances, however, it is worthwhile to improve on them. If l is large, the function $\bar{u}(\zeta)$ reaches its asymptotic behavior at large values of ζ only. We will develop a formula to obtain $\delta_l(\varepsilon)$ by evaluating $\bar{u}$ at not so large values of ζ.

Consider a region where the variable ζ is large enough for the potential $v(\zeta)$ to be negligible. By this we mean that $v(\zeta)$ is negligible not only when compared with the term ε in (8.38), but also when compared with $l(l+1)/\zeta^2$. In other words we assume that $v(\zeta)$ very rapidly decreases when ζ is large.

In that region, solutions of (8.38) are essentially spherical Bessel and Neumann functions multiplied by ζ. Take, in the region where $v(\zeta)$ is negligible, the following linear combination of solutions of (8.38):

$$F(\zeta) = A\sqrt{\varepsilon}\zeta \left\{ \cos\delta\, j_l\left(\sqrt{\varepsilon}\zeta\right) + \sin\delta\, n_l\left(\sqrt{\varepsilon}\zeta\right) \right\}. \tag{8.55}$$

Should one only keep the first term in the asymptotic expansion of j_l and n_l, function $F(\zeta)$ would coincide with the asymptotic form $\bar{u}(\zeta)$ given in (8.50).

---

[10] The function arctan determines the phase shift value in the interval $[0,\pi)$. Examination of signs of $\bar{u}_n$ and $(\mathrm{d}\bar{u}/\mathrm{d}\zeta)_n$ determines $\delta_l(\varepsilon)$ uniquely in the interval $[0,2\pi)$.

[11] We can also extract $\delta_l(\varepsilon)$ by evaluating $\bar{u}$ at a larger number of points [205].

Hence δ coincides with the phase shift $\delta_l(\varepsilon)$. If, on the other hand, all terms are kept in the asymptotic expansion, one finds that

$$\begin{aligned}\bar{u}(\zeta) &= F(\zeta)\\ &= A\left\{f_\nu^2\left(\sqrt{\varepsilon}\zeta\right)+g_\nu^2\left(\sqrt{\varepsilon}\zeta\right)\right\}^{1/2}\\ &\quad\times\sin\left\{\sqrt{\varepsilon}\zeta-\frac{\pi l}{2}+\delta_l(\varepsilon)-\arctan\left[\frac{g_\nu\left(\sqrt{\varepsilon}\zeta\right)}{f_\nu\left(\sqrt{\varepsilon}\zeta\right)}\right]\right\},\end{aligned}$$

where f_ν and g_ν $(\nu = l + 1/2)$ are given in (7.191) and (7.192). Rearranging one gets (Exercise 8.7)

$$\begin{aligned}\bar{u}(\zeta) &= A\left\{1+\frac{l(l+1)}{4\varepsilon\zeta^2}+O\left[\left(\sqrt{\varepsilon}\zeta\right)^{-4}\right]\right\}\\ &\quad\times\sin\left\{\sqrt{\varepsilon}\zeta-\frac{\pi l}{2}+\delta_l(\varepsilon)+\frac{l(l+1)}{2\sqrt{\varepsilon}\zeta}+O\left[\left(\sqrt{\varepsilon}\zeta\right)^{-3}\right]\right\}.\end{aligned}\tag{8.56}$$

Comparison of this asymptotic form with values of $\bar{u}_n$ for large n yields, following methods similar to the ones just described, a determination of $\delta_l(\varepsilon)$. The asymptotic form (8.56) applies earlier than (8.50), that is, starting with smaller values of ζ. Looking at the sine's argument shows that the asymptotic behavior of $\bar{u}(\zeta)$ really begins at values of ζ for which the correction of order ζ^{-1} is small, that is, where

$$\zeta \gg \frac{l(l+1)}{2\sqrt{\varepsilon}}.\tag{8.57}$$

Table 8.1. Phase shifts associated with the spherical square well

ε	l	$\delta(8.56)$	$\delta(8.50)$	$\delta(\text{exact})$
0.2	0	3.155 88	3.155 88	3.155 88
	1	2.982 70	3.023 18	2.982 72
	3	0.000 40	0.239 97	0.000 21
1	0	3.145 53	3.145 53	3.145 53
	1	2.384 99	2.425 82	2.385 01
	3	0.062 53	0.301 39	0.062 33
5	0	1.840 31	1.840 31	1.840 31
	1	1.869 12	1.909 44	1.869 14
	3	1.654 64	1.896 42	1.654 44
20	0	1.006 02	1.006 02	1.006 02
	1	1.100 10	1.139 36	1.100 12
	3	1.027 06	1.263 29	1.026 88

Table 8.1 displays results of a numerical calculation of phase shifts for the spherical square well defined as

$$v(\zeta) = \begin{cases} -5 & \text{if } \zeta < 2\,, \\ 0 & \text{if } \zeta > 2\,. \end{cases}$$

Equations (8.50) and (8.56) are used at values of ζ such that $\sqrt{\varepsilon}\zeta \approx 25$. If $l = 0$, (8.50) and (8.56) are equivalent and coincide with the exact form of $\bar{u}(\zeta)$. Values shown in Table 8.1 also coincide. If $l \neq 0$, values computed from (8.56) are much more precise. Exact values of the phase shifts are obtained in Sect. 11.7.

Exercises

8.1. Let $F_n = F(n\Delta)$, where Δ is small. Use the Taylor series expansion to show that

$$\frac{F_{n+1} - 2F_n + F_{n-1}}{\Delta^2} = \left(\frac{\mathrm{d}^2F}{\mathrm{d}\xi^2}\right)_{\xi_n} + \frac{\Delta^2}{12}\left(\frac{\mathrm{d}^4F}{\mathrm{d}\xi^4}\right)_{\xi_n} + O(\Delta^4)\,,$$
$$\frac{-F_{n+2} + 8F_{n+1} - 8F_{n-1} + F_{n-2}}{12\Delta} = \left(\frac{\mathrm{d}F}{\mathrm{d}\xi}\right)_{\xi_n} + O(\Delta^4)\,.$$

8.2. Let $v(\xi)$ be a function which for large $|\xi|$ goes to zero faster than $|\xi|^{-1}$. Show that the asymptotic form of the solutions of (8.12) is then given by (8.14).

8.3. Write a computer program that computes the Hamiltonian's eigenvalues and eigenfunctions for a particle in one dimension, following the method developed in Sect. 8.2.

8.4. Let $v(\xi)$ be a one-dimensional potential. One can show that if $v(\xi)$ is an even function, normalizable solutions of the eigenvalue equation of H are either even functions $[\psi(\xi) = \psi(-\xi)]$ or odd functions $[\psi(\xi) = -\psi(-\xi)]$.
a) Show that if $\psi(\xi)$ is an odd function, then $\psi(0) = 0$.
b) Show that if $\psi(\xi)$ is an even function with a continuous derivative, then $(d\psi|d\xi)_0 = 0$.
c) Use these observations to outline an algorithm that looks for the Hamiltonian's eigenvalues for an even potential, by means of numerically integrating the eigenvalue equation from $\xi = 0$ to $\xi \to \infty$ ([158], Sect. 12.3).

8.5. Let $\bar{u}(\zeta)$ be a solution of (8.38).
a) Show that if $w(\zeta)$ is defined so that $\bar{u}(\zeta) = \zeta^{l+1}w(\zeta)$, then $w(\zeta)$ satisfies the equation

$$\frac{\mathrm{d}^2w}{\mathrm{d}\zeta^2} + \frac{2(l+1)}{\zeta}\frac{\mathrm{d}w}{\mathrm{d}\zeta} + [\varepsilon - v(\zeta)]\,w(\zeta) = 0\,.$$

b) Let

$$w(\zeta) = \sum_{k=0}^{\infty} a_k \zeta^k, \qquad v(\zeta) = \sum_{j=-1}^{\infty} b_j \zeta^j,$$

where a_k and b_j are constants. Show that the a_k satisfy the recurrence equation (8.46).

8.6. Write a computer program that computes eigenvalues and functions $\bar{u}(\zeta)$ associated with the Hamiltonian of a particle in three dimensions, following one of the methods developed in Sect. 8.4 ([158], Sects. 12.4 and 12.5; [202], Chap. 13).

8.7. Derive (8.56) from (8.55) and asymptotic properties of spherical Bessel and Neumann functions.

9 The Central-Field Model

The search for eigenvalues and eigenfunctions of an atom's Hamiltonian is a very complex problem. The central-field model simplifies it while remaining fairly close to physical reality. The model assumes each electron moves in a spherically symmetric potential due to the nucleus and all other electrons. Moreover, it introduces in a simple way the notion of electronic spin and the Pauli exclusion principle. The model leads to the construction of the periodic table of elements. Although it has limited quantitative value, it can account for basic properties of atomic structure.

9.1 The Hamiltonian of an Atom

An atom is a complex quantum system. It is made up of a nucleus of Z protons and N neutrons, and of a "cloud" of Q electrons.[1] Strictly speaking, it is a system of $Z + N + Q$ particles.

Energies associated with electronic motion range from a few eV (for outer electrons) to a few tens of KeV (for inner electrons). Energies associated with nuclear motion, on the other hand, are of the order of MeV. Moreover, the nuclear radius is about 10^5 times smaller than the atomic radius. Thus our first approximation will consist in viewing the nucleus as a single particle of mass m_{N} and charge $-Zq_e$.[2] We can reasonably think that this approximation is adequate if energies are restricted to values substantially lower than 1 MeV.

Thus we are dealing with a system of $Q + 1$ particles. In addition to the time coordinate, wave functions will depend on $3Q + 3$ coordinates that can be denoted by $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, …, $\boldsymbol{r}_Q$ for electrons and $\boldsymbol{r}_{\mathrm{N}}$ for the nucleus.

The system's Hamiltonian, the operator corresponding to energy, should represent the total energy of the atom. Consider first the kinetic energy of the $Q + 1$ particles. Generalizing from the case of a single particle, one can write the Hamiltonian's kinetic energy terms as

---

[1] We mean "atom" in a broad sense, so as to cover ions as well as neutral atoms. In principle positive ions exist for all Q such that $0 \leq Q < Z$. There seems to be stable negative ions for $Q = Z + 1$ only.

[2] Recall that q_e, the electron's charge, is negative, so that $-Zq_e$ is positive.

L. Marchildon, *Quantum Mechanics*

$$\frac{1}{2m_{\mathrm{N}}}\boldsymbol{P}_{\mathrm{N}}^2+\frac{1}{2m}\sum_{i=1}^{Q}\boldsymbol{P}_i^2 \rightarrow -\frac{\hbar^2}{2m_{\mathrm{N}}}\nabla_{r_{\mathrm{N}}}^2-\frac{\hbar^2}{2m}\sum_{i=1}^{Q}\nabla_{r_i}^2 . \tag{9.1}$$

Here m stands for the mass of each of the Q electrons. The coordinate representation was used in the right-hand side, where symbols $\nabla_{r_{\mathrm{N}}}^2$ and $\nabla_{r_i}^2$ represent the Laplacian operator relative to coordinates $\boldsymbol{r}_{\mathrm{N}}$ and $\boldsymbol{r}_i$. Explicitly

$$\nabla_{r_i}^2 \equiv \frac{\partial^2}{\partial x_i^2}+\frac{\partial^2}{\partial y_i^2}+\frac{\partial^2}{\partial z_i^2} . \tag{9.2}$$

Let us now turn to the electrostatic potential energy. It comes from the interaction of each electron with the nucleus, and of each electron with each other electron. In the coordinate representation it can thus be written as

$$\sum_{i=1}^{Q}\frac{1}{4\pi\varepsilon_0}\frac{(-Zq_e)q_e}{|\boldsymbol{r}_i-\boldsymbol{r}_{\mathrm{N}}|}+\sum_{i<j}\frac{1}{4\pi\varepsilon_0}\frac{q_e^2}{|\boldsymbol{r}_i-\boldsymbol{r}_j|} . \tag{9.3}$$

The inequality $i < j$ appearing under the summation symbol means that the sum is carried out on all electron pairs. It is convenient to introduce a positive number e defined so that[3]

$$e^2=\frac{q_e^2}{4\pi\varepsilon_0} . \tag{9.4}$$

The electrostatic potential energy then becomes

$$\sum_{i=1}^{Q}\frac{(-Ze^2)}{|\boldsymbol{r}_i-\boldsymbol{r}_{\mathrm{N}}|}+\sum_{i<j}\frac{e^2}{|\boldsymbol{r}_i-\boldsymbol{r}_j|} . \tag{9.5}$$

The kinetic energy and the electrostatic potential energy are the two most important terms of an atom's Hamiltonian. But they are not the only ones. At this stage we briefly mention three other contributions.

The first one is magnetic. An electron moving in an atom has in general orbital angular momentum. Hence according to (7.137), it also has orbital magnetic moment. Besides, each electron has spin and intrinsic magnetic moment. So is it with the nucleus. The interaction of all these moments produces a magnetic energy that must be represented in the exact Hamiltonian. The main contribution comes from the interaction of an electron's spin magnetic moment with the same electron's orbital magnetic moment. It is particularly important in large atoms. It will be investigated in Chaps. 15 and 18.

The second contribution is relativistic. As we pointed out, an electron's energy is typically between a few eV and a few tens of KeV. This means that

---

[3] The quantity e has dimensions of $(\mathrm{kg\ m^3\ s^{-2}})^{1/2}$. Translated into $(\mathrm{g\ cm^3\ s^{-2}})^{1/2}$, its numerical value coincides with the fundamental charge in Gaussian units.

it can amount to a few percent of the relativistic energy of an electron at rest, which is about 511 KeV. The exact Hamiltonian must therefore include relativistic corrections associated with electronic motion. Here again the effect is more important in large atoms. We will come back to it in Chap. 15.

The third contribution is quantitatively small but conceptually very complicated. It comes from quantization of the electromagnetic field. Strictly speaking, the field in an atom should also be treated by quantum-mechanical methods. The appropriate formalism is, however, much more complex than the quantum mechanics of a finite number of particles.

These three additional contributions have the effect not only to add terms to the Hamiltonian, but also to change the structure of the state space. The modification is rather simple in the case of spin and some relativistic effects, but major in the case of field quantization.

The approximation we are going to make consists in including in the Hamiltonian only kinetic and electrostatic potential energy terms. In the coordinate representation we can thus write

$$H=-\frac{\hbar^2}{2m_{\mathrm{N}}}\nabla^2_{r_{\mathrm{N}}}-\frac{\hbar^2}{2m}\sum_{i=1}^{Q}\nabla^2_{r_i}-Ze^2\sum_{i=1}^{Q}\frac{1}{|\boldsymbol{r}_i-\boldsymbol{r}_{\mathrm{N}}|}+e^2\sum_{i<j}\frac{1}{|\boldsymbol{r}_i-\boldsymbol{r}_j|}\,. \tag{9.6}$$

It is very convenient to switch to relative and center-of-mass coordinates. Let us define

$$\boldsymbol{R}=\frac{1}{M}\left[m_{\mathrm{N}}\boldsymbol{r}_{\mathrm{N}}+\sum_{i=1}^{Q}m\boldsymbol{r}_i\right], \tag{9.7}$$

$$\boldsymbol{r}_{i\mathrm{N}}=\boldsymbol{r}_i-\boldsymbol{r}_{\mathrm{N}}\,. \tag{9.8}$$

Here $M=m_{\mathrm{N}}+Qm$ is the atom's total mass. Performing this change of variables we easily find (Exercise 9.1) that (9.6) becomes

$$\begin{aligned}H=&-\frac{\hbar^2}{2M}\nabla^2_R-\frac{\hbar^2}{2\mu}\sum_{i=1}^{Q}\nabla^2_{r_{i\mathrm{N}}}-\frac{\hbar^2}{2m_{\mathrm{N}}}\sum_{i\neq j}\boldsymbol{\nabla}_{r_{i\mathrm{N}}}\cdot\boldsymbol{\nabla}_{r_{j\mathrm{N}}}\\&-Ze^2\sum_{i=1}^{Q}\frac{1}{r_{i\mathrm{N}}}+e^2\sum_{i<j}\frac{1}{|\boldsymbol{r}_{i\mathrm{N}}-\boldsymbol{r}_{j\mathrm{N}}|}\,.\end{aligned} \tag{9.9}$$

The quantity μ, called the *reduced mass* of the electron, is defined as

$$\mu=\frac{m\,m_{\mathrm{N}}}{m+m_{\mathrm{N}}}\,. \tag{9.10}$$

Let us now write down the Hamiltonian's eigenvalue equation. We have

$$-\frac{\hbar^2}{2M}\nabla_R^2\Psi-\frac{\hbar^2}{2\mu}\sum_{i=1}^{Q}\nabla_{r_{i\mathrm{N}}}^2\Psi-\frac{\hbar^2}{2m_\mathrm{N}}\sum_{i\neq j}\boldsymbol{\nabla}_{r_{i\mathrm{N}}}\cdot\boldsymbol{\nabla}_{r_{j\mathrm{N}}}\Psi$$
$$-Ze^2\sum_{i=1}^{Q}\frac{1}{r_{i\mathrm{N}}}\Psi+e^2\sum_{i<j}\frac{1}{|\boldsymbol{r}_{i\mathrm{N}}-\boldsymbol{r}_{j\mathrm{N}}|}\Psi=E\Psi\,, \tag{9.11}$$

where $\Psi=\Psi(\boldsymbol{R},\boldsymbol{r}_{1\mathrm{N}},\ldots,\boldsymbol{r}_{Q\mathrm{N}})$. We look for a solution of the form

$$\Psi\left(\boldsymbol{R},\boldsymbol{r}_{1\mathrm{N}},\ldots,\boldsymbol{r}_{Q\mathrm{N}}\right)=\Psi_{\mathrm{CM}}\left(\boldsymbol{R}\right)\psi(\boldsymbol{r}_{1\mathrm{N}},\ldots,\boldsymbol{r}_{Q\mathrm{N}}). \tag{9.12}$$

Substituting this expression in the eigenvalue equation, one can check that the latter separates in the following way:

$$-\frac{\hbar^2}{2\mu}\sum_{i=1}^{Q}\nabla_{r_{i\mathrm{N}}}^2\psi-\frac{\hbar^2}{2m_\mathrm{N}}\sum_{i\neq j}\boldsymbol{\nabla}_{r_{i\mathrm{N}}}\cdot\boldsymbol{\nabla}_{r_{j\mathrm{N}}}\psi$$
$$-Ze^2\sum_{i=1}^{Q}\frac{1}{r_{i\mathrm{N}}}\psi+e^2\sum_{i<j}\frac{1}{|\boldsymbol{r}_{i\mathrm{N}}-\boldsymbol{r}_{j\mathrm{N}}|}\psi=\bar{E}\psi\,, \tag{9.13}$$

$$-\frac{\hbar^2}{2M}\nabla_R^2\Psi_{\mathrm{CM}}=(E-\bar{E})\Psi_{\mathrm{CM}}\,. \tag{9.14}$$

Equation (9.14) coincides with the eigenvalue equation of a free-particle Hamiltonian, whose solution is immediate:

$$\Psi_{\mathrm{CM}}(\boldsymbol{R})=\exp\{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{R}\}\,, \tag{9.15}$$

where

$$\boldsymbol{k}\cdot\boldsymbol{k}=\frac{2M(E-\bar{E})}{\hbar^2}\,. \tag{9.16}$$

Equation (9.13), on the other hand, is complex, but we are going to simplify it a little. One can reasonably think that terms like $\boldsymbol{\nabla}_{r_{i\mathrm{N}}}\cdot\boldsymbol{\nabla}_{r_{j\mathrm{N}}}\psi$ are roughly of the same magnitude as terms like $\nabla^2_{r_{i\mathrm{N}}}\psi$. But typically, the coefficient $(m_\mathrm{N})^{-1}$ is 10^4 or 10^5 times smaller than μ^{-1}. Hence we will neglect terms $\boldsymbol{\nabla}_{r_{i\mathrm{N}}}\cdot\boldsymbol{\nabla}_{r_{j\mathrm{N}}}\psi$.[4] Equation (9.13) thus becomes

$$-\frac{\hbar^2}{2\mu}\sum_{i=1}^{Q}\nabla_{r_{i\mathrm{N}}}^2\psi-Ze^2\sum_{i=1}^{Q}\frac{1}{r_{i\mathrm{N}}}\psi+e^2\sum_{i<j}\frac{1}{|\boldsymbol{r}_{i\mathrm{N}}-\boldsymbol{r}_{j\mathrm{N}}|}\psi=\bar{E}\psi. \tag{9.17}$$

This is the eigenvalue equation of the operator

$$H=-\frac{\hbar^2}{2\mu}\sum_{i=1}^{Q}\nabla_{r_{i\mathrm{N}}}^2-Ze^2\sum_{i=1}^{Q}\frac{1}{r_{i\mathrm{N}}}+e^2\sum_{i<j}\frac{1}{|\boldsymbol{r}_{i\mathrm{N}}-\boldsymbol{r}_{j\mathrm{N}}|}\,, \tag{9.18}$$

---

[4] The effect of such terms can be evaluated by the methods we will develop in Chap. 10.

which, in the limit of our approximations, we shall take as the atom's Hamiltonian. The fundamental problem of atomic physics consists in finding the solutions of (9.17).

9.2 The Hydrogen Atom and Hydrogen-Like Ions

The hydrogen atom is made up of a nucleus of charge $-q_e$ and an electron of charge q_e. Depending on the isotope the nucleus can have one proton, one proton and one neutron (deuterium), or one proton and two neutrons (tritium). Hydrogen-like ions are atoms with Z protons from which $Z-1$ electrons have been removed. They have just one electron left.

The analysis of Sect. 9.1 applies very simply to atoms with only one electron. In this case there are no terms like $\boldsymbol{\nabla}_{r_{i\mathrm{N}}} \cdot \boldsymbol{\nabla}_{r_{j\mathrm{N}}}\psi$, so that this additional approximation is not required. The sum over $i < j$ is void. If we simply denote by $\boldsymbol{r}$ the coordinate of the electron relative to the nucleus, (9.13) can be written as

$$-\frac{\hbar^2}{2\mu}\nabla^2\psi(\boldsymbol{r}) - \frac{Ze^2}{r}\psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}), \tag{9.19}$$

where E now stands for what used to be denoted by $\bar{E}$. Inasmuch as the Hamiltonian is given solely by nonrelativistic kinetic and electrostatic potential energy terms, this equation is exact.

Formally, (9.19) is identical with the Hamiltonian's eigenvalue equation for a particle of mass μ in the potential $V(r) = -Ze^2/r$. In Sect. 7.5 we have found the discrete-spectrum exact solutions of this equation. We saw that solutions can be indexed by three quantum numbers n, l and m, so that

$$\psi_{nlm}(\boldsymbol{r}) = u_{nl}(r)Y_{lm}(\theta, \phi).$$

Functions $Y_{lm}(\theta, \phi)$ and $u_{nl}(r)$ were given in Sects. 7.9.1 and 7.9.2. The energy E depends only on n and is equal to

$$E_n = -\frac{\mu Z^2 e^4}{2\hbar^2 n^2}. \tag{9.20}$$

This is the formula Bohr obtained in 1913 and which was given in (1.2) for the case where $Z = 1$.

Recall that $\mu = m m_{\mathrm{N}}(m + m_{\mathrm{N}})^{-1}$, where m is the electron's mass and m_{N} the nucleus's mass. Even though μ is very close to m, values of E_n for a given Z depend on the mass of the nucleus. Spectral lines of two distinct isotopes of a hydrogen-like ion are therefore slightly separated (Exercise 9.2).

The constant $a_0 = \hbar^2/\mu e^2$ appearing in the expression of $u_{nl}(r)$ also depends on the nucleus's mass. The *Bohr radius*, denoted by a_0^∞, is defined as the value of a_0 that corresponds to an infinite-mass nucleus, that is,

$$a_0^\infty = \lim_{m_\mathrm{N}\to\infty} a_0 = \frac{\hbar^2}{me^2} = 0.529\,2\times 10^{-10}\,\mathrm{m}. \tag{9.21}$$

One also defines the *rydberg* (Ry) as a unit of energy equal to the absolute value of the lowest energy of hydrogen for an infinite-mass nucleus. Hence

$$1\,\mathrm{rydberg} = \frac{me^4}{2\hbar^2} = \frac{e^2}{2a_0^\infty} = 13.61\,\mathrm{eV}. \tag{9.22}$$

The first three energy levels of hydrogen are shown in Fig. 9.1. Also indicated are triplets (nlm) labeling wave functions in each level.

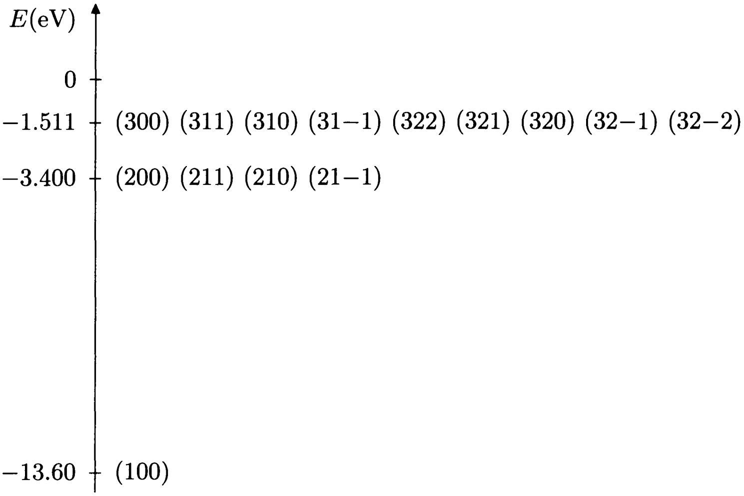

Fig. 9.1. First three energy levels of the hydrogen atom and quantum numbers (nlm) of associated states

Only the lowest energy of hydrogen and hydrogen-like ions is not degenerate.[5] Since $0 \le l < n$ and $-l \le m \le l$, one easily checks that the degree of degeneracy of energy E_n is equal to

$$\sum_{l=0}^{n-1} (2l+1) = n^2. \tag{9.23}$$

9.3 The Central Field

Let us now come back to the general case, i.e. the Hamiltonian's eigenvalue equation for an atom with Q electrons and Z protons:

---

[5] The lowest energy is degenerate, however, if spin is taken in account.

$$-\frac{\hbar^2}{2m}\sum_{i=1}^{Q}\nabla_i^2\psi - Ze^2\sum_{i=1}^{Q}\frac{1}{r_i}\psi + e^2\sum_{i<j}\frac{1}{r_{ij}}\psi = E\psi. \tag{9.24}$$

Here we have simplified notations used in (9.17), writing m for μ, r_i for $r_{i\mathrm{N}}$, ∇_i^2 for $\nabla^2_{r_{i\mathrm{N}}}$, r_{ij} for $|\boldsymbol{r}_{i\mathrm{N}} - \boldsymbol{r}_{j\mathrm{N}}|$ and E for $\bar{E}$. The wave function $\psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_Q)$ is the solution of a partial differential equation with $3Q$ independent variables.

The basic idea of the central-field model is simple. One assumes that the action, on a given electron, of the nucleus and all other electrons can approximately be represented by a spherically symmetric potential $V(r)$, which only depends on the coordinate of the electron considered.

Insofar as properties of electron k are entirely determined by the potential $V(r_k)$, there is no correlation between the position of k and the position of other electrons. The wave function can therefore be written as[6]

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_Q) = \psi_1(\boldsymbol{r}_1)\psi_2(\boldsymbol{r}_2)\ldots\psi_Q(\boldsymbol{r}_Q). \tag{9.25}$$

Each $\psi_k(\boldsymbol{r}_k)$ represents the wave function of a single electron. Following the central-field hypothesis, these functions satisfy an eigenvalue equation associated with the potential $V(r_k)$, that is,

$$-\frac{\hbar^2}{2m}\nabla_k^2\psi_k(\boldsymbol{r}_k) + V(r_k)\psi_k(\boldsymbol{r}_k) = \lambda_k\psi_k(\boldsymbol{r}_k). \tag{9.26}$$

For the time being we do not know the exact form of $V(r_k)$. From the simple fact that it is spherically symmetric, however, we know that in the discrete spectrum solutions $\psi_k(\boldsymbol{r}_k)$ have the form

$$\psi_k(\boldsymbol{r}_k) = u_{ln}(r_k)Y_{lm}(\theta_k, \phi_k).$$

Constant λ_k depends on indices n and l, but not on m.[7] It can be written as λ_{nl}. Function $u_{nl}(r_k)$ satisfies the radial equation associated with the potential $V(r_k)$.

The wave function $\psi_k(\boldsymbol{r}_k)$ is called an *orbital.* Its shape mostly depends on indices n and l. Orbitals with the same n and l and different m are called *equivalent.* The set of all orbitals with the same n and l makes up an *electronic shell.*[8] In spectroscopy orbitals are labeled by a number (the value of n) and by a lower case letter. Letters s, p, d, f, g, ... are respectively associated

---

[6] In Chap. 17 we will see how to alter this expression so as to take into account the identity of all electrons.

[7] Let E_I^k be the ionization energy of electron k, that is, the energy required to extract it. Although λ_k and $-E_I^k$ are connected, they do not coincide. Their relation will be made more precise in Chaps. 16 and 17. The central-field model nonetheless assumes that λ_k represents well enough the total energy of electron k.

[8] The name "shell" is sometimes applied to the set of all orbitals with the same n. The set of all orbitals with the same n and l is then called a *subshell.*

with values 0, 1, 2, 3, 4, ... of the quantum number l. So we have orbitals $1s$, $2p$, $3d$, $3p$, and so on.

The existence of the potential $V(r_k)$ is the basic postulate of the central-field model. Two additional hypotheses are nonetheless necessary to its successful application. They are the existence of electronic spin and the Pauli exclusion principle.

Let us first consider the electronic spin. We know that an electron is a spin 1/2 system. Hence a complete description of an electronic state not only requires the spatial wave function $\psi_k(\boldsymbol{r}_k)$, but also a vector in a two-dimensional complex space. That vector can be denoted by $\chi_k(k)$. This means that an electron described by the spatial wave function ψ_k can be in two different, even orthogonal, states.

The *Pauli exclusion principle* will be given a general formulation in Chap. 17. For the present discussion it can be expressed simply by stipulating that two electrons in a given atom are necessarily in orthogonal states. This means that if two electrons k and k' $(k \neq k')$ are described by wave functions indexed by triplets (nlm) and $(n'l'm')$, then either the triplets are different, or the triplets are identical, the spin vectors in the latter case being orthogonal.

The existence of electronic spin and the Pauli exclusion principle have the effect of limiting the number of electrons in an orbital or a shell. For given n and l, there are $2l+1$ values of the quantum number m; and for a given m, there are two spin vectors. The maximum number of electrons in a shell nl is thus equal to $2(2l+1)$. Hence there can be at most two electrons in an s shell, six electrons in a p shell, ten electrons in a d shell, fourteen electrons in an f shell, and so on.

The central-field model is now completely set up. Of course, using the model to obtain quantitative results requires to determine the potential $V(r_k)$. Phenomenological considerations can yield an approximate form of the potential. As a first approximation though, we will rather use the form of $V(r_k)$ derived from the Thomas–Fermi equation.

9.4 The Thomas–Fermi Equation

The Thomas–Fermi equation provides a way to determine the electrostatic potential $\phi(r)$ everywhere in an atom. We assume here that the number of electrons is equal to the number of protons. Let $\varrho(r)$ be the total charge density, due to the nucleus and each of the $Q = Z$ electrons. Following the laws of electrostatics

$$\nabla^2 \phi = -\frac{\varrho}{\varepsilon_0} . \tag{9.27}$$

This equation yields ϕ if ϱ is known. But neither ϱ nor ϕ is known a priori.

The potential energy $V(r)$ of an electron at the point r is negative. We assume it is equal to $q_e\phi(r)$. This hypothesis is not entirely correct, since $\phi(r)$ represents the electrostatic potential due to the nucleus and all electrons, including the one whose potential energy we are looking for. But we can reasonably think that the larger Z is, the better the assumption.

The Thomas–Fermi equation makes use of the notion of volume density of states of a free particle. So we will first take a closer look at that density. Consider a cubic box of length $2L$ with periodic boundary conditions. We denote by $n(E)$ the volume density of states, with energy lower than E, of a free particle.

In the box the free-particle Hamiltonian's eigenfunctions are given by

$$\psi(\boldsymbol{r}) = (2L)^{-3/2}\exp\{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}\}. \tag{9.28}$$

Here,

$$k_x = \frac{\pi n_x}{L}, \qquad k_y = \frac{\pi n_y}{L}, \qquad k_z = \frac{\pi n_z}{L}, \tag{9.29}$$

where n_x, n_y and n_z are integers. The Hamiltonian's eigenvalues are equal to

$$E = \frac{\hbar^2}{2m}\boldsymbol{k}\cdot\boldsymbol{k} = \frac{\hbar^2\pi^2}{2mL^2}\boldsymbol{n}\cdot\boldsymbol{n}. \tag{9.30}$$

To any triplet (n_x, n_y, n_z) there corresponds a free-particle state. States whose energy is lower than a given value E are those whose corresponding triplets satisfy the inequality

$$\left(n_x^2 + n_y^2 + n_z^2\right)^{1/2} < \frac{L}{\pi\hbar}(2mE)^{1/2} \equiv r(E). \tag{9.31}$$

There is, on the average, one point with integral coordinates per unit volume. Hence the number of states with energy lower than E is equal to the volume $V(E)$ of a sphere of radius $r(E)$:

$$V(E) = \frac{4L^3}{3\pi^2\hbar^3}(2mE)^{3/2}. \tag{9.32}$$

Finally, the volume density of states with energy lower than E is equal to the number of states divided by the volume of the box, that is,

$$n(E) = \frac{1}{(2L)^3}V(E) = \frac{1}{6\pi^2\hbar^3}(2mE)^{3/2}. \tag{9.33}$$

Let us come back to Thomas–Fermi. An electron's kinetic energy at r cannot be higher than $-V(r) = -q_e\phi(r)$ for, otherwise, the electron's total energy would be positive and it would escape. Thomas and Fermi assumed electrons occupy all available negative-energy states.[9] More precisely they

---

[9] The maximum kinetic energy is called the *Fermi energy* and is denoted by E_F. Likewise k_F denotes the length of the vector $\boldsymbol{k}$ associated with E_F through (9.30). Note that E_F et k_F depend on $\boldsymbol{r}$, i.e. on position inside the atom.

postulated that the electronic density at r, given (for $r \neq 0$) by $\varrho(r)/q_e$, is equal to twice the free-particle volume density of states with energy lower than $-q_e\phi(r)$. This means that

$$\frac{1}{q_e}\varrho(r) = \frac{2}{6\pi^2\hbar^3}\left\{-2mq_e\phi(r)\right\}^{3/2}. \tag{9.34}$$

Making use of (9.34) and (9.27), one can eliminate $\varrho(r)$. One thus obtains

$$\nabla^2\phi = \frac{-q_e}{3\pi^2\hbar^3\varepsilon_0}(-2mq_e\phi)^{3/2}. \tag{9.35}$$

It is convenient to rewrite (9.35) in terms of dimensionless variables. Let $a_0 = \hbar^2/me^2$ and define

$$\Phi = \frac{-4\pi\varepsilon_0 r\phi}{Zq_e}, \qquad \xi = \frac{4r}{a_0}\left(\frac{2Z}{9\pi^2}\right)^{1/3}. \tag{9.36}$$

In this way we get (Exercise 9.5)

$$\frac{\mathrm{d}^2\Phi}{\mathrm{d}\xi^2} = \xi^{-1/2}\Phi^{3/2}. \tag{9.37}$$

This is the *Thomas–Fermi* equation. It is a second-order differential equation for one independent and one dependent variable. If $r \to \infty$, it is clear that $\phi(r) \to 0$ faster than r^{-1}. If $r \to 0$, on the other hand, $\phi(r) \to -Zq_e/4\pi\varepsilon_0 r$, because then the potential comes from the nucleus only. This entails that $\Phi(0) = 1$ and $\Phi(\infty) = 0$. With these conditions, the solution of the Thomas–Fermi equation is unique.

The Thomas–Fermi equation is universal, in the sense that it does not depend on Z. Thus it is the same for all atoms. Although it cannot be solved exactly, it can be solved numerically. Values of the Thomas–Fermi function are tabulated, for example in [214], p. 323. Latter [142] gave the following approximate representation:

$$\begin{aligned}\Phi(\xi) = \Big[&1 + 0.027\,47\,\xi^{1/2} + 1.243\,\xi - 0.148\,6\,\xi^{3/2} \\ &+ 0.230\,2\,\xi^2 + 0.007\,298\,\xi^{5/2} + 0.006\,944\,\xi^3\Big]^{-1}.\end{aligned} \tag{9.38}$$

Let us now examine more closely the assumptions from which the Thomas–Fermi equation was derived. The factor 2 that multiplies the volume density of states clearly comes from the electronic spin: there are two electronic states for each spatial wave function. The hypothesis stating that all states with kinetic energy lower than $-q_e\phi(r)$ are occupied corresponds to the use of a zero-temperature Fermi–Dirac statistical distribution. Note that the temperature associated with 1 eV is about 11 600 K. Room temperature is therefore much lower than the temperature associated with electronic energies, of the order of several eV.

Finally, let us turn to the hypothesis identifying the electronic density to a free-particle density. It is justified if the distance over which the potential changes appreciably is much larger than the radius of a sphere that contains on the average one electron, that is, if

$$\phi(r)\left|\frac{\mathrm{d}\phi}{\mathrm{d}r}\right|^{-1} \gg \left(\frac{q_e}{\varrho(r)}\right)^{1/3} . \tag{9.39}$$

In other words

$$\phi(r) \gg \left|\frac{\mathrm{d}\phi}{\mathrm{d}r}\right| \frac{(3\pi^2)^{1/3}\hbar}{\left[-2mq_e\phi(r)\right]^{1/2}} . \tag{9.40}$$

In terms of variables Φ and ξ, this inequality is written as (Exercise 9.5)

$$\left(\frac{\Phi}{\xi}\right)^{3/2} \gg \left(\frac{4\pi}{Z}\right)^{1/3} \left|\frac{\mathrm{d}}{\mathrm{d}\xi}\left(\frac{\Phi}{\xi}\right)\right| . \tag{9.41}$$

The larger Z is, the better the inequality is satisfied. It is nonetheless remarkable that the Thomas–Fermi potential provides a good approximation even for small atoms.

For any value of Z, the electrostatic potential $\phi(r)$ and the potential energy $q_e\phi(r)$ of an electron at r can be obtained from tabulated values of $\Phi(\xi)$ or from the representation (9.38). This provides a good initial approximation for the spherically symmetric potential $V(r)$ of a neutral atom. In Sect. 9.6 we will see how to pick a starting potential for ions.[10]

9.5 The Periodic Table of Elements

Let us now come back to the central-field model. Orbitals, or one-electron wave functions, have the form

$$\psi(\boldsymbol{r}) = u_{nl}(r)Y_{lm}(\theta, \phi) ,$$

where $u_{nl}(r)$ satisfies the radial equation for the spherically symmetric potential $V(r)$, that is,

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left\{r^2\frac{\mathrm{d}}{\mathrm{d}r}u_{nl}(r)\right\} - \left\{\frac{l(l+1)}{r^2} + \frac{2m}{\hbar^2}\left[V(r) - \lambda_{nl}\right]\right\} u_{nl}(r) = 0. \tag{9.42}$$

An *electronic configuration* consists in putting each electron of an atom in a given electronic shell. In general an electronic configuration involves

---

[10] The Thomas–Fermi equation can also be adapted to ions ([59], Sects. 7.8 and 7.16) but then the function Φ looses its universal character.

several distinct quantum states, corresponding to different values of magnetic quantum numbers m_i and spin quantum numbers m_{s_i} allowed by the Pauli exclusion principle. In the central-field model, all quantum states associated with a given configuration have the same energy.

The purpose of this section is to determine the lowest-energy electronic configuration of a neutral atom. As a starting point, we assume the central field is given in terms of the Thomas–Fermi potential, that is,

$$V(r) = q_e \phi = -\frac{Ze^2}{r} \varPhi . \tag{9.43}$$

Let us write

$$r = \left(\frac{9\pi^2}{2Z} \right)^{1/3} \frac{a_0}{4} \xi = a_0 \zeta , \qquad \lambda_{nl} = \frac{\hbar^2}{2ma_0^2} \varepsilon_{nl} . \tag{9.44}$$

The radial equation then becomes

$$\frac{1}{\zeta^2} \frac{\mathrm{d}}{\mathrm{d}\zeta} \left\{ \zeta^2 \frac{\mathrm{d}}{\mathrm{d}\zeta} u_{nl}(\zeta) \right\} + \left\{ -\frac{l(l+1)}{\zeta^2} + \frac{2Z}{\zeta} \varPhi(1.1295\, Z^{1/3} \zeta) + \varepsilon_{nl} \right\} u_{nl}(\zeta) = 0 . \tag{9.45}$$

For each value of Z, this equation is easily solved by the numerical methods developed in Chap. 8.[11] Radial functions $u_{nl}(r)$ and electronic energies λ_{nl} are thus obtained for any Z.

Quite independently of the value of Z, one finds that electronic energies can be ordered increasingly as [142]:[12]

$$\lambda_{1s} < \lambda_{2s} < \lambda_{2p} < \lambda_{3s} < \lambda_{3p} < \lambda_{3d} < \lambda_{4s} < \cdots \tag{9.46}$$

The $1s$ orbital corresponds to the lowest energy. As pointed out in Sect. 9.3, it has room for two electrons. Lowest-energy configurations, for hydrogen and helium, are thus given respectively by $1s$ and $1s^2$. The notation $1s^2$ means that there are two electrons in the $1s$ orbital.

The $2s$ orbital has higher energy than the $1s$. It also has room for two electrons. Hence lowest-energy configurations, for lithium and beryllium, are given by $1s^2 2s$ and $1s^2 2s^2$.

Next come, in increasing order of energy, $2p$ orbitals. They have room for six electrons. Lowest-energy configurations of boron, carbon, nitrogen,

---

[11] The change of variable $\bar{u}_{nl}(\zeta) = \zeta u_{nl}(\zeta)$ transforms (9.45) so that the Numerov method applies. Examples of the computation of electronic wave functions by these numerical methods are given in [38], [101] and [179].

[12] Recall that energies λ_{nl} are all negative. For $Z \leq 27$ one often has $\lambda_{4s} < \lambda_{3d}$. Latter changes the Thomas–Fermi potential so that for large r one has $V(r) \to -e^2/r$.

oxygen, fluorine and neon are thus $1s^2 2s^2 2p$, $1s^2 2s^2 2p^2$, $1s^2 2s^2 2p^3$, $1s^2 2s^2 2p^4$, $1s^2 2s^2 2p^5$ and $1s^2 2s^2 2p^6$, respectively.

In this way, lowest-energy configurations of all elements of the periodic table can be established. Up to small discrepancies, they correspond pretty well to experimentally determined configurations. Before making the comparison, however, we will develop a method of systematically improving the initial potential.

9.6 Self-Consistent Field

By means of the Thomas–Fermi potential, one can find an approximate value for the wave function of electron i. The statistical interpretation of the wave function entails that the mean charge density associated with electron i is given by $q_e|\psi_i(\boldsymbol{r}_i)|^2$.

With the laws of electromagnetism one can compute the electrostatic potential due to a constant charge distribution ([40], p. 141; [123], p. 30). Specifically, the electrostatic potential produced at $\boldsymbol{r}_k$ by the charge density $q_e|\psi_i(\boldsymbol{r}_i)|^2$ is equal to

$$\frac{q_e}{4\pi\varepsilon_0}\int \mathrm{d}\boldsymbol{r}_i\, \frac{1}{r_{ik}}|\psi_i(\boldsymbol{r}_i)|^2 . \tag{9.47}$$

Hence the electrostatic potential due to the nucleus and all electrons other than electron k is given at $\boldsymbol{r}_k$ by

$$\frac{-Zq_e}{4\pi\varepsilon_0 r_k}+\frac{q_e}{4\pi\varepsilon_0}\sum_{i\neq k}\int \mathrm{d}\boldsymbol{r}_i\, \frac{1}{r_{ik}}|\psi_i(\boldsymbol{r}_i)|^2 . \tag{9.48}$$

The potential energy of electron k is equal to its charge (q_e) multiplied by the potential (9.48). The equation for the wave function of electron k can thus be written as

$$\begin{aligned} &-\frac{\hbar^2}{2m}\nabla_k^2\psi_k(\boldsymbol{r}_k)-\frac{Ze^2}{r_k}\psi_k(\boldsymbol{r}_k)\\ &\quad+e^2\left\{\sum_{i\neq k}\int \mathrm{d}\boldsymbol{r}_i\,\frac{1}{r_{ik}}|\psi_i(\boldsymbol{r}_i)|^2\right\}\psi_k(\boldsymbol{r}_k)=\lambda_k\psi_k(\boldsymbol{r}_k) . \end{aligned} \tag{9.49}$$

These equations (there is one for each value of k) are called the *Hartree equations*. They make up a system of integro-differential equations for the wave functions $\psi_k(\boldsymbol{r}_k)$.

To obtain Hartree's equations, we have assumed that each electron's potential energy is due to the mean charge density of the other electrons and of the nucleus. Furthermore, the atom's total wave function is chosen, once again, as a product of one-electron wave functions. In other words, there is no correlation between electronic positions. Thus the Hartree equations are certainly not exact. Hartree wave functions can, however, be better approximations than wave functions obtained from the Thomas–Fermi potential.

As expected, the Hartree equations do not have closed-form solutions. One can, however, obtain approximate solutions recursively, as suggested indeed by the procedure just outlined. The Thomas–Fermi potential yields a first approximation for the charge density and electronic wave functions. These wave functions are then used to compute a second (presumably better) approximation for the charge density associated with electrons other than the one considered. Hartree's equations are next solved with the electrostatic potential produced by the improved charge density. This yields a second approximation for wave functions. One can then compute a third approximation for the charge density, a third one for wave functions, and so on until convergence.

The procedure just described is called *self-consistent* because ultimately, values obtained for the charge density and electrostatic potential are stable. The setup of this procedure with Hartree's equations, however, meets with technical problems. Let us indeed examine the electrostatic potential energy of electron k, proportional to (9.48). Clearly, it depends on electron k, and it depends on polar and azimuthal angles of $\boldsymbol{r}_k$. Hartree's equations, therefore, are outside the formalism of the central-field model.

One can, however, slightly modify the Hartree equations so as to integrate them in the central-field model. Let us denote by N_{nlm} the number of electrons associated with the quantum numbers n, l and m ($N_{nlm} = 0$, 1 or 2). The total electronic charge density can then be written as

$$\sum_{i=1}^{Q} q_e |\psi_i(\boldsymbol{r})|^2 = q_e \sum_{n,l,m} N_{nlm} \left|u_{nl}(r) Y_{lm}(\theta, \phi)\right|^2 . \tag{9.50}$$

In general (9.50) depends on θ and ϕ, except in the special case where all electronic shells are *closed* (i.e. completely filled). This can be checked directly from the addition formula (7.175) for spherical harmonics. But then, inasmuch as an atom is made up of a few electrons outside closed shells, its charge density is near spherically symmetric.

We will thus replace density (9.50) by its average taken on all solid angle. Making use of the normalization relation for spherical harmonics, one gets

$$\frac{1}{4\pi} \int \mathrm{d}\Omega \, q_e \sum_{i=1}^{Q} |\psi_i(\boldsymbol{r})|^2 = \frac{q_e}{4\pi} \sum_{n,l} N_{nl} |u_{nl}(r)|^2 , \tag{9.51}$$

where $d\Omega$ stands for an element of solid angle and $N_{nl} = \sum_m N_{nlm}$ is the number of electrons in shell nl. This angle-independent electronic charge density produces an electrostatic potential equal to (Exercise 9.8)

$$\phi(r) = \frac{q_e}{4\pi\varepsilon_0}\sum_{n,l} N_{nl}\left\{\frac{1}{r}\int_0^r \mathrm{d}r'\,(r')^2|u_{nl}(r')|^2 + \int_r^\infty \mathrm{d}r'\,r'|u_{nl}(r')|^2\right\}. \tag{9.52}$$

This electrostatic potential comes from the total electronic charge density. Consequently, an electron's potential energy $V(r)$ will be given by a sum of three terms: (i) the energy of electrostatic interaction with the nucleus, (ii) $q_e\phi(r)$ and (iii) a correction term, denoted by $V_{\mathrm{corr}}(r)$, which subtracts the effect of the electron considered included in $\phi(r)$. Thus

$$\begin{aligned} V(r) = -\frac{Ze^2}{r} + e^2\sum_{n,l} N_{nl}&\left\{\frac{1}{r}\int_0^r \mathrm{d}r'\,(r')^2|u_{nl}(r')|^2\right. \\ &\left.+\int_r^\infty \mathrm{d}r'\,r'|u_{nl}(r')|^2\right\} + V_{\mathrm{corr}}(r)\,. \end{aligned} \tag{9.53}$$

This is the potential we want. Under some hypotheses, the term $V_{\mathrm{corr}}(r)$ will be computed in Chap. 17, with the following result:

$$V_{\mathrm{corr}}(r) = -e^2\left[\frac{3}{4\pi^2}\sum_{n,l} N_{nl}|u_{nl}(r)|^2\right]^{1/3}. \tag{9.54}$$

In numerical calculations, it is good to write down the radial equation (9.42) with dimensionless variables. First let

$$r = a_0\zeta, \qquad V(r) = \frac{\hbar^2}{2ma_0^2}v(\zeta), \qquad \lambda_{nl} = \frac{\hbar^2}{2ma_0^2}\varepsilon_{nl}. \tag{9.55}$$

Define next

$$\tilde{u}_{nl}(\zeta) = a_0^{3/2}u_{nl}(r)\,, \tag{9.56}$$

so that

$$\int_0^\infty \mathrm{d}\zeta\,\zeta^2|\tilde{u}_{nl}(\zeta)|^2 = 1\,. \tag{9.57}$$

Equation (9.42) thus becomes

$$\frac{1}{\zeta^2}\frac{\mathrm{d}}{\mathrm{d}\zeta}\left\{\zeta^2\frac{\mathrm{d}}{\mathrm{d}\zeta}\tilde{u}_{nl}(\zeta)\right\} - \left\{\frac{l(l+1)}{\zeta^2} + v(\zeta) - \varepsilon_{nl}\right\}\tilde{u}_{nl}(\zeta) = 0, \tag{9.58}$$

where the spherically symmetric potential $v(\zeta)$ is given by[13]

$$v(\zeta) = -\frac{2Z}{\zeta} + 2\sum_{n',l'} N_{n'l'}\left\{\frac{1}{\zeta}\int_0^\zeta \mathrm{d}\zeta'\,(\zeta')^2|\tilde{u}_{n'l'}(\zeta')|^2\right.$$

---

[13] An additional correction on $v(\zeta)$, for large values of ζ, will be made in Sect. 17.5.

$$+ \int_{\zeta}^{\infty} \mathrm{d}\zeta' \, \zeta' |\tilde{u}_{n'l'}(\zeta')|^2 \Bigg\} - \left[\frac{6}{\pi^2} \sum_{n',l'} N_{n'l'} |\tilde{u}_{n'l'}(\zeta)|^2 \right]^{1/3} . \quad (9.59)$$

Let us summarize the self-consistent calculation of atomic orbitals and electronic energies in the central-field model. An initial potential is used, for instance Thomas–Fermi's. This spherically symmetric potential represents the effect on an electron of the nucleus and the other electrons. By methods developed in Chap. 8, (9.58) is solved and radial functions $\tilde{u}_{nl}$ as well as energies ε_{nl} are found. To construct the ground state of a neutral atom, Z electrons are assigned to the lowest-energy orbitals, in accordance with the Pauli exclusion principle. An improved approximation of the potential is then obtained from (9.59). From this we get an improved approximation of radial functions $\tilde{u}_{nl}$ and energies ε_{nl}. The procedure is carried out until convergence.

Once the lowest-energy atomic configuration has been found, the wave function of an excited state can be obtained by assigning electrons to corresponding orbitals. The self-consistent calculation can then proceed. In fact when a single electron is excited, one can use orbitals obtained in the calculation of the lowest-energy configuration and simply move an electron to a higher orbital.

The central-field model can also be used to determine ionic configurations. For this purpose, the neutral-atom orbitals are first obtained. The ion's Q electrons are next assigned to appropriate orbitals, and the charge density associated with these Q electrons is computed. A potential is obtained from that charge density, and the self-consistent calculation can proceed.

9.7 Validity of the Central-Field Model

Results for eigenvalues λ_{nl}, computed by the method developed in this chapter, are presented in Table 9.1 (pp. 200 and 201), for the ground states of neutral atoms up to krypton. Also displayed are the experimental electronic configuration and ionization energies.[14]

The central-field model gives a good idea of ground-state electronic configurations. Indeed experimental configurations are correctly predicted (with the exception of $3d$ and $4s$ orbitals) if electrons are systematically assigned to lowest-energy orbitals. An atom's ionization energy is of the order of the

---

[14] Eigenvalues λ_{nl} as well as corresponding radial functions are computed in [156]. Ionization energies are taken from [154]. [Ar] stands for the ground-state configuration of argon. All values are expressed in multiples of $e^2/2a_0 = me^4/2\hbar^2$, where m is the electron's reduced mass in each atom. The λ_{nl} associated with hydrogen come from the exact solution. Values in parentheses are associated with unoccupied orbitals.

absolute value of the energy associated with the least strongly bound electron.[15]

A look at Table 9.1 reveals that the more electrons there are in the outer shell, the higher the ionization energy. Outer-shell electrons have much smaller energies (in absolute value) than inner-shell electrons. As the chemical properties of elements are determined by outer-shell electrons (possibly the last two shells, if they have close energies), elements that have similar outer-shell electronic configurations have similar chemical properties.

In addition to giving approximation of ionization energies, the central-field model yields an estimate of electronic excitation energies, at least in some simple cases. Consider an electronic configuration obtained from the ground state by moving a single electron from one to another orbital. The model then predicts that the excitation energy is equal to the difference of orbital energies. As a matter of fact, values so obtained can differ from experimental values rather significantly. The difference is even more important if configurations with more than one excited electron are introduced.[16]

Whether quantitative values are accurate or not, the qualitative structure of atomic energy levels is very simple in the central-field model. To each electronic configuration there corresponds one and only one energy. However with a few exceptions (i.e. closed-shell configurations or configurations that differ from closed shells by one electron only), this is in total contradiction with experiment. The structure of atomic energy levels is in general much more complex. Different quantum states belonging to the same configuration usually have different energies, even though differences are normally much smaller than the energy gap between distinct configurations.

To account for the complexity of atomic spectra, it is essential to add two ingredients to the central-field model. On the one hand, wave functions must be used that reflect the identity of all electrons. In accordance with the fact that the atom's total angular momentum commutes with the Hamiltonian, on the other hand, wave functions must be picked from eigenfunctions of angular momentum. In Chap. 18 we will carry out such a detailed analysis.

---

[15] Spherically symmetric potentials that yield better approximations for ionization energies can be found by using different forms of the term V_{corr}. Other atomic parameters may then be represented less accurately. At any rate, the correct way to compute ionization energies will be developed in Chap. 18.

[16] The energy necessary to excite two or more electrons is not equal to the sum of energies required to excite them separately. In fact, an atom's total energy is not equal to the sum of values λ_{nl} corresponding to each electron. Indeed such a sum would doubly evaluate the contribution from electronic potential energy. No matter what configuration is investigated, the total energy of an atom must be calculated by evaluating the mean value of the Hamiltonian in the state involved. Such a computation correctly yields the ground-state configuration. We will come back to this in Chaps. 17 and 18.

Table 9.1. Experimental configurations, electronic energies (Ry) and ionisation energies (Ry) of light atoms

	Configuration	$-\lambda_{1s}$	$-\lambda_{2s}$	$-\lambda_{2p}$	$-\lambda_{3s}$
H	$1s$	1.000	(0.250)	(0.250)	
He	$1s^2$	2.330	(0.340)	(0.254)	
Li	$1s^2 2s$	5.624	0.475	(0.302)	(0.163)
Be	$1s^2 2s^2$	10.264	0.851	(0.523)	(0.207)
B	$1s^2 2s^2 2p$	16.291	1.281	0.781	(0.240)
C	$1s^2 2s^2 2p^2$	23.653	1.749	1.052	(0.268)
N	$1s^2 2s^2 2p^3$	32.363	2.260	1.336	(0.293)
O	$1s^2 2s^2 2p^4$	42.429	2.813	1.633	(0.315)
F	$1s^2 2s^2 2p^5$	53.854	3.410	1.943	(0.335)
Ne	$1s^2 2s^2 2p^6$	66.642	4.051	2.266	(0.353)
Na	$[\mathrm{Ne}]3s$	82.350	5.840	3.724	0.460
Mg	$[\mathrm{Ne}]3s^2$	99.715	7.792	5.344	0.722
Al	$[\mathrm{Ne}]3s^2 3p$	118.907	10.096	7.309	1.039
Si	$[\mathrm{Ne}]3s^2 3p^2$	139.780	12.609	9.478	1.361
P	$[\mathrm{Ne}]3s^2 3p^3$	162.344	15.345	11.865	1.691
S	$[\mathrm{Ne}]3s^2 3p^4$	186.603	18.309	14.474	2.031
Cl	$[\mathrm{Ne}]3s^2 3p^5$	212.561	21.505	17.309	2.383
Ar	$[\mathrm{Ne}]3s^2 3p^6$	240.217	24.934	20.370	2.747
K	$[\mathrm{Ar}]4s$	270.389	29.373	24.439	3.699
Ca	$[\mathrm{Ar}]4s^2$	302.357	34.119	28.812	4.681
Sc	$[\mathrm{Ar}]4s^2 3d$	335.553	38.529	32.850	5.257
Ti	$[\mathrm{Ar}]4s^2 3d^2$	370.460	43.159	37.102	5.839
V	$[\mathrm{Ar}]4s^2 3d^3$	407.079	48.013	41.573	6.432
Cr	$[\mathrm{Ar}]4s 3d^5$	444.820	52.435	45.611	6.517
Mn	$[\mathrm{Ar}]4s^2 3d^5$	485.468	58.407	51.187	7.656
Fe	$[\mathrm{Ar}]4s^2 3d^6$	527.241	63.951	56.333	8.289
Co	$[\mathrm{Ar}]4s^2 3d^7$	570.734	69.729	61.708	8.937
Ni	$[\mathrm{Ar}]4s^2 3d^8$	615.948	75.741	67.313	9.601
Cu	$[\mathrm{Ar}]4s 3d^{10}$	662.171	81.212	72.376	9.641
Zn	$[\mathrm{Ar}]4s^2 3d^{10}$	711.544	88.472	79.215	10.978
Ga	$[\mathrm{Ar}]4s^2 3d^{10} 4p$	762.899	96.233	86.551	12.555
Ge	$[\mathrm{Ar}]4s^2 3d^{10} 4p^2$	816.073	104.335	94.223	14.200
As	$[\mathrm{Ar}]4s^2 3d^{10} 4p^3$	871.074	112.786	102.240	15.923
Se	$[\mathrm{Ar}]4s^2 3d^{10} 4p^4$	927.905	121.585	110.602	17.726
Br	$[\mathrm{Ar}]4s^2 3d^{10} 4p^5$	986.565	130.733	119.309	19.610
Kr	$[\mathrm{Ar}]4s^2 3d^{10} 4p^6$	1 047.056	140.231	128.361	21.575

Table 9.1. Experimental configurations, electronic energies (Ry) and ionisation energies (Ry) of light atoms (continued)

$-\lambda_{3p}$	$-\lambda_{3d}$	$-\lambda_{4s}$	$-\lambda_{4p}$	E_I	
				1.000	H
				1.807	He
				0.396	Li
				0.685	Be
(0.176)				0.610	B
(0.189)				0.828	C
(0.199)				1.068	N
(0.207)				1.001	O
(0.213)				1.281	F
(0.218)				1.585	Ne
(0.253)	(0.063)	(0.159)		0.378	Na
(0.379)	(0.117)	(0.187)		0.562	Mg
0.568	(0.123)	(0.211)	(0.150)	0.440	Al
0.766	(0.129)	(0.231)	(0.163)	0.599	Si
0.971	(0.133)	(0.250)	(0.172)	0.771	P
1.184	(0.135)	(0.266)	(0.181)	0.761	S
1.404	(0.136)	(0.281)	(0.188)	0.953	Cl
1.632	(0.135)	(0.295)	(0.194)	1.158	Ar
2.414	(0.470)	0.391	(0.229)	0.319	K
3.228	(0.951)	0.578	(0.333)	0.449	Ca
3.653	1.092	0.616	(0.346)	0.482	Sc
4.080	1.226	0.650	(0.356)	0.502	Ti
4.514	1.354	0.682	(0.365)	0.496	V
4.447	1.016	0.552	(0.279)	0.497	Cr
5.403	1.598	0.740	(0.378)	0.546	Mn
5.862	1.716	0.767	(0.383)	0.581	Fe
6.330	1.830	0.793	(0.387)	0.579	Co
6.808	1.942	0.818	(0.390)	0.562	Ni
6.668	1.464	0.659	(0.296)	0.568	Cu
7.798	2.159	0.867	(0.394)	0.690	Zn
9.162	3.079	1.152	0.567	0.441	Ga
10.590	4.062	1.429	0.744	0.581	Ge
12.093	5.119	1.704	0.921	0.719	As
13.673	6.250	1.979	1.098	0.717	Se
15.331	7.457	2.256	1.277	0.868	Br
17.069	8.742	2.537	1.459	1.029	Kr

Exercises

9.1. Carry out transformations (9.7) and (9.8) in (9.6), thereby obtaining (9.9).

9.2. A hydrogen atom goes from a level with principal quantum number n_i to a level $n_f < n_i$. The radiation wavelength is equal to $ch/(E_i - E_f)$. Compute relative wavelength differences between normal hydrogen, deuterium and tritium. [Assume equal proton and neutron masses, 1840 times bigger than the electron mass.]

9.3. How many free electron states with energy lower than 10 eV are there in a cubic box of length 3 Å?

9.4. Consider a free particle in three dimensions.
a) Find all solutions of the Hamiltonian's eigenvalue equation that vanish on the faces of a cube of length $2L$, i.e. such that

$$\begin{aligned}\psi(0,y,z) &= 0 = \psi(2L,y,z)\,,\\ \psi(x,0,z) &= 0 = \psi(x,2L,z)\,,\\ \psi(x,y,0) &= 0 = \psi(x,y,2L)\,,\end{aligned}$$

for all x, y and z.
b) Show that the total number of states with energy lower than E is, as in the case of periodic boundary conditions, equal to

$$\frac{4L^3}{3\pi^2\hbar^3}(2mE)^{3/2}\,.$$

9.5. Perform the change of variables (9.36) and obtain (9.37) and (9.41).

9.6.
a) Show that the function $\Phi = A\xi^{-n}$ is an exact solution of the Thomas–Fermi equation for a specific value of A and a specific value of n, and find these values. Note, however, that the condition at $\xi = 0$ is not satisfied.
b) On the other hand, show that the function

$$\Phi = 1 + \alpha\xi + \frac{4}{3}\xi^{3/2}$$

is an approximate solution of the Thomas–Fermi equation for $\xi \ll 1$ and for any value of α. The condition at $\xi = 0$ is satisfied but as ξ becomes large, the approximation deteriorates.

9.7.
a) Write the electronic density in the Thomas–Fermi model in terms of the function Φ and variable ξ.
b) Show that the number of electrons inside a sphere of radius R is given by

$$Z\left[\xi\frac{\mathrm{d}\Phi}{\mathrm{d}\xi}-\Phi\right]_0^{R/\beta}, \quad \text{where} \quad \beta=\frac{a_0}{4}\left(\frac{9\pi^2}{2Z}\right)^{1/3}.$$

c) Deduce from this that the radius of a sphere that contains a fixed fraction of electrons is proportional to $Z^{-1/3}$, and that a sphere of infinite radius indeed contains Z electrons. [To evaluate the quantity in square brackets, use solutions found in Exercise (9.6), in the region where they are valid.]

9.8. Let $\varrho(r)$ be an electronic charge density that is bounded and restricted to a finite region of space. Let $\phi(r)$ be defined as

$$\phi(r)=\frac{1}{\varepsilon_0 r}\int_0^r \mathrm{d}r'\,(r')^2\varrho(r')+\frac{1}{\varepsilon_0}\int_r^\infty \mathrm{d}r'\,r'\varrho(r').$$

Show that $\phi(r)\to 0$ as $r\to\infty$ and that $\phi(r)$ satisfies Poisson's equation, that is,

$$\nabla^2\phi(r)=-\frac{\varrho(r)}{\varepsilon_0}.$$

9.9. Let a total charge q be uniformly distributed inside a sphere of radius r_0, that is,

$$\varrho(r)=\begin{cases}\frac{3q}{4\pi r_0^3} & \text{if } r<r_0\,,\\ 0 & \text{if } r>r_0\,.\end{cases}$$

Use the result of Exercise (9.8) to show that the electrostatic potential $\phi(r)$ produced by this charge density is given by

$$\phi(r)=\begin{cases}\frac{q}{8\pi\varepsilon_0 r_0}\left[3-\left(\frac{r}{r_0}\right)^2\right] & \text{if } r<r_0\,,\\ \frac{q}{4\pi\varepsilon_0 r} & \text{if } r>r_0\,.\end{cases}$$

10 Stationary Perturbations

Stationary perturbation theory is one of the main approximation methods in quantum mechanics. It applies to quantum systems with a Hamiltonian which, in a sense that will be made clear, is close to a Hamiltonian whose eigenvalues and eigenvectors are known. We will develop the formalism and use it to investigate the effect of the spatial extension of the nucleus on electronic energies; the Stark effect and the van der Waals forces in hydrogen; and the rotation-vibration energies of a diatomic molecule.[1]

10.1 Small Perturbation

Consider a quantum system governed by a Hamiltonian H_0 that does not depend on time. Let eigenvalues of H_0 be denoted by E_i, where the index i covers both the discrete and the continuous spectrum. In the former, eigenvalues can be ordered so that

$$E_1 < E_2 < E_3 < \cdots$$

In the latter, i takes a suitable form which we won't need to specify further.

Let eigenvectors of H_0 be denoted by $|E_i^\alpha\rangle$. The purpose of exponent α is to distinguish, in the case where eigenvalue E_i is degenerate, different eigenvectors corresponding to that eigenvalue. The eigenvalue equation of H_0 can be written as

$$H_0|E_i^\alpha\rangle = E_i|E_i^\alpha\rangle\,. \tag{10.1}$$

We assume that kets $|E_i^\alpha\rangle$ are orthonormal. In the discrete spectrum this means that

$$\langle E_i^\alpha|E_j^\beta\rangle = \delta_{ij}\delta_{\alpha\beta}\,. \tag{10.2}$$

In the continuous spectrum Kronecker deltas should be replaced by appropriate Dirac deltas.

---

[1] Stationary perturbation theory is developed in most books on quantum mechanics. Our presentation more particularly follows [53], Chap. 11 and [200], Sects. 31–33.

Let us now examine a quantum system slightly different from the one we just considered. By this we mean a quantum system whose state space is the same as the one of the original system, but whose Hamiltonian is given by

$$H = H_0 + W \,. \tag{10.3}$$

Here W is a Hermitian operator called the *perturbation.* In this chapter we will always assume that matrix elements of W in the $|E_i^\alpha\rangle$ basis are small compared with differences $E_{i+1} - E_i$. In this sense the perturbation is small. We will call H_0 and H the *unperturbed Hamiltonian* and the *perturbed Hamiltonian*, respectively.

The formalism we will develop yields, in an approximate way, the eigenvalues and eigenvectors of the perturbed Hamiltonian from the eigenvalues and eigenvectors of the unperturbed Hamiltonian. To achieve this it is very useful to introduce a real parameter λ and an operator $H(\lambda)$ defined as

$$H(\lambda) = H_0 + \lambda W \,. \tag{10.4}$$

Clearly $H = H(1)$. The eigenvalue equation of the operator $H(\lambda)$ can be written as

$$H(\lambda)|\chi(\lambda)\rangle = E(\lambda)|\chi(\lambda)\rangle \,. \tag{10.5}$$

Inasmuch as the perturbation is weak, one can think that eigenvalues $E(\lambda)$ are close to the E_i and that eigenvectors $|\chi(\lambda)\rangle$ are close to the vectors $|E_i^\alpha\rangle$. We will assume that for $0 \le \lambda \le 1$ one can expand $E(\lambda)$ and $|\chi(\lambda)\rangle$ in series of powers of λ, that is,

$$E(\lambda) = \varepsilon_0 + \lambda\varepsilon_1 + \lambda^2\varepsilon_2 + \cdots \tag{10.6}$$

$$|\chi(\lambda)\rangle = |\chi_0\rangle + \lambda|\chi_1\rangle + \lambda^2|\chi_2\rangle + \cdots \tag{10.7}$$

Here ε_0 is one of the E_i, say E_n. The ket $|\chi_0\rangle$ is a linear combination of the $|E_n^\alpha\rangle$.

The existence of the series expansions (10.6) and (10.7) is by no means obvious, especially when the state space has infinite dimension [131]. The formalism we will develop can nonetheless be useful even if the series in powers of λ do not converge, provided that the first few terms in the expansions make up a good approximation of $E(\lambda)$ and $|\chi(\lambda)\rangle$.

We have pointed out that the unperturbed Hamiltonian's spectrum can be in part discrete and in part continuous. We will assume that the eigenvalue E_n, whose change due to the perturbation we are looking for, belongs to the discrete spectrum. One can then set $\langle\chi(\lambda)|\chi(\lambda)\rangle = 1$ for every λ between 0 and 1.

Let us now substitute expansions (10.6) and (10.7) in (10.5). One gets

$$(H_0 + \lambda W)\left\{\sum_{l=0}^{\infty}\lambda^l|\chi_l\rangle\right\} = \left\{\sum_{l=0}^{\infty}\varepsilon_l\lambda^l\right\}\left\{\sum_{l=0}^{\infty}\lambda^l|\chi_l\rangle\right\}. \tag{10.8}$$

This equation holds for any λ between 0 et 1. Hence coefficients of the same powers of λ, on the left and right-hand sides, must be equal. Thus

$$
\begin{aligned}
H_0|\chi_0\rangle &= \varepsilon_0|\chi_0\rangle\,, & (10.9)\\
H_0|\chi_1\rangle + W|\chi_0\rangle &= \varepsilon_0|\chi_1\rangle + \varepsilon_1|\chi_0\rangle\,, & (10.10)\\
H_0|\chi_2\rangle + W|\chi_1\rangle &= \varepsilon_0|\chi_2\rangle + \varepsilon_1|\chi_1\rangle + \varepsilon_2|\chi_0\rangle\,, & (10.11)\\
\cdots &= \cdots
\end{aligned}
$$

Equation (10.9) merely represents the eigenvalue equation of the unperturbed Hamiltonian. As we pointed out, ε_0 is equal to E_n and $|\chi_0\rangle$ is a linear combination of the kets $|E_n^\alpha\rangle$. In the next two sections we will investigate first the case where E_n is not degenerate, and then the case where it is.

But before, notice that the normalization condition of $|\chi(\lambda)\rangle$ entails that

$$
\begin{aligned}
1 &= \langle\chi(\lambda)|\chi(\lambda)\rangle\\
&= \langle\chi_0|\chi_0\rangle + \lambda\left\{\langle\chi_0|\chi_1\rangle + \langle\chi_1|\chi_0\rangle\right\}\\
&\quad + \lambda^2\left\{\langle\chi_0|\chi_2\rangle + \langle\chi_1|\chi_1\rangle + \langle\chi_2|\chi_0\rangle\right\} + O(\lambda^3)\,.
\end{aligned}
$$

This equation holds for every λ. Since $\langle\chi_0|\chi_0\rangle = 1$, one must have

$$
\begin{aligned}
\langle\chi_0|\chi_1\rangle + \langle\chi_1|\chi_0\rangle &= 2\,\mathrm{Re}\left\{\langle\chi_0|\chi_1\rangle\right\} = 0\,, & (10.12)\\
\langle\chi_0|\chi_2\rangle + \langle\chi_1|\chi_1\rangle + \langle\chi_2|\chi_0\rangle &= 2\,\mathrm{Re}\left\{\langle\chi_0|\chi_2\rangle\right\} + \langle\chi_1|\chi_1\rangle = 0\,. & (10.13)
\end{aligned}
$$

The amplitude of $|\chi(\lambda)\rangle$ has already been determined by the normalization condition. Without loss of generality, the phase of $|\chi(\lambda)\rangle$ can be picked so that $\langle\chi_0|\chi(\lambda)\rangle$ is real for all λ. Making use of the series expansion of $|\chi(\lambda)\rangle$, one immediately concludes that $\langle\chi_0|\chi_1\rangle$, $\langle\chi_0|\chi_2\rangle$, etc., must be real. Equations (10.12) and (10.13) then imply that

$$
\begin{aligned}
\langle\chi_0|\chi_1\rangle &= 0\,, & (10.14)\\
\langle\chi_0|\chi_2\rangle &= -\frac{1}{2}\langle\chi_1|\chi_1\rangle\,. & (10.15)
\end{aligned}
$$

10.2 Nondegenerate Energy

We first investigate the case where the eigenvalue E_n of the unperturbed Hamiltonian is not degenerate. Then

$$
\varepsilon_0 = E_n\,, \qquad |\chi_0\rangle = |E_n\rangle\,. \qquad (10.16)
$$

These equations represent the zeroth-order approximation of the eigenvalue $E(\lambda)$ and eigenvector $|\chi(\lambda)\rangle$ of the Hamiltonian $H(\lambda)$.

Let us now compute the first-order correction, i.e. ε_1 and $|\chi_1\rangle$. To find ε_1 we project (10.10) on the bra $\langle\chi_0|$. Thus

$$\langle\chi_0|H_0|\chi_1\rangle+\langle\chi_0|W|\chi_0\rangle = \langle\chi_0|\varepsilon_0|\chi_1\rangle+\langle\chi_0|\varepsilon_1|\chi_0\rangle .$$

Making use of the fact that $|\chi_0\rangle$ is a normalized eigenvector of H_0, one gets

$$\varepsilon_1 = \langle\chi_0|W|\chi_0\rangle . \tag{10.17}$$

To find $|\chi_1\rangle$ we project (10.10) on $\langle E_i^\alpha|$, for $i \neq n$. One easily gets that

$$(E_i - \varepsilon_0)\,\langle E_i^\alpha|\chi_1\rangle = -\langle E_i^\alpha|W|\chi_0\rangle . \tag{10.18}$$

Since $\langle E_n|\chi_1\rangle = 0$ one can write

$$|\chi_1\rangle = |E_n\rangle\langle E_n|\chi_1\rangle+\mathsf{S}_{i\neq n,\alpha}|E_i^\alpha\rangle\langle E_i^\alpha|\chi_1\rangle = \mathsf{S}_{i\neq n,\alpha}\frac{\langle E_i^\alpha|W|\chi_0\rangle}{E_n - E_i}|E_i^\alpha\rangle . \tag{10.19}$$

We have obtained ε_1 and $|\chi_1\rangle$, the first-order corrections of the Hamiltonian's eigenvalue and eigenvector. Therefore

$$E(\lambda = 1) = E_n + \langle E_n|W|E_n\rangle + \cdots \tag{10.20}$$

$$|\chi(\lambda = 1)\rangle = |E_n\rangle + \mathsf{S}_{i\neq n,\alpha}\frac{\langle E_i^\alpha|W|E_n\rangle}{E_n - E_i}|E_i^\alpha\rangle + \cdots \tag{10.21}$$

It is clear that for first-order corrections to be small, matrix elements $\langle E_i^\alpha|W|E_n\rangle$ should be small compared with differences $E_n - E_i$.

Higher-order corrections are obtained like ε_1 and $|\chi_1\rangle$, except that expressions become more and more complicated. To find ε_2 we project (10.11) on $\langle\chi_0|$. We get

$$\langle\chi_0|W|\chi_1\rangle = \varepsilon_2 , \tag{10.22}$$

which entails that

$$\varepsilon_2 = \mathsf{S}_{i\neq n,\alpha}\frac{\langle E_i^\alpha|W|\chi_0\rangle}{E_n - E_i}\langle\chi_0|W|E_i^\alpha\rangle = \mathsf{S}_{i\neq n,\alpha}\frac{|\langle E_n|W|E_i^\alpha\rangle|^2}{E_n - E_i} . \tag{10.23}$$

Furthermore, we find in Exercise (10.2) that

$$|\chi_2\rangle = \mathsf{S}_{i\neq n,\alpha}\left\{\mathsf{S}_{j\neq n,\beta}\frac{\langle E_i^\alpha|W|E_j^\beta\rangle\langle E_j^\beta|W|E_n\rangle}{(E_n - E_j)(E_n - E_i)} - \frac{\langle E_i^\alpha|W|E_n\rangle\langle E_n|W|E_n\rangle}{(E_n - E_i)^2}\right\}|E_i^\alpha\rangle - \frac{1}{2}\mathsf{S}_{i\neq n,\alpha}\frac{|\langle E_n|W|E_i^\alpha\rangle|^2}{(E_n - E_i)^2}|E_n\rangle . \tag{10.24}$$

The constant ε_2 and the vector $|\chi_2\rangle$ are the second-order corrections of the energy E_n and the state vector $|E_n\rangle$, produced by the perturbation.

How good an approximation the expression $\varepsilon_0 + \varepsilon_1$ is of the exact eigenvalue $E(\lambda = 1)$ of the perturbed Hamiltonian? To answer this question we will find an upper bound on $|\varepsilon_2|$. Letting $\Delta E = \min|(E_n - E_i)|_{n \neq i}$, one certainly has

$$\begin{aligned}
|\varepsilon_2| &\leq \mathsf{S}_{i \neq n, \alpha} \frac{1}{\Delta E} \left|\langle E_n|W|E_i^\alpha\rangle\right|^2 \\
&\leq \frac{1}{\Delta E} \mathsf{S}_{i \neq n, \alpha} \langle E_n|W|E_i^\alpha\rangle\langle E_i^\alpha|W|E_n\rangle \\
&\leq \frac{1}{\Delta E} \langle E_n|W| \left\{I - |E_n\rangle\langle E_n|\right\} |W|E_n\rangle \\
&\leq \frac{1}{\Delta E} \left\{\langle E_n|W^2|E_n\rangle - \left(\langle E_n|W|E_n\rangle\right)^2\right\} .
\end{aligned} \tag{10.25}$$

The expression in curly brackets coincides with the variance of W in state $|E_n\rangle$. Hence if

$$\frac{1}{\Delta E} \left[\mathrm{Var}(W)\right]_{|E_n\rangle} \ll |\varepsilon_1| , \tag{10.26}$$

one can think that $\varepsilon_0 + \varepsilon_1$ is a good approximation of $E(\lambda = 1)$.

10.3 Degenerate Energy

Let us now turn to the case where an eigenvalue E_n of H_0 is g_n-fold degenerate ($g_n > 1$). There are g_n orthonormal eigenvectors of H_0 that correspond to the eigenvalue E_n. They are denoted by $|E_n^\alpha\rangle$ ($\alpha = 1, 2, \ldots, g_n$) and they make up a subspace $\mathcal{V}_n$ of the state space.

The vector $|\chi_0\rangle$ here is not uniquely determined by (10.9). At any rate, assume that $|\chi_0\rangle$ is one of the vectors obeying that equation. Let us project (10.10) on $\langle E_n^\alpha|$. One finds that

$$\langle E_n^\alpha|W|\chi_0\rangle = \varepsilon_1 \langle E_n^\alpha|\chi_0\rangle . \tag{10.27}$$

Introducing the identity operator appropriately and making use of the fact that $|\chi_0\rangle$ belongs to $\mathcal{V}_n$, one easily obtains that

$$\sum_{\beta=1}^{g_n} \langle E_n^\alpha|W|E_n^\beta\rangle\langle E_n^\beta|\chi_0\rangle = \varepsilon_1 \langle E_n^\alpha|\chi_0\rangle . \tag{10.28}$$

Denote $\langle E_n^\alpha|W|E_n^\beta\rangle$ by $W_{\alpha\beta}$ and $\langle E_n^\alpha|\chi_0\rangle$ by c_α. Equation (10.28) becomes

$$\sum_{\beta=1}^{g_n} W_{\alpha\beta} c_\beta = \varepsilon_1 c_\alpha . \tag{10.29}$$

This is just the eigenvalue equation of matrix $W_{\alpha\beta}$, which has dimension $g_n \times g_n$. We have found that first-order corrections to energy due to the perturbation coincide with eigenvalues of the matrix $W_{\alpha\beta}$.

The complete solution of the eigenvalue equation of $W_{\alpha\beta}$ yields g_n orthonormal vectors $|\chi_0^\alpha\rangle$, which are linear combinations of the $|E_n^\alpha\rangle$ and for which

$$\langle\chi_0^\alpha|W|\chi_0^\beta\rangle = \varepsilon_{1,\alpha}\delta_{\alpha\beta}\,. \tag{10.30}$$

Constants $\varepsilon_{1,\alpha}$ are eigenvalues of $W_{\alpha\beta}$. Instead of the $|E_n^\alpha\rangle$, it is much more convenient to use the kets $|\chi_0^\alpha\rangle$ as eigenvectors of the unperturbed Hamiltonian.

Whenever all $\varepsilon_{1,\alpha}$ are distinct, one says the degeneracy is *removed* in first order. Otherwise the degeneracy remains partly or completely. It may then be removed in a higher order, but such is not always the case.

The first-order correction to $|\chi_0^\alpha\rangle$ can be denoted by $|\chi_1^\alpha\rangle$. To evaluate it we project (10.10) on $\langle E_i^\beta|$ $(i \neq n)$. One easily finds that

$$\langle E_i^\beta|\chi_1^\alpha\rangle = \frac{\langle E_i^\beta|W|\chi_0^\alpha\rangle}{E_n - E_i}\,. \tag{10.31}$$

Equation (10.14) implies that $\langle\chi_0^\alpha|\chi_1^\alpha\rangle = 0$. Hence we get

$$|\chi_1^\alpha\rangle = \mathrm{S}_{i\neq n,\beta}\frac{\langle E_i^\beta|W|\chi_0^\alpha\rangle}{E_n - E_i}|E_i^\beta\rangle + \sum_{\beta\neq\alpha} d_\beta^\alpha|\chi_0^\beta\rangle\,. \tag{10.32}$$

Equation (10.10) is consistent with any choice of constants d_β^α. To determine these it is necessary to examine the second-order corrections to energy.

To do this project (10.11) on $\langle\chi_0^\gamma|$. Substituting (10.32) in the result and manipulating, one gets

$$\mathcal{W}_{\gamma\alpha} + \sum_{\beta\neq\alpha} d_\beta^\alpha(\varepsilon_{1,\gamma} - \varepsilon_{1,\alpha})\delta_{\gamma\beta} = \varepsilon_{2,\alpha}\delta_{\alpha\gamma}\,, \tag{10.33}$$

where matrix elements $\mathcal{W}_{\gamma\alpha}$ are given by

$$\mathcal{W}_{\gamma\alpha} = \mathrm{S}_{i\neq n,\beta}\frac{\langle\chi_0^\gamma|W|E_i^\beta\rangle\langle E_i^\beta|W|\chi_0^\alpha\rangle}{E_n - E_i}\,. \tag{10.34}$$

We will consider the case where all $\varepsilon_{1,\alpha}$ are distinct, and the one where they are all equal. Intermediate cases can be analyzed in a similar way.

If all $\varepsilon_{1,\alpha}$ are distinct, (10.33) can be written as

$$\mathcal{W}_{\alpha\alpha} = \varepsilon_{2,\alpha}\,, \tag{10.35}$$

$$\mathcal{W}_{\gamma\alpha} = d_\gamma^\alpha(\varepsilon_{1,\alpha} - \varepsilon_{1,\gamma})\,, \qquad \gamma \neq \alpha\,. \tag{10.36}$$

Second-order corrections to energy are then obtained, as well as all coefficients of first-order corrections to the state vector.

If all $\varepsilon_{1,\alpha}$ are equal, (10.33) simply becomes

$$\mathcal{W}_{\gamma\alpha} = \varepsilon_{2,\alpha}\delta_{\alpha\gamma} \, . \tag{10.37}$$

One can see that this is an eigenvalue equation that determines corrections $\varepsilon_{2,\alpha}$ and (inasmuch as degeneracy is removed) vectors $|\chi_0^\alpha\rangle$. Constants d_β^α, however, are at this stage still not determined.

As a matter of fact, stationary perturbation theory is rarely pushed beyond second order. It is useful precisely in those situations where the first few terms in expansions (10.6) and (10.7) adequately represent the perturbed Hamiltonian's eigenvalues and eigenvectors.

In certain circumstances it is possible to extend to distinct energy levels the stationary perturbation formalism applicable to degenerate energy. This happens with close levels. To be more specific, consider two energy levels E_1 and E_2 associated with subspaces $\mathcal{V}_1$ and $\mathcal{V}_2$ of the state space. Assume that $E_2 - E_1$ is small compared with differences $E_i - E_1$ and $E_i - E_2$ $(1 \neq i \neq 2)$, but comparable with matrix elements $\langle E_1^\alpha|W|E_1^\beta\rangle$ and $\langle E_2^\alpha|W|E_2^\beta\rangle$. It is then convenient to redefine the perturbation so as to include in it the energy difference between the two levels, and to diagonalize it in the space $\mathcal{V}_1 \oplus \mathcal{V}_2$ rather than in $\mathcal{V}_1$ and $\mathcal{V}_2$ separately. Examples of this situation occur in atomic physics, in particular, where some perturbations are diagonalized in the set of all states associated with a given electronic configuration.

10.4 Spatial Extension of the Atomic Nucleus

In the chapter devoted to the central-field model, the atomic nucleus was always taken as a point charge. Such is not exactly the case in the real world, where the nucleus has spatial extension. Although nuclear dimensions ($\approx 10^{-15}$ m) are small compared with atomic dimensions ($\approx 10^{-10}$ m), the spatial extension of the nucleus can have a nonnegligible effect on the energy of some orbitals, particularly with big or exotic atoms.

We will assume the nucleus is a uniformly charged sphere of radius r_0. Its charge density $\varrho(r)$ is thus given by

$$\varrho(r) = \begin{cases} -\frac{3Zq_e}{4\pi r_0^3} & \text{if } r < r_0 \, , \\ 0 & \text{if } r > r_0 \, . \end{cases} \tag{10.38}$$

Following the laws of electrostatics, the potential energy of an electron in the field produced by this charge density is equal to

$$\begin{cases} -\frac{Ze^2}{2r_0}\left\{3 - \left(\frac{r}{r_0}\right)^2\right\} & \text{if } r < r_0 \, , \\ -\frac{Ze^2}{r} & \text{if } r > r_0 \, , \end{cases}$$

where (9.4) and the result of Exercise (9.9) were used. This expression differs from $-Ze^2/r$ if $r < r_0$. Let H_0 be the Hamiltonian of an electron in the central-field model. The effect of the spatial extension of the nucleus can be represented by taking $H = H_0 + W$ instead of H_0, where W is a perturbation given in the coordinate representation by

$$W(r) = \begin{cases} \frac{Ze^2}{2r_0}\left\{\left(\frac{r}{r_0}\right)^2 - 3 + \frac{2r_0}{r}\right\} & \text{if } r < r_0\,, \\ 0 & \text{if } r > r_0\,. \end{cases} \tag{10.39}$$

We will compute the first-order correction to electronic energies produced by this perturbation. For an electron in an orbital nl, the energy (with spin left out) is $(2l+1)$-fold degenerate, different states corresponding to different values of m. Hence one has to evaluate matrix elements $\langle nlm|W|nlm'\rangle$, given by

$$\begin{aligned}\langle nlm|W|nlm'\rangle &= \int \mathrm{d}\boldsymbol{r}\, \psi^*_{nlm}(\boldsymbol{r})W(r)\psi_{nlm'}(\boldsymbol{r}) \\ &= \int_0^\infty \mathrm{d}r\, r^2 u^*_{nl}(r)W(r)u_{nl}(r) \int \mathrm{d}\Omega\, Y^*_{lm}(\theta,\phi)Y_{lm'}(\theta,\phi) \\ &= \delta_{mm'} \int_0^\infty \mathrm{d}r\, r^2 |u_{nl}(r)|^2 W(r)\,.\end{aligned}$$

The result is very simple. For a given orbital, the perturbation is diagonal in the $\psi_{nlm}(\boldsymbol{r})$ basis, with all diagonal elements equal. The correction to the energy of an orbital nl in first-order perturbation theory is therefore equal to

$$\varepsilon_1 = \int_0^\infty \mathrm{d}r\, r^2 |u_{nl}(r)|^2 W(r)\,. \tag{10.40}$$

Note that since $W(r) \geq 0$ for all r, the correction ε_1 is never negative.

Equation (10.40) is also correct for the hydrogen atom and hydrogen-like ions, even though there is an additional degeneracy in l. Indeed matrix elements $\langle nlm|W|nl'm'\rangle$ vanish unless $l = l'$ and $m = m'$, so that again the perturbation is diagonal.

In most situations, the typical size of orbital nl is much larger than r_0. One can then replace the function $u_{nl}(r)$ in the integral (10.40) by its value at $r = 0$. One thus obtains

$$\varepsilon_1 = |u_{nl}(0)|^2 \int_0^\infty \mathrm{d}r\, r^2 W(r) = |u_{nl}(0)|^2 \frac{Ze^2}{10} r_0^2\,. \tag{10.41}$$

We know that if $l \neq 0$, then $u_{nl}(0) = 0$. Hence to first order, only s orbital energies are significantly changed. For the 1s orbital of a hydrogen-like ion, one has

$$u_{10}(0) = 2\left(\frac{Z}{a_0}\right)^{3/2}\,,$$

so that

$$\varepsilon_1 = \frac{2}{5}\frac{Z^4 e^2 r_0^2}{a_0^3} = \frac{4Z^2}{5}\frac{r_0^2}{a_0^2}|E| . \tag{10.42}$$

Here

$$|E| = \frac{Z^2 e^2}{2a_0} \tag{10.43}$$

is the absolute value of the unperturbed energy. For hydrogen the ratio $\varepsilon_1/|E|$ is of order 10^{-10}. For hydrogen-like ions the relative correction increases as $Z^{8/3}$, r_0 being roughly proportional to $Z^{1/3}$. The correction is particularly important for inner orbitals of large atoms, as well as for some exotic atoms (Exercise 10.5).

10.5 Dipole Moment of Atoms and Stark Effect

Consider an atom in a constant electric field of amplitude E. We assume E is weak and use perturbation theory to compute, in the central-field model, the correction to orbital energies produced by the field.

Let $\hat{z}$ be the electric field direction. The corresponding electrostatic potential is equal to $-Ez = -Er\cos\theta$. Following results of Sect. 7.7, the Hamiltonian of an electron acquires an additional term given by

$$W(\boldsymbol{r}) = -q_e E r \cos\theta . \tag{10.44}$$

We will take $W(\boldsymbol{r})$ as a perturbation of the Hamiltonian H_0 of an electron, in the central-field model.[2]

Let us compute the first-order correction to electronic energies due to the perturbation. Once again we must evaluate matrix elements of W between $|nlm\rangle$ and $|nlm'\rangle$. We have

$$\begin{aligned} \langle nlm|W|nlm'\rangle &= \int \mathrm{d}\boldsymbol{r}\, \psi^*_{nlm}(\boldsymbol{r}) W(\boldsymbol{r}) \psi_{nlm'}(\boldsymbol{r}) \\ &= \int_0^\infty \mathrm{d}r\, r^2(-q_e E r)|u_{nl}(r)|^2 \int \mathrm{d}\Omega\, Y^*_{lm}(\theta,\phi)\cos\theta\, Y_{lm'}(\theta,\phi) . \end{aligned} \tag{10.45}$$

Making the change of variables $\theta \to \pi - \theta$ and $\phi \to \phi + \pi$ and using property (7.169) of spherical harmonics, one easily concludes that the integral on solid angle vanishes. Therefore

---

[2] Strictly speaking, we should point out that since $W(\boldsymbol{r}) \to -\infty$ if $r\cos\theta \to -\infty$, the perturbed Hamiltonian no longer has a discrete spectrum. In other words, electrons can reach beyond the potential barrier that holds them. But the probability that an electron escapes is negligible if the field is weak ([32], Sect. 101).

$$\langle nlm|W|nlm'\rangle = 0\,. \tag{10.46}$$

To first order then an electron's energy, and thus the atom's energy, are not changed by the presence of an electric field. In general, however, the atom's energy will be changed in second order. Note that the first-order correction to energy is proportional to the electric field amplitude, whereas the second-order correction is proportional to the square of the amplitude.

In a constant electric field an atom's energy thus contains no term proportional to the field amplitude. This means the atom does not have a permanent electric dipole moment. Indeed in an electric field, an electric dipole acquires an additional energy proportional to E. However, even if the atom does not have a permanent dipole moment, the electric field gives rise to an *induced* dipole moment, by distorting to some extent the electronic distributions. The induced dipole moment is proportional to E, and its energy in the field is proportional to E^2. The effect of the induced dipole moment is represented by second-order corrections to orbitals' energies.

That first-order corrections vanish follows from the fact that orbitals' energies are all different, in which case degeneracy is correctly taken care of. And here hydrogen and hydrogen-like ions are important exceptions, since there orbitals with the same n and different l have the same energy. Take for instance orbitals with $n = 2$. There are four states with the same energy, represented by $|200\rangle$, $|210\rangle$, $|211\rangle$ and $|21{-}1\rangle$. One easily sees that 14 out of 16 matrix elements vanish. The two remaining ones are given by

$$\langle 200|W|210\rangle = -q_e E M = \langle 210|W|200\rangle^*\,, \tag{10.47}$$

where (Exercise 10.6)

$$M = \int_0^\infty \mathrm{d}r\, r^3 u_{20}^*(r) u_{21}(r) \int \mathrm{d}\Omega\, Y_{00}^*(\theta,\phi)\cos\theta\, Y_{10}(\theta,\phi) = -\frac{3a_0}{Z}\,. \tag{10.48}$$

Matrix elements of W can thus be displayed as

$$\begin{array}{cc} \begin{matrix} (200) \\ (210) \\ (211) \\ (21{-}1) \end{matrix} & \begin{pmatrix} 0 & 3q_e E a_0/Z & 0 & 0 \\ 3q_e E a_0/Z & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}\,,$$

where the numbering of rows is shown in parentheses.

Eigenvalues and eigenvectors of this matrix yield first-order corrections $\varepsilon_{1,\alpha}$ and zeroth-order kets $|\chi_0^\alpha\rangle$ as

$$\begin{aligned} \varepsilon_{1,1} &= -3q_e E a_0/Z\,, \qquad |\chi_0^1\rangle = \frac{1}{\sqrt{2}}\{|200\rangle - |210\rangle\}\,; \\ \varepsilon_{1,2} &= 3q_e E a_0/Z\,, \qquad |\chi_0^2\rangle = \frac{1}{\sqrt{2}}\{|200\rangle + |210\rangle\}\,; \end{aligned} \tag{10.49}$$

$$\varepsilon_{1,3} = 0\,, \qquad |\chi_0^3\rangle = |211\rangle\,;$$

$$\varepsilon_{1,4} = 0\,, \qquad |\chi_0^4\rangle = |21{-1}\rangle\,.$$

In excited states, hydrogen and hydrogen-like ions can thus have an additional energy proportional to the amplitude of the applied field. There is then a permanent electric dipole moment.

The *Stark effect* is the splitting of spectral lines due to the electric field's partial removal of degeneracy. The effect is proportional to the field amplitude in some states of hydrogen and hydrogen-like ions, and to the square of the amplitude in other atoms.

10.6 Van der Waals Forces

When two atoms are not too far, they exert on one another rather complicated forces. At short distances the interaction of electronic clouds entails strong repulsion. At larger distances, on the other hand, the interaction is weak and attractive. These attractive forces come mainly form charge polarization and the resulting dipoles' electrostatic interaction. They are called *van der Waals forces.*

We will investigate van der Waals forces in the simplest case, namely two hydrogen atoms. We assume that both nuclei are rest or, equivalently, that they have infinite mass.

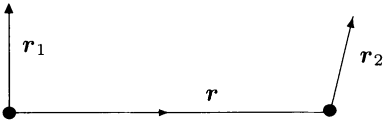

Fig. 10.1. Two hydrogen atoms separated by a distance r

The coordinates we will use are shown in Fig. 10.1. In the first atom $\boldsymbol{r}_1$ stands for the electron's coordinates with respect to the nucleus. Likewise with $\boldsymbol{r}_2$ in the second atom. The symbol $\boldsymbol{r}$ ($\neq \boldsymbol{r}_2 - \boldsymbol{r}_1$) denotes the vector going from the first to the second nucleus.

Clearly, if interaction between particles of the first atom and particles of the second one could be neglected, the system's Hamiltonian would be given by[3]

$$H_0 = -\frac{\hbar^2}{2m}\nabla_1^2 - \frac{e^2}{r_1} - \frac{\hbar^2}{2m}\nabla_2^2 - \frac{e^2}{r_2}\,. \tag{10.50}$$

Taking into account the interaction between the two atoms, one can see that the Hamiltonian becomes $H = H_0 + W$, where

---

[3] Once again spin and relativistic effects are neglected.

$$W = \frac{e^2}{r} + \frac{e^2}{|\boldsymbol{r} + \boldsymbol{r}_2 - \boldsymbol{r}_1|} - \frac{e^2}{|\boldsymbol{r} + \boldsymbol{r}_2|} - \frac{e^2}{|\boldsymbol{r} - \boldsymbol{r}_1|} . \tag{10.51}$$

These four terms represent the electrostatic interaction between the two nuclei, the two electrons, the first nucleus and the second electron, and the second nucleus and the first electron.

Van der Waals forces, as we said, occur when the distance between atoms is relatively large, i.e. large compared with the mean distance between each electron and its nucleus. This means we can let $r \gg \langle r_1 \rangle$ and $r \gg \langle r_2 \rangle$. Matrix elements of W are then small compared with the lowest eigenvalues of H_0.

We will evaluate van der Waals forces by means of perturbation theory. We will first consider the case with two atoms in their ground state, and next the case with one atom in its ground state and the other atom in its first excited state.

The wave function of two hydrogen atoms in their ground state is given by

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \psi_{100}(\boldsymbol{r}_1)\psi_{100}(\boldsymbol{r}_2) . \tag{10.52}$$

The corresponding energy (with spin neglected) is not degenerate. The first-order correction to energy due to the perturbation can be written in obvious notation by

$$\varepsilon_1 = \langle 100; 100 | W | 100; 100 \rangle . \tag{10.53}$$

One can see that $\varepsilon_1 = 0$. This follows from an expansion of each term in the expression of W in a series of spherical harmonics, according to (7.173) and (7.174). The four terms involving Y_{00} cancel each other, so that there only remain spherical harmonics Y_{lm} with $l > 0$. Clearly none of these functions contributes to ε_1.

Let us now turn to the second-order correction to energy. To compute it we will keep only the leading term in the expansion of W in powers of r^{-1}.[4] One can show (Exercise 10.8) that this term is given by

$$W \rightarrow W_{dd} = \frac{e^2}{r^3} \{ \boldsymbol{r}_1 \cdot \boldsymbol{r}_2 - 3(\boldsymbol{r}_1 \cdot \hat{r})(\boldsymbol{r}_2 \cdot \hat{r}) \} . \tag{10.54}$$

This is just the interaction energy of two dipoles separated by a distance r.

The second-order correction to energy is given by

$$\varepsilon_2 = \mathop{\mathrm{S}}\nolimits_{i \neq n, \alpha} \frac{|\langle E_n | W_{dd} | E_i^\alpha \rangle|^2}{E_n - E_i} . \tag{10.55}$$

---

[4] There is always a small probability to find an electron very far from its nucleus. The expansion of W in powers of r^{-1} is not valid in that part of configuration space. But for all practical purposes, the exponential decrease of wave functions eliminates the contributions we are neglecting.

Here $|E_n\rangle$ coincides with $|100;100\rangle$. The summation is carried out on the discrete and on the continuous spectrum. Clearly ε_2 is proportional to r^{-6}.

Although the sum cannot be computed explicitly, it is obvious that ε_2 is negative since $E_n < E_i$ for all $i \neq n$. Moreover, one can see that $\langle E_n|W_{dd}|E_i^\alpha\rangle$ vanishes if $|E_i^\alpha\rangle$ represents a state where one of the hydrogen atoms is in its ground state. The lowest value of E_i that brings a nonvanishing contribution to ε_2 thus coincides with the energy of two hydrogen atoms in their first excited state. Denoting that energy by E' one sees that

$$\begin{aligned}\varepsilon_2 &\geq \frac{1}{E_n - E'}\mathsf{S}_{i\neq n,\alpha}\,|\langle E_n|W_{dd}|E_i^\alpha\rangle|^2 \\ &\geq \frac{1}{E_n - E'}\left\{\mathsf{S}_{i,\alpha}\,|\langle E_n|W_{dd}|E_i^\alpha\rangle|^2 - |\langle E_n|W_{dd}|E_n\rangle|^2\right\} \\ &\geq \frac{1}{E_n - E'}\langle E_n|W_{dd}^2|E_n\rangle\ .\end{aligned}$$

The right-hand side is evaluated in Exercise (10.9). Ones find that

$$\varepsilon_2 \geq -\frac{8e^2a_0^5}{r^6}\,. \tag{10.56}$$

Thus we have shown that if two hydrogen atoms in their ground state are separated by a distance r large enough, their interaction energy is negative and proportional to r^{-6}.[5] Such energy corresponds to an attractive force. The r^{-6} character corresponds to what one intuitively expects for the interaction energy of two induced dipoles. Indeed a dipole moment fluctuation in one of the atoms induces in the other an electric field and a dipole moment proportional to r^{-3}. These two moments' interaction energy is proportional to r^{-6}.

So much for two hydrogen atoms in their ground state. Let us now investigate van der Waals forces in the case where one atom is in its ground state, while the other one is in its first excited state.

Here energy is eightfold degenerate. The eight vectors $|E_n^\alpha\rangle$ are given by $|100;200\rangle$, $|100;21m\rangle$, $|200;100\rangle$ and $|21m;100\rangle$, where $m = -1, 0, 1$.

We will soon see that, in contrast with the previous case, first-order corrections to energy do not vanish and have a term going like r^{-3}. So we will at once replace W by W_{dd}. The matrix $\langle E_n^\alpha|W_{dd}|E_n^\beta\rangle$ has 64 elements. Without loss of generality, one can let $\hat{r} = \hat{z}$ in the expression for W_{dd}. A careful analysis then shows that of the 64 matrix elements, only six do not vanish.

---

[5] One can show ([178], Sect. 47; [200], Sect. 32) that the leading term in the interaction energy is approximately given by $-6.5\,e^2a_0^5r^{-6}$. This result, however, no longer holds if r is really very large. In that case one must take into account retardation effects in the propagation of electromagnetic interactions ([141], Sect. 90). Due to these effects, the interaction energy at very large distance is proportional to r^{-7}. Note that retardation effects must be taken into account if r is of the order of wavelengths associated with atomic transitions, that is if $r/a_0 > 10^3$.

Of these, three have the form $\langle 100; 21m|W_{dd}|21m; 100\rangle$ while the other three are the transpose elements. The 8×8 matrix is block-diagonal, with three non-vanishing blocks like

$$\begin{pmatrix} 0 & \alpha_m r^{-3} \\ \alpha_m^* r^{-3} & 0 \end{pmatrix} . \tag{10.57}$$

Here we have let

$$\frac{\alpha_m}{r^3} = \langle 100; 21m|W_{dd}|21m; 100\rangle . \tag{10.58}$$

The constants α_m are computed in Exercise (10.10). They are real.

Clearly the first-order correction to energies is given by $\alpha_m r^{-3}$ and $-\alpha_m r^{-3}$. The zeroth-order associated eigenvectors are given by

$$\frac{1}{\sqrt{2}} \{|100; 21m\rangle + |21m; 100\rangle\} \quad \text{and} \quad \frac{1}{\sqrt{2}} \{|100; 21m\rangle - |21m; 100\rangle\} . \tag{10.59}$$

The correction to energies can be either positive or negative. This entails that depending on the state of the system, forces can be either repulsive or attractive.

10.7 Rotation and Vibration of Diatomic Molecules

A diatomic molecule is made of two atomic nuclei with electrons going around them. Inner electrons occupy orbitals that are very similar to atomic orbitals. Outer electrons, on the other hand, occupy very different orbitals. They are not associated with a specific nucleus.

Denote by $\boldsymbol{r}$ the vector from one nucleus to the other. Under certain hypotheses one can associate with the nuclei a wave function $\psi(\boldsymbol{r})$ satisfying the equation[6]

$$-\frac{\hbar^2}{2M}\nabla^2\psi(\boldsymbol{r}) + V(r)\psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}) . \tag{10.60}$$

Here M is the reduced mass of the nuclei, $V(r)$ is a spherically symmetric potential and E is the molecule's energy in the center-of-mass frame. Since electrons are in a given energy level, E represents, up to an additive constant, the energy of motion of nuclei.

The potential has a minimum at $r = r_0$ so that one can write

$$V(r) = V(r_0) + \frac{1}{2}M\omega_0^2(r - r_0)^2 + \cdots \tag{10.61}$$

---

[6] In Chap. 20 we will see the conditions under which the equation for $\psi(\boldsymbol{r})$, and the one we will obtain for $\bar{u}(r)$, are valid.

In the following discussion, we will show that energies associated with this potential are much like the ones of a harmonic oscillator with frequency ω_0. This motivates the notation used in the quadratic term. In every molecule, ω_0 is such that the oscillation amplitude in the first few vibrational levels is much smaller than r_0. Moreover $\omega_0 \gg \hbar M^{-1} r_0^{-2}$. This means that vibrational energies ($\approx \hbar\omega_0$) are much larger than rotational energies ($\approx \hbar^2 M^{-1} r_0^{-2}$).

Since the potential is spherically symmetric one can let

$$\psi(\boldsymbol{r}) = \frac{\bar{u}(r)}{r} Y_{LM}(\theta, \phi).$$

The equation for $\bar{u}(r)$ was obtained in Sect. 8.4. One has

$$-\frac{\hbar^2}{2M}\frac{\mathrm{d}^2\bar{u}}{\mathrm{d}r^2} + V_{\mathrm{eff}}(r)\bar{u}(r) = E\bar{u}(r), \tag{10.62}$$

where

$$V_{\mathrm{eff}}(r) = V(r) + \frac{\hbar^2 L(L+1)}{2Mr^2}. \tag{10.63}$$

The function $V_{\mathrm{eff}}(r)$ has a minimum at a value of r that can be denoted by $\bar{r}$. We will see that if L is not too large, $\bar{r}$ is very close to r_0. At any rate one can write

$$V_{\mathrm{eff}}(r) = V_{\mathrm{eff}}(\bar{r}) + \frac{1}{2}M\omega^2(r-\bar{r})^2 + a_3(r-\bar{r})^3 + a_4(r-\bar{r})^4 + \cdots \tag{10.64}$$

The coefficients a_3 and a_4 are easily obtained if the explicit form of $V(r)$ is known. In any case let $x = r - \bar{r}$. Equation (10.62) can then be written as

$$-\frac{\hbar^2}{2M}\frac{\mathrm{d}^2\bar{u}}{\mathrm{d}x^2} + \left\{\frac{1}{2}M\omega^2 x^2 + a_3 x^3 + a_4 x^4\right\}\bar{u} = \{E - V_{\mathrm{eff}}(\bar{r})\}\,\bar{u}. \tag{10.65}$$

With two differences, this is the equation of a one-dimensional harmonic oscillator. The first difference lies in the addition of cubic and quartic terms. The second one comes from the fact that the domain of the variable x is the interval $(-\bar{r}, \infty)$ rather than the interval $(-\infty, \infty)$.

The second difference is of no consequence if only the first few vibrational levels are considered. Since the amplitude of vibration is then much smaller than $r_0 \approx \bar{r}$, it is essentially equivalent to have $\bar{u}(x)$ vanish at $-\bar{r}$ or at $-\infty$. As for the cubic and quartic terms, we will take them as a perturbation of the harmonic oscillator potential. Here again the assumption is justified in the first few vibrational levels.

In diatomic molecules the size of the coefficients a_3 and a_4 is such that $a_3^2 \approx a_4 M\omega^2$. This entails that second-order corrections to energy due to the cubic term are of the same order of magnitude as first-order corrections due to the quartic term. One can show (Exercise 10.12) that first-order corrections due to the cubic term vanish. Second-order corrections are given by

$$\varepsilon_2(a_3) = -a_3^2 \frac{\hbar^2}{M^3\omega^4} \left[\frac{15}{4} \left(n + \frac{1}{2} \right)^2 + \frac{7}{16} \right] . \tag{10.66}$$

Here n is the quantum number associated with the vibrational level considered. On the other hand, first-order corrections due to the quartic term are given by

$$\varepsilon_1(a_4) = a_4 \frac{3\hbar^2}{2M^2\omega^2} \left[\left(n + \frac{1}{2} \right)^2 + \frac{1}{4} \right] . \tag{10.67}$$

Let us now go back to (10.65) for the function $\bar{u}$. Obviously the constant $E - V_{\text{eff}}(\bar{r})$ is equal to the harmonic oscillator's energy, corrected with the terms just obtained. The energy of motion of the nuclei is thus given by

$$\begin{aligned} E = V_{\text{eff}}(\bar{r}) + \hbar\omega \left(n + \frac{1}{2} \right) + \frac{\hbar^2}{M^3\omega^4} \left\{ \frac{3}{2} M\omega^2 a_4 \left[\left(n + \frac{1}{2} \right)^2 + \frac{1}{4} \right] \right. \\ \left. - a_3^2 \left[\frac{15}{4} \left(n + \frac{1}{2} \right)^2 + \frac{7}{16} \right] \right\} . \end{aligned} \tag{10.68}$$

Clearly this result holds only for moderate values of n. Otherwise corrections (proportional to n^2) are more important than the unperturbed energy (proportional to n). In general the correction term is negative, so that vibrational levels come closer as n increases.

To correctly interpret (10.68), we must examine the way the "constants" $\bar{r}$ and ω depend on L, the quantum number associated with the molecule's angular momentum. By definition $\bar{r}$ is the value of r that minimizes $V_{\text{eff}}(r)$. Hence

$$\left(\frac{\mathrm{d}V_{\text{eff}}}{\mathrm{d}r} \right)_{\bar{r}} \equiv \left(\frac{\mathrm{d}V}{\mathrm{d}r} \right)_{\bar{r}} - \frac{\hbar^2 L(L+1)}{M\bar{r}^3} = 0. \tag{10.69}$$

But for $\bar{r} \approx r_0$

$$\left(\frac{\mathrm{d}V}{\mathrm{d}r} \right)_{\bar{r}} = \left(\frac{\mathrm{d}V}{\mathrm{d}r} \right)_{r_0} + (\bar{r} - r_0) \left(\frac{\mathrm{d}^2V}{\mathrm{d}r^2} \right)_{r_0} = (\bar{r} - r_0) M\omega_0^2 . \tag{10.70}$$

From (10.69) and (10.70) one gets that

$$\bar{r} = r_0 \left\{ 1 + \frac{\hbar^2 L(L+1)}{M^2\omega_0^2 r_0^4} \right\} . \tag{10.71}$$

As expected the distance between nuclei increases with angular momentum. The faster the molecule rotates, the more it stretches.

To express ω as a function of ω_0, one computes in two different ways the second derivative of V_{eff}. Thus

$$\left(\frac{\mathrm{d}^2 V_{\text{eff}}}{\mathrm{d}r^2}\right)_{r_0} = \left(\frac{\mathrm{d}^2 V}{\mathrm{d}r^2}\right)_{r_0} + \frac{3\hbar^2 L(L+1)}{M r_0^4} = M\omega^2 + 6a_3(r_0 - \bar{r}).$$

Manipulating one finds

$$\omega = \omega_0 \left\{ 1 + \frac{3\hbar^2 L(L+1)}{2M^2\omega_0^2 r_0^4} \left[1 + \frac{2a_3 r_0}{M\omega_0^2} \right] \right\}. \tag{10.72}$$

The oscillator's frequency therefore changes with angular momentum. In general the expression between square brackets is negative, so that frequency decreases as L grows larger.

The minimum value of V_{eff} also depends on L. To evaluate it we write

$$V_{\text{eff}}(r_0) = V(r_0) + \frac{\hbar^2 L(L+1)}{2M r_0^2} = V_{\text{eff}}(\bar{r}) + \frac{1}{2} M\omega^2 (r_0 - \bar{r})^2 .$$

Substituting ω by ω_0 in the last term and making use of (10.71) we get

$$V_{\text{eff}}(\bar{r}) = V(r_0) + \frac{\hbar^2 L(L+1)}{2M r_0^2} - \frac{1}{2M} \left(\frac{\hbar^2 L(L+1)}{M\omega_0 r_0^3} \right)^2 . \tag{10.73}$$

In the right-hand side of (10.73), the second term's numerator represents the molecule's angular momentum, whereas the denominator represents twice the moment of inertia with respect to an axis going through the center of mass and perpendicular to the segment joining the nuclei. Hence the second term coincides with a rotational energy. The latter is corrected (always negatively) by the third term of order L^4.

To obtain the final expression for the energy of motion of the nuclei, one substitutes (10.73) in (10.68) and one replaces ω, in the second term on the right-hand side of (10.68), by (10.72). We will not write down the result explicitly.

The analysis of rotational and vibrational spectra of molecules constitutes an important field in experimental chemistry.[7] The example of a diatomic molecule shows that perturbation theory helps understanding the arrangement of rotational and vibrational energies. In Chap. 20 we will see what transitions are allowed between these levels.

Exercises

10.1. On each side of (10.8), obtain the coefficient of the term proportional to λ^k.

---

[7] See for instance [15] and [109], Chap. 3.

10.2. Find the second-order correction $|\chi_2\rangle$ of the state vector, in the case of a nondegenerate energy [i.e. (10.24)].

10.3. Let H_0 be an unperturbed Hamiltonian and W a perturbation given, in an orthonormal basis $|E_1\rangle, |E_2\rangle$ and $|E_3\rangle$, by

$$[H_0] = \begin{pmatrix} E_1 & 0 & 0 \\ 0 & E_2 & 0 \\ 0 & 0 & E_3 \end{pmatrix}, \qquad [W] = \begin{pmatrix} 0 & a & 0 \\ a^* & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Assume the three values E_1, E_2 and E_3 are all different. Compute in first-order perturbation theory the corrections to eigenvalues and eigenvectors.

10.4. Let H_0 be an unperturbed Hamiltonian and W a perturbation given by $(E_1 \neq E_3)$

$$[H_0] = \begin{pmatrix} E_1 & 0 & 0 \\ 0 & E_1 & 0 \\ 0 & 0 & E_3 \end{pmatrix}, \qquad [W] = \begin{pmatrix} 0 & a & b \\ a^* & 0 & 0 \\ b^* & 0 & 0 \end{pmatrix}.$$

a) Find in first-order perturbation theory the eigenvectors of the perturbed Hamiltonian $H_0 + W$.
b) Find eigenvalues to second order, and compare the result obtained by computing, to this order, the eigenvalues of the matrix $[H_0] + [W]$.

10.5. A nucleus of uniformly distributed charge $-Zq_e$ and radius $r_0 = 1.5\,Z^{1/3} \times 10^{-15}$ m forms a bound state with $Z-1$ electrons and one negative muon. The muon is in a $1s$ orbital. Its mass is equal to $207\,m_e$, where m_e is the electron's mass, and its charge is equal to the electron's. One can take the nuclear mass infinite.
a) What is the muon's energy, if nuclear extension is neglected?
b) What is the first-order correction to energy due to nuclear extension?
c) For what value of Z does the correction to the muon's energy become comparable to the energy itself? [For the muon one can use a hydrogen-like orbital.]

10.6. Compute the integral (10.48).

10.7. In this problem we will study the Stark effect in the third energy level ($n = 3$) of hydrogen. The energy is ninefold degenerate. Thus the matrix $W_{\alpha\beta}$ has 81 elements. Fortunately, there are only 8 elements different from zero.
a) What are these 8 elements (no need to compute them explicitly)?
b) Since $W_{\alpha\beta}$ is Hermitian and real, the eight nonzero elements are pairwise equal. Moreover, the explicit form of spherical harmonics leads to an additional equality between these elements. Write the 8 nonzero elements of $W_{\alpha\beta}$ in terms of three parameters (calculable, but not to be computed here).

c) Find the eigenvalues and eigenvectors of the matrix $W_{\alpha\beta}$.
d) Make a diagram of perturbed energy levels, in first-order perturbation theory. In each case indicate the corresponding zeroth-order eigenvectors.

10.8. Show that for $r \gg r_1$ and $r \gg r_2$, the leading term of (10.51) is given by (10.54).

10.9. Evaluate $\langle 100; 100|W_{dd}^2|100; 100\rangle$, where W_{dd} is given in (10.54).

10.10. Show that the constants α_m defined in (10.58) are given by

$$\alpha_1 = \alpha_{-1} = -\frac{1}{2}\alpha_0 = e^2 a_0^2 \left(\frac{8}{9}\right)^5 .$$

10.11. A one-dimensionnal harmonic oscillator with angular frequency ω is perturbed by a term λX. Use perturbation theory to find to second order in λ the new value of the ground-state energy. Compare with the exact value of the ground-state energy of such a perturbed oscillator.

10.12. Let H_0 be the harmonic oscillator Hamiltonian. Find the first-order correction to energy due to a perturbation proportional to X^4. Obtain the second-order correction due to a perturbation proportional to X^3.

11 Stationary Scattering States

Scattering is one of the two most important methods for the experimental investigation of atomic and molecular properties (the other being spectroscopy). We will treat scattering by means of the Hamiltonian's eigenvalue equation, focussing on the continuous spectrum associated with a given potential. After defining the scattering cross section, we will transform the eigenvalue equation into an integral equation for the wave function. The integral equation can be solved by a series expansion which, when truncated, gives rise to the Born approximation. For a spherically symmetric potential it is convenient to go back to the eigenvalue equation, and solve it by a series of Legendre polynomials. We will investigate analytical properties of the coefficients of that expansion as well as certain conditions giving rise to resonance phenomena. Finally we will analyze scattering by a periodic potential.[1]

11.1 Scattering Cross Section

A typical scattering setup is shown in Fig. 11.1. A number of type A particles, called *scattering centers*, are subject to a flux of type B particles, called *incoming particles.* The scattering centers, all identical, are for instance molecules, atoms, ions or atomic nuclei. The incoming particles, also assumed all identical, are for instance photons, electrons, neutrons, alpha particles, etc. We assume the incoming beam is made up of a constant flux of n_B particles per unit area and unit time, all moving in the same direction. Incoming particles interact with scattering centers, to produce a number of *scattered particles.* We will assume the incoming and scattered particles are of the same type. Actually this implies that the incoming particles' kinetic energies are small compared with energies associated with their mass or their dissociation.

Let $\mathrm{d}n(\theta, \phi)$ be the number of particles scattered per unit time in an element of solid angle $\mathrm{d}\Omega$ at angles (θ, ϕ). Under hypotheses soon to be specified one can write

---

[1] Quantum-mechanical scattering theory is a very wide subject. For additional material the reader is referred to [169] and, for a particularly pedagogical exposition, to [224].

L. Marchildon, *Quantum Mechanics*

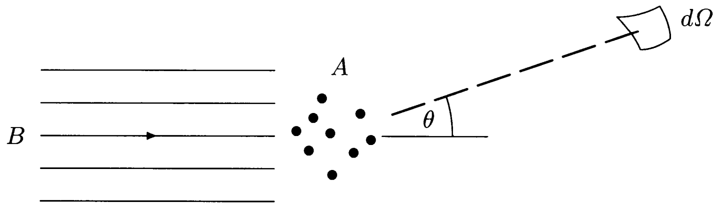

Fig. 11.1. Scattering of particles of type B by particles of type A

$$\mathrm{d}n(\theta,\phi) = n_B N_A \sigma(\theta,\phi)\mathrm{d}\Omega\,, \tag{11.1}$$

where N_A is the total number of scattering centers. The function $\sigma(\theta,\phi)$ is called the *differential scattering cross section.* This is the quantity extracted from every scattering experiment. It is of cardinal importance.

Let us now examine the conditions of validity of (11.1).

i) The number of scattered particles will be proportional to the flux of incoming particles if the latter are independent of one another, and interact only one at a time with the scattering centers.
ii) $\mathrm{d}n$ will be proportional to the number of scattering centers if these are all subject to the incoming flux and if N_A is small enough that an incoming particle almost never interacts with two scattering centers.[2]
iii) $\mathrm{d}n$ will be proportional to the solid angle element $\mathrm{d}\Omega$ if the latter is small enough and if the distance between scattering centers and detector is much larger than linear dimensions of the volume occupied by scattering centers. Of course, the detector should not be subject to the incoming flux directly.

We have defined $\mathrm{d}n(\theta,\phi)$ as the number of scattered particles per unit time in the element $\mathrm{d}\Omega$ at angles (θ,ϕ). Clearly in many experimental situations, a more detailed analysis can be made. Scattered particles can be sorted out with respect to energy, value of a spin component, etc. Thus one is led to define a differential cross section that depends not only on angles but also on energy, spin, etc. In this chapter we will not introduce such additional variables. The energy of incoming and scattered particles will be the same (this is *elastic* scattering) and internal degrees of freedom will not be considered.

The *total scattering cross section,* denoted by σ, is defined as the integral of the differential cross section over all solid angle:

$$\sigma = \int \mathrm{d}\Omega\, \sigma(\theta,\phi)\,. \tag{11.2}$$

Note that measurement of $\sigma(\theta,\phi)$ near $\theta = 0$ and $\theta = \pi$ is particularly complicated, due to the presence of the incoming beam. Values of $\sigma[(\theta = 0,\pi),\phi]$ are actually obtained by extrapolation from neighboring angles.

---

[2] For $\mathrm{d}n$ to be proportional to N_A, the positions of scattering centers should also be randomly distributed. We will come back to this in Sect. 11.9.

The differential cross section (like the total cross section) has dimensions of an area. Specifically $N_A\sigma(\theta,\phi)$ is the area which, on the incoming beam, corresponds to the number of particles detected per unit solid angle in direction (θ,ϕ).

We have not yet mentioned the recoil of scattering centers after interaction with incoming particles. The effect is negligible if the mass of scattering centers is much larger than the mass of incoming particles. If, however, masses are comparable, the effect is important. In this case the scattering process can favorably be viewed in the center-of-mass reference frame. The result is then translated in the coordinates of the laboratory frame. In this chapter we will focus exclusively on potential scattering. If the potential represents an interaction between two particles, the formalism can be interpreted either as the scattering of a particle of mass m by an infinitely massive particle in the laboratory frame, or as the scattering of a system with reduced mass m in the center-of-mass frame.

To close this section, let us come back to the second condition of validity mentioned above, which can now be expressed quantitatively. If there are N_A scattering centers in volume V, the total area hidden by these is of order $N_A\sigma$. The probability that an incoming particle interacts twice is small compared with the probability that is interacts once if the latter is much smaller than 1. The probability of a single interaction is of the order of the area hidden by scattering centers divided by the area of the bulk ($\approx V^{2/3}$). The second condition of validity can therefore be written as

$$\frac{N_A\sigma}{V^{2/3}} \ll 1\,. \tag{11.3}$$

11.2 Stationary Scattering States

The scattering of a particle by another is a dynamic process which involves time in an essential way. Nevertheless, scattering by a potential can be treated by means of the Hamiltonian's eigenvalue equation. Assume two particles interact through a potential $V(\boldsymbol{r}_1 - \boldsymbol{r}_2)$. Let m be the system's reduced mass. The discussion in Sect. 9.1 shows that in terms of relative coordinates $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$, the eigenvalue equation is given by

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\boldsymbol{r})+V(\boldsymbol{r})\psi(\boldsymbol{r}) = E\psi(\boldsymbol{r})\,. \tag{11.4}$$

This equation also holds for the scattering of a particle of mass m by a potential $V(\boldsymbol{r})$.

At this stage we should point out that the results we will obtain do not necessarily hold with all kinds of potential. Throughout this chapter we will assume the potential satisfies the following conditions:

i) $V(\boldsymbol{r})$ is not more singular than r^{-1} around $r = 0$;

ii) $V(\boldsymbol{r})$ decreases faster than r^{-3} as $r \to \infty$;
iii) $V(\boldsymbol{r})$ is piecewise continuous and has finite discontinuities only.

Unless otherwise indicated, all the results we will obtain will hold for such potentials. Some results may apply to more general situations, for instance to potentials that decrease faster than r^{-2} at infinity or that have a singularity stronger than r^{-1} at the origin.

We are concerned here with positive values of E. In the coming sections we will show that for large r the eigenvalue equation has solutions like

$$\psi(\boldsymbol{r}) = \exp\{\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r}\} + \frac{1}{r} f(\theta, \phi) \exp(\mathrm{i}kr). \tag{11.5}$$

Here θ and ϕ are the polar and azimuthal angles of $\boldsymbol{r}$ with respect to the direction of $\boldsymbol{k}$, and $\boldsymbol{k}$ is any real vector such that

$$E = \frac{\hbar^2}{2m} \boldsymbol{k} \cdot \boldsymbol{k} = \frac{\hbar^2 k^2}{2m}. \tag{11.6}$$

The function $f(\theta, \phi)$ is called the *scattering amplitude.*

A solution like (11.5) represents the superposition of a plane wave directed along $\hat{k}$ and of an outgoing wave centered on $r = 0$. As it stands, this solution does not quite represent a scattering process, since a beam of incoming particles is never infinitely extended in the plane perpendicular to its direction of propagation. To represent the scattering process correctly, the term $\exp\{\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r}\}$ ought to be modulated by a factor $F(u, v)$, where u and v are coordinates in the plane perpendicular to $\boldsymbol{k}$. The factor F should have the following properties:

i) $F(u, v) = 1$ if $|u|$ and $|v|$ are not too large;
ii) $F(u, v) \to 0$ if $|u| \to \infty$ or if $|v| \to \infty$.

We will not explicitly write this factor but we will assume it is there. In other words we assume our treatment could be generalized so as to include it, without significantly changing the results we are about to obtain.[3]

Let us now compute the probability current associated with $\psi(\boldsymbol{r})$. We have

$$\boldsymbol{j}(\boldsymbol{r}) = -\frac{\hbar}{m} \mathrm{Im}\,(\psi \boldsymbol{\nabla} \psi^*).$$

The probability current associated with the wave function (11.5) is complicated. It involves many superposition terms. If, however, one takes into account the factor F introduced above, one can draw the following conclusions:

---

[3] For a treatment of the scattering process that takes into account the finite size of incoming particles and the fact that their energy is not perfectly well-defined, see [224], Chap. 3.

i) If r is large and $\theta = 0$ or π, then the term in r^{-1} is negligible compared with the term $\exp\{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}\}$. The probability current is given by

$$\boldsymbol{j}(\boldsymbol{r}) = -\frac{\hbar}{m}\mathrm{Im}\left\{\exp\left(\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}\right)\boldsymbol{\nabla}\left[\exp\left(-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}\right)\right]\right\} = \frac{\hbar}{m}\boldsymbol{k}. \tag{11.7}$$

ii) If r if large and θ is not too close to 0 or π, the term $\exp\{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}\}$ disappears due to the multiplicative factor F. One then finds (Exercise 11.1)

$$\begin{aligned}\boldsymbol{j}(\boldsymbol{r}) &= -\frac{\hbar}{m}\mathrm{Im}\left\{\frac{f}{r}\mathrm{e}^{\mathrm{i}kr}\,\boldsymbol{\nabla}\left(\frac{f^*}{r}\mathrm{e}^{-\mathrm{i}kr}\right)\right\}\\ &= \frac{\hbar}{m}\left\{\frac{k}{r^2}|f|^2\hat{r} - \frac{1}{r^2}\mathrm{Im}\left(f\boldsymbol{\nabla}f^*\right)\right\}.\end{aligned} \tag{11.8}$$

Since $\boldsymbol{\nabla}f^*$ is perpendicular to $\hat{r}$ one gets

$$\boldsymbol{j}(\boldsymbol{r})\cdot\hat{r} = \frac{\hbar k}{m}\frac{|f(\theta,\phi)|^2}{r^2}. \tag{11.9}$$

Thus $\psi(\boldsymbol{r})$ represents the superposition of an incoming plane wave along $\hat{k}$ and of an outgoing wave, associated with scattered particles, with decreasing intensity proportional to r^{-2}. The number of incoming particles per unit surface and time is equal to $\boldsymbol{j}(\boldsymbol{r})\cdot\hat{k}$, with $\boldsymbol{j}(\boldsymbol{r})$ given in (11.7). The number of particles scattered per unit time in the solid angle $\mathrm{d}\Omega$ about (θ,ϕ) is equal to $\boldsymbol{j}(\boldsymbol{r})\cdot\hat{r}\,r^2\mathrm{d}\Omega$, with $\boldsymbol{j}(\boldsymbol{r})$ given in (11.8). Hence the differential cross section is obtained from (11.1) as $(N_A = 1)$

$$\sigma(\theta,\phi) = \frac{1}{\boldsymbol{j}\cdot\hat{k}}\frac{1}{\mathrm{d}\Omega}\boldsymbol{j}(\boldsymbol{r})\cdot\hat{r}\,r^2\mathrm{d}\Omega = |f(\theta,\phi)|^2. \tag{11.10}$$

Thus the cross section is directly extracted from the solution $\psi(\boldsymbol{r})$, provided the latter is known explicitly.

For a spherically symmetric potential, the solutions (11.5) can be developed (formally at least) directly from the differential equation. This we will do in Sect. 11.6. But before, we will investigate the more general case of a potential that is not necessarily spherically symmetric. There it is more convenient to transform the differential equation into an integral equation. For this purpose we will use the technique of Green's operators and functions. They are particularly important in quantum mechanics and we now proceed to define them.

11.3 Green's Operators and Functions

Let H be the Hamiltonian of a particle in three dimensions. One can write $H = H_0 + V$, where H_0 is the free-particle Hamiltonian. As above we assume that V does not depend on time and that $V(\infty) = 0$.

Let λ be a complex variable. The *Green's operator* associated with the Hamiltonian H is defined as

$$G(\lambda) = (\lambda I - H)^{-1} . \tag{11.11}$$

Here I stands for the identity operator. The Green's operator is a function of the variable λ.

The Green's operator is related to the evolution operator. Indeed for any λ in the upper half-plane one can write

$$\begin{aligned} G(\lambda) &= -\frac{\mathrm{i}}{\hbar} \int_0^\infty \mathrm{d}T \exp\left\{ \frac{\mathrm{i}T}{\hbar}(\lambda I - H) \right\} \\ &= -\frac{\mathrm{i}}{\hbar} \int_0^\infty \mathrm{d}T \exp\left\{ \frac{\mathrm{i}T\lambda}{\hbar} \right\} U(T,0) . \end{aligned} \tag{11.12}$$

The spectral decomposition of the Green's operator can be obtained from (2.50). One finds that

$$G(\lambda) = \mathrm{S}_E \frac{1}{\lambda - E} |E\rangle\langle E| . \tag{11.13}$$

Kets $|E\rangle$ are eigenvectors of H. The notation keeps up with the fact that the spectrum of H is, in general, partly discrete and partly continuous.

Equation (11.13) is singular if λ coincides with an eigenvalue of H, that is, a negative value corresponding to a bound-state energy or any positive value.[4]

In the same way as $G(\lambda)$ was defined through the Hamiltonian H, one can define a Green's operator $G_0(\lambda)$ through the free-particle Hamiltonian. Thus

$$G_0(\lambda) = (\lambda I - H_0)^{-1} . \tag{11.15}$$

Equation (11.13) associated with $G_0(\lambda)$ is singular at positive values of λ. Furthermore

$$G_0^{-1}(\lambda) - G^{-1}(\lambda) = V ,$$

---

[4] If the Hamiltonian has a discrete spectrum only [i.e. if $V(\infty) = \infty$], its eigenvalues can be characterized by computing in two different ways the trace of $G(\lambda)$. One uses, on the one hand, (11.13) and the orthonormal basis $\{|E_n\rangle\}$; and, on the other hand, (11.12) and the orthonormal basis $\{|x\rangle\}$. One gets

$$\sum_n \frac{1}{\lambda - E_n} = -\frac{\mathrm{i}}{\hbar} \int_0^\infty \mathrm{d}T \exp\left\{ \frac{\mathrm{i}T\lambda}{\hbar} \right\} \int_{-\infty}^\infty \mathrm{d}x \, U(x,T;x,0) . \tag{11.14}$$

The Hamilotnian's eigenvalues thus coincide with the poles of the right-hand side of (11.14).

which entails that operators G_0 and G satisfy the following relation, called the *Lippmann–Schwinger equation*:

$$G(\lambda) = G_0(\lambda) + G_0(\lambda) V G(\lambda) . \tag{11.16}$$

We will see that this equation leads to a series expansion of the wave function associated with a scattering process.

Let $|\phi\rangle$ be an arbitrary vector. For any value of λ different from an eigenvalue of H, one can define a vector $|\psi\rangle$ as

$$|\psi\rangle = \{I + G(\lambda)V\} |\phi\rangle . \tag{11.17}$$

It is easy to see that

$$(H - \lambda I) |\psi\rangle = (H - \lambda I) \{I + G(\lambda)V\} |\phi\rangle \tag{11.18}$$

$$= (H - \lambda I - V) |\phi\rangle = (H_0 - \lambda I) |\phi\rangle . \tag{11.19}$$

As a matter of fact, this equation is also true for real and positive values of λ, provided that $|\psi\rangle$ is defined appropriately. By construction, we will see that operators $G_0(\lambda)$ and $G(\lambda)$ can be analytically continued to real and positive values of λ. The result, however, is not single-valued. This means that the analytic continuation from values of λ in the upper half-plane does not coincide with the continuation from values in the lower half-plane. For E real and positive and for $\eta > 0$,

$$\lim_{\eta \to 0} G(E + \mathrm{i}\eta) \neq \lim_{\eta \to 0} G(E - \mathrm{i}\eta) .$$

Define two vectors $|\psi_+\rangle$ and $|\psi_-\rangle$ as

$$|\psi_\pm\rangle = \lim_{\eta \to 0} \{I + G(E \pm \mathrm{i}\eta)V\} |\phi\rangle . \tag{11.20}$$

One has

$$\begin{aligned}(H - EI)|\psi_\pm\rangle &= \lim_{\eta \to 0} \{H - (E \pm \mathrm{i}\eta)I\} \lim_{\eta \to 0} \{I + G(E \pm \mathrm{i}\eta)V\} |\phi\rangle \\ &= \lim_{\eta \to 0} [\{H - (E \pm \mathrm{i}\eta)I\} \{I + G(E \pm \mathrm{i}\eta)V\} |\phi\rangle] \\ &= \lim_{\eta \to 0} [\{H_0 - (E \pm \mathrm{i}\eta) I\} |\phi\rangle] = (H_0 - EI)|\phi\rangle . \end{aligned} \tag{11.21}$$

If, therefore, $|\phi\rangle$ is an eigenvector of H_0, then $|\psi_+\rangle$ and $|\psi_-\rangle$ are eigenvectors of H. Hence (11.20) allows, in principle, the construction of eigenvectors of H form eigenvectors of H_0.

The problem is that $G(\lambda)$ is not known explicitly. On the other hand, we will soon see that $G_0(\lambda)$ can be known exactly. From (11.16) it is clear that, if λ is not real,

$$\{I - G_0(\lambda)V\} \{I + G(\lambda)V\} = I . \tag{11.22}$$

If $\lambda = E > 0$, (11.22) remains valid in the sense of analytic continuation. Thus one finds that

$$\lim_{\eta \to 0} \{I - G_0(E \pm \mathrm{i}\eta)V\} |\psi_\pm\rangle = \{I - G_0(E \pm \mathrm{i}0)V\} |\psi_\pm\rangle = |\phi\rangle. \tag{11.23}$$

The notation $i0$ is an abbreviation of the limiting process indicated.

In Sect. 11.4 we will write (11.23) in the coordinate representation, and see that it is really an integral equation for the wave function $\psi_\pm(\boldsymbol{r})$. But (11.23) also leads to a series representation of the vector $|\psi_\pm\rangle$. Indeed

$$\begin{aligned} |\psi_\pm\rangle &= \{I - G_0(E \pm \mathrm{i}0)V\}^{-1} |\phi\rangle \\ &= |\phi\rangle + G_0(E \pm \mathrm{i}0)V|\phi\rangle + [G_0(E \pm \mathrm{i}0)V]^2 |\phi\rangle + \cdots \end{aligned} \tag{11.24}$$

Whether the integral equation or the series representation is used, the determination of $|\psi_\pm\rangle$ from $|\phi\rangle$ requires knowledge of the Green's operator $G_0(E \pm \mathrm{i}0)$. To evaluate $G_0(E \pm \mathrm{i}0)$, it is convenient to go to the coordinate representation. Let $\eta > 0$. For $G_0(E + \mathrm{i}\eta)$, one has

$$\begin{aligned} &\langle \boldsymbol{r}|G_0(E + \mathrm{i}\eta)|\boldsymbol{r}'\rangle \\ &= \int \mathrm{d}\boldsymbol{p}\, \langle \boldsymbol{r}|G_0(E + \mathrm{i}\eta)|\boldsymbol{p}\rangle\langle \boldsymbol{p}|\boldsymbol{r}'\rangle \\ &= \int \mathrm{d}\boldsymbol{p}\, \langle \boldsymbol{r}| \{(E + \mathrm{i}\eta)I - H_0\}^{-1} |\boldsymbol{p}\rangle\langle \boldsymbol{p}|\boldsymbol{r}'\rangle \\ &= \frac{1}{(2\pi\hbar)^3} \int \mathrm{d}\boldsymbol{p} \left[E + \mathrm{i}\eta - \frac{p^2}{2m}\right]^{-1} \exp\left\{\frac{\mathrm{i}}{\hbar}\boldsymbol{p} \cdot (\boldsymbol{r} - \boldsymbol{r}')\right\} \\ &= -\frac{m}{4\pi^3\hbar^2} \int \mathrm{d}\boldsymbol{q}\, \frac{1}{q^2 - k^2 - \mathrm{i}\eta} \exp\{\mathrm{i}\boldsymbol{q} \cdot (\boldsymbol{r} - \boldsymbol{r}')\}. \end{aligned} \tag{11.25}$$

Here we have let $\boldsymbol{q} = \hbar^{-1}\boldsymbol{p}$, and k is related to E through (11.6).[5] Let us introduce the spherical coordinates (q, θ_q, ϕ_q) of the vector $\boldsymbol{q}$ with respect to the direction of $\boldsymbol{r} - \boldsymbol{r}'$. Performing the integration on $\mathrm{d}\theta_q$ and $\mathrm{d}\phi_q$ one finds that

$$\begin{aligned} &\langle \boldsymbol{r}|G_0(E + \mathrm{i}\eta)|\boldsymbol{r}'\rangle \\ &= -\frac{m}{4\pi^3\hbar^2} \int_0^\infty q^2 \mathrm{d}q \int \mathrm{d}\Omega_q\, \frac{1}{q^2 - k^2 - \mathrm{i}\eta} \exp\{-\mathrm{i}q|\boldsymbol{r} - \boldsymbol{r}'| \cos\theta_q\} \\ &= \frac{m}{2\pi^2\hbar^2} \frac{1}{\mathrm{i}|\boldsymbol{r} - \boldsymbol{r}'|} \int_0^\infty \mathrm{d}q\, \frac{q}{q^2 - k^2 - \mathrm{i}\eta} \left\{\mathrm{e}^{-\mathrm{i}q|\boldsymbol{r}-\boldsymbol{r}'|} - \mathrm{e}^{\mathrm{i}q|\boldsymbol{r}-\boldsymbol{r}'|}\right\} \\ &= \frac{m}{2\pi^2\hbar^2} \frac{1}{\mathrm{i}|\boldsymbol{r} - \boldsymbol{r}'|} \int_{-\infty}^\infty \mathrm{d}q\, \frac{q}{q^2 - k^2 - \mathrm{i}\eta} \mathrm{e}^{-\mathrm{i}q|\boldsymbol{r}-\boldsymbol{r}'|}. \end{aligned}$$

---

[5] If C is a positive and finite constant, we make no distinction between η and $C\eta$, since eventually the limit $\eta \to 0$ is taken.

To compute this integral it is convenient to consider q as a complex variable, shown in Fig. 11.2. It is easy to show that adding to the integration contour an infinite semicircle in the lower half-plane produces no additional contribution to the integral. Within the closed contour the only singularity is a simple pole at $q = -k - \mathrm{i}\eta$. Applying the residue theorem then yields[6]

$$\langle \boldsymbol{r}|G_0(E+\mathrm{i}\eta|\boldsymbol{r}'\rangle = \frac{m}{2\pi^2\hbar^2}\frac{(-2\pi\mathrm{i})}{\mathrm{i}|\boldsymbol{r}-\boldsymbol{r}'|}\lim_{q\to-k-\mathrm{i}\eta}\left\{\frac{q}{q-k-\mathrm{i}\eta}\mathrm{e}^{-\mathrm{i}q|\boldsymbol{r}-\boldsymbol{r}'|}\right\}.$$

Taking the limit shown and letting η go to zero, one finally obtains

$$\langle \boldsymbol{r}|G_0(E+\mathrm{i}0)|\boldsymbol{r}'\rangle = -\frac{1}{4\pi}\frac{2m}{\hbar^2}\frac{\mathrm{e}^{\mathrm{i}k|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{11.26}$$

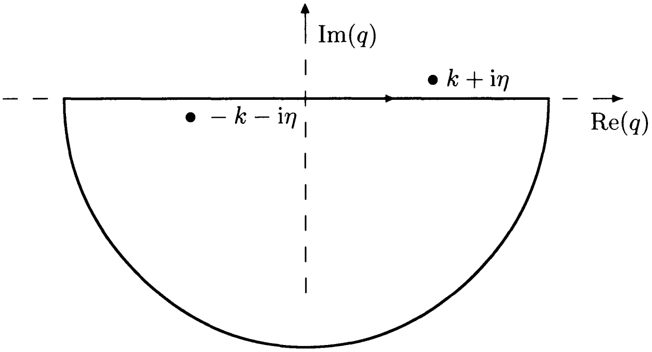

Fig. 11.2. Integration contour in the complex plane

Similarly, for $G_0(E - \mathrm{i}\eta)$ one gets (Exercise 11.2)

$$\langle \boldsymbol{r}|G_0(E-\mathrm{i}0)|\boldsymbol{r}'\rangle = -\frac{1}{4\pi}\frac{2m}{\hbar^2}\frac{\mathrm{e}^{-\mathrm{i}k|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}. \tag{11.27}$$

Matrix elements of Green's operators in the coordinate representation are called *Green's functions.* The result of Exercise (11.3) implies that functions (11.26) and (11.27) satisfy the differential equation

$$\frac{\hbar^2}{2m}\left\{\nabla^2 + k^2\right\}\langle \boldsymbol{r}|G_0(E\pm\mathrm{i}0)|\boldsymbol{r}'\rangle = \delta(\boldsymbol{r}-\boldsymbol{r}'). \tag{11.28}$$

This equation is often used to define G_0.

---

[6] For the residue theorem and basic notions of complex variables see, for instance, [65], [180] and [236].

11.4 Scattering Integral Equation

Equation (11.23) implicitly determines $|\psi_\pm\rangle$ in terms of $|\phi\rangle$:

$$\{I - G_0(E \pm \mathrm{i}0)V\}\,|\psi_\pm\rangle = |\phi\rangle\,.$$

In the coordinate representation this becomes

$$\langle \boldsymbol{r}|\psi_\pm\rangle - \int \mathrm{d}\boldsymbol{r}'\,\langle \boldsymbol{r}|G_0(E\pm\mathrm{i}0)|\boldsymbol{r}'\rangle\langle \boldsymbol{r}'|V|\psi_\pm\rangle = \langle \boldsymbol{r}|\phi\rangle. \tag{11.29}$$

We know that if $|\phi\rangle$ is an eigenvector of H_0 with eigenvalue E, then $|\psi_+\rangle$ and $|\psi_-\rangle$ are eigenvectors of H with the same eigenvalue. Let $\boldsymbol{k}$ be a vector obeying (11.6) and set

$$\phi(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\,. \tag{11.30}$$

Then $\psi_\pm(\boldsymbol{r})$, which is a solution of the integral equation

$$\psi_\pm(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} - \frac{1}{4\pi}\int \mathrm{d}\boldsymbol{r}'\,\frac{\mathrm{e}^{\pm\mathrm{i}k|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\,\frac{2m}{\hbar^2}V(\boldsymbol{r}')\psi_\pm(\boldsymbol{r}'), \tag{11.31}$$

is an eigenfunction of H with eigenvalue E. Equation (11.31) is called the *scattering integral equation.*

We now show that if r is large, the wave function $\psi_+(\boldsymbol{r})$ has the form (11.5). Because $V(\boldsymbol{r}')$ rapidly decreases for $r' \to \infty$, the integral on $\mathrm{d}\boldsymbol{r}'$ gets a significant contribution only where $r' \ll r$. In that case one can write

$$|\boldsymbol{r}-\boldsymbol{r}'| = r - \hat{r}\cdot\boldsymbol{r}' + O\left[(r')^2/r\right],$$

so that

$$\frac{\mathrm{e}^{\mathrm{i}k|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|} \approx \frac{1}{r}\mathrm{e}^{\mathrm{i}kr}\mathrm{e}^{-\mathrm{i}k\hat{r}\cdot\boldsymbol{r}'}\,. \tag{11.32}$$

Substituting (11.32) in (11.31), one gets

$$\begin{aligned}\psi_+(\boldsymbol{r}) &= \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} - \frac{1}{4\pi}\frac{\mathrm{e}^{\mathrm{i}kr}}{r}\int \mathrm{d}\boldsymbol{r}'\,\mathrm{e}^{-\mathrm{i}k\hat{r}\cdot\boldsymbol{r}'}\frac{2m}{\hbar^2}V(\boldsymbol{r}')\psi_+(\boldsymbol{r}')\\ &= \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} + \frac{f(\theta,\phi)}{r}\mathrm{e}^{\mathrm{i}kr}\,,\end{aligned} \tag{11.33}$$

where

$$f(\theta,\phi) = -\frac{1}{4\pi}\int \mathrm{d}\boldsymbol{r}'\,\mathrm{e}^{-\mathrm{i}k\hat{r}\cdot\boldsymbol{r}'}\frac{2m}{\hbar^2}V(\boldsymbol{r}')\psi_+(\boldsymbol{r}'). \tag{11.34}$$

As the notation suggests, the scattering amplitude depends on $\boldsymbol{r}$ only through the angular variables $\hat{r}$. Variables θ and ϕ can be chosen as the polar and azimuthal angles of $\boldsymbol{r}$ with respect to $\boldsymbol{k}$.

It is not difficult to show that for large r the solution $\psi_-(\boldsymbol{r})$ is given by

$$\psi_-(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} - \frac{1}{4\pi}\frac{\mathrm{e}^{-\mathrm{i}kr}}{r}\int \mathrm{d}\boldsymbol{r}'\,\mathrm{e}^{\mathrm{i}k\hat{r}\cdot\boldsymbol{r}'}\frac{2m}{\hbar^2}V(\boldsymbol{r}')\psi_-(\boldsymbol{r}'). \tag{11.35}$$

This wave function represents the superposition of a plane wave in the $\hat{k}$ direction and an incoming wave centered on $r = 0$. It is clear that for all practical purposes, such a situation cannot be realized physically. It would require, everywhere on the surface of a large sphere, to control the phase of the incoming wave with a precision of the order of atomic dimensions. For this reason the solution $\psi_-(\boldsymbol{r})$ will not be considered further. The solution $\psi_+(\boldsymbol{r})$ will henceforth be denoted by $\psi(\boldsymbol{r})$.

In Sect. 11.2 we saw that the differential scattering cross section is equal to the absolute square of the scattering amplitude. The latter has just been obtained, but in terms of the wave function $\psi(\boldsymbol{r})$, itself unknown. In principle $\psi(\boldsymbol{r})$ can be calculated from the scattering integral equation. Under certain hypotheses, one can also evaluate $\psi(\boldsymbol{r})$ by means of an infinite series. To do this, one just writes (11.24) in the coordinate representation. One thus finds

$$\begin{aligned}\psi(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} &- \frac{1}{4\pi}\int \mathrm{d}\boldsymbol{r}'\,\frac{\mathrm{e}^{\mathrm{i}k|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\frac{2m}{\hbar^2}V(\boldsymbol{r}')\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}'} \\ &+ \left(\frac{-1}{4\pi}\right)^2\int \mathrm{d}\boldsymbol{r}'\,\mathrm{d}\boldsymbol{r}''\,\frac{\mathrm{e}^{\mathrm{i}k|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\frac{2m}{\hbar^2}V(\boldsymbol{r}')\frac{\mathrm{e}^{\mathrm{i}k|\boldsymbol{r}'-\boldsymbol{r}''|}}{|\boldsymbol{r}'-\boldsymbol{r}''|}\frac{2m}{\hbar^2}V(\boldsymbol{r}'')\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}''} \\ &+ \cdots \end{aligned} \tag{11.36}$$

This series solution can also be obtained recursively from the scattering integral equation.

In general the series just found does not converge for all potentials, even if they satisfy the hypotheses stated in Sect. 11.2. One can show, however, that the series converges if the energy is high enough. Moreover, if $V(\boldsymbol{r})$ is a potential for which the series does not converge, then there exists a $\kappa_0 > 0$ such that if $|\kappa| < \kappa_0$, the series converges for the potential $\kappa V(\boldsymbol{r})$ ([169], Sects. 9.1 and 10.3; [224], Sect. 9a).

11.5 The Born Approximation

When the series (11.36) converges, it happens that it does so rapidly. The *Born approximation* consists in replacing, in the expression (11.34) for $f(\theta,\phi)$, the wave function $\psi(\boldsymbol{r}')$ by $\exp\{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}'\}$. Thus only the first term is retained in the series expansion of $\psi(\boldsymbol{r}')$. In this way we get

$$\begin{aligned} f(\theta,\phi) &= -\frac{1}{4\pi}\int \mathrm{d}\boldsymbol{r}'\,\mathrm{e}^{-\mathrm{i}k\hat{r}\cdot\boldsymbol{r}'}\frac{2m}{\hbar^2}V(\boldsymbol{r}')\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}'} \\ &= -\frac{1}{4\pi}\int \mathrm{d}\boldsymbol{r}'\,\mathrm{e}^{-\mathrm{i}\Delta\boldsymbol{k}\cdot\boldsymbol{r}'}\frac{2m}{\hbar^2}V(\boldsymbol{r}')\,. \end{aligned} \tag{11.37}$$

Here $\Delta \boldsymbol{k} = k\hat{\boldsymbol{r}} - \boldsymbol{k}$ is the difference between wave vectors of the scattered and of the incoming wave. The integral (11.37) is familiar: in the Born approximation, the scattering amplitude is essentially the Fourier transform of the potential.

Let us investigate the conditions under which the Born approximation is valid. The approximation was obtained by replacing in the expression for the scattering amplitude the wave function $\psi(\boldsymbol{r})$ by $\exp\{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}\}$. From the scattering integral equation we see that

$$\left|\psi(\boldsymbol{r}) - \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\right| = \frac{1}{4\pi}\left|\int \mathrm{d}\boldsymbol{r}' \, \frac{\mathrm{e}^{\mathrm{i}k|\boldsymbol{r}-\boldsymbol{r}'|}}{|\boldsymbol{r}-\boldsymbol{r}'|}\frac{2m}{\hbar^2}V(\boldsymbol{r}')\psi(\boldsymbol{r}')\right| .$$

For the Born approximation to be valid, the substitution of $\exp\{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}\}$ to $\psi(\boldsymbol{r})$ must be legitimate around $\boldsymbol{r} = 0$. Indeed most contributions to the integral in the expression for $f(\theta,\phi)$ come from that region. Hence it is likely that the Born approximation will be valid provided that

$$\left|\frac{1}{4\pi}\int \mathrm{d}\boldsymbol{r}' \, \frac{\mathrm{e}^{\mathrm{i}(kr'+\boldsymbol{k}\cdot\boldsymbol{r}')}}{r'}\frac{2m}{\hbar^2}V(\boldsymbol{r}')\right| \ll 1 . \tag{11.38}$$

This inequality is satisfied if the potential is weak or k is large. These conditions, of course, are similar to the ones that determine the convergence of the series representation of $\psi(\boldsymbol{r})$.

As an application we evaluate the differential scattering cross section of charged particles by neutral atoms. Consider an atom with atomic number Z. Let $\varrho(\boldsymbol{r})$ be the electronic charge density computed, for instance, in the central-field model. We assume a pointlike nucleus, with charge density equal to $-Zq_\mathrm{e}\delta(\boldsymbol{r})$. The potential energy of a charge q in the field of the atom at point $\boldsymbol{r}'$ is given by

$$V(\boldsymbol{r}') = \frac{q}{4\pi\varepsilon_0}\int \mathrm{d}\boldsymbol{r}'' \, \frac{1}{|\boldsymbol{r}'-\boldsymbol{r}''|}\left\{-Zq_\mathrm{e}\delta(\boldsymbol{r}'') + \varrho(\boldsymbol{r}'')\right\} . \tag{11.39}$$

In the Born approximation the scattering amplitude is therefore equal to

$$\begin{aligned} f(\theta,\phi) &= -\frac{1}{4\pi}\int \mathrm{d}\boldsymbol{r}' \, \mathrm{e}^{-\mathrm{i}\Delta\boldsymbol{k}\cdot\boldsymbol{r}'}\frac{2m}{\hbar^2}V(\boldsymbol{r}') \\ &= -\frac{mq}{8\pi^2\varepsilon_0\hbar^2}\int \mathrm{d}\boldsymbol{r}' \, \mathrm{e}^{-\mathrm{i}\Delta\boldsymbol{k}\cdot\boldsymbol{r}'}\int \mathrm{d}\boldsymbol{r}'' \, \frac{1}{|\boldsymbol{r}'-\boldsymbol{r}''|}\left\{-Zq_\mathrm{e}\delta(\boldsymbol{r}'') + \varrho(\boldsymbol{r}'')\right\} . \end{aligned} \tag{11.40}$$

It would be convenient here to change the order of integration, and first perform the integral over $\mathrm{d}\boldsymbol{r}'$. Unfortunately an indefinite result would follow, the inversion not being allowed in integral (11.40). One can nonetheless bypass the problem by introducing a *convergence factor*. Let $\alpha > 0$ and write

$$f(\theta,\phi) = -\frac{mq}{8\pi^2\varepsilon_0\hbar^2}\lim_{\alpha\to 0}\int d\boldsymbol{r}'\,e^{-i\Delta\boldsymbol{k}\cdot\boldsymbol{r}'} \times \int d\boldsymbol{r}''\,\frac{e^{-\alpha|\boldsymbol{r}'-\boldsymbol{r}''|}}{|\boldsymbol{r}'-\boldsymbol{r}''|}\left\{-Zq_e\delta(\boldsymbol{r}'')+\varrho(\boldsymbol{r}'')\right\}.$$

The introduction of the factor $\exp\{-\alpha|\boldsymbol{r}'-\boldsymbol{r}''|\}$ does not change the value of the integral when the limit $\alpha\to 0$ is taken. This comes from the fact that it differs from 1 only in regions that essentially do not contribute to the integral. But now the change of the order of integration is allowed and one can write

$$\begin{aligned} f(\theta,\phi) &= -\frac{mq}{8\pi^2\varepsilon_0\hbar^2}\int d\boldsymbol{r}''\left\{-Zq_e\delta(\boldsymbol{r}'')+\varrho(\boldsymbol{r}'')\right\} \\ &\quad\times\lim_{\alpha\to 0}\int d\boldsymbol{r}'\,e^{-i\Delta\boldsymbol{k}\cdot\boldsymbol{r}'}\frac{e^{-\alpha|\boldsymbol{r}'-\boldsymbol{r}''|}}{|\boldsymbol{r}'-\boldsymbol{r}''|} \\ &= -\frac{mq}{8\pi^2\varepsilon_0\hbar^2}\int d\boldsymbol{r}''\left\{-Zq_e\delta(\boldsymbol{r}'')+\varrho(\boldsymbol{r}'')\right\}e^{-i\Delta\boldsymbol{k}\cdot\boldsymbol{r}''} \\ &\quad\times\lim_{\alpha\to 0}\int d\boldsymbol{s}\,\frac{1}{s}\exp\left\{-\alpha s - i\Delta\boldsymbol{k}\cdot\boldsymbol{s}\right\}. \end{aligned}$$

In the last equality the change of variable $\boldsymbol{r}'=\boldsymbol{s}+\boldsymbol{r}''$ was made. The integral on $d\boldsymbol{s}$ can be evaluated and one gets (Exercise 11.4)

$$\lim_{\alpha\to 0}\int\frac{d\boldsymbol{s}}{s}\exp\left\{-\alpha s - i\Delta\boldsymbol{k}\cdot\boldsymbol{s}\right\} = \lim_{\alpha\to 0}\frac{4\pi}{\alpha^2+(\Delta\boldsymbol{k})^2} = \frac{4\pi}{(\Delta\boldsymbol{k})^2}.$$

Hence

$$f(\theta,\phi) = -\frac{mq}{2\pi\varepsilon_0\hbar^2(\Delta\boldsymbol{k})^2}\int d\boldsymbol{r}''\,e^{-i\Delta\boldsymbol{k}\cdot\boldsymbol{r}''}\left\{-Zq_e\delta(\boldsymbol{r}'')+\varrho(\boldsymbol{r}'')\right\}. \tag{11.41}$$

This is the scattering amplitude of charged particles by atoms in the Born approximation.

To simplify formula (11.41) further, assume, as in the central-field model, that the electronic charge density $\varrho(\boldsymbol{r}'')$ depends only on r''. In this case one can write

$$\int d\boldsymbol{r}''\,e^{-i\Delta\boldsymbol{k}\cdot\boldsymbol{r}''}\left\{-Zq_e\delta(\boldsymbol{r}'')+\varrho(\boldsymbol{r}'')\right\} = -Zq_e+q_eF(\theta), \tag{11.42}$$

where $F(\theta)$, called the *atomic form factor*, is a dimensionless quantity equal to

$$F(\theta) = \frac{1}{q_e}\frac{4\pi}{|\Delta\boldsymbol{k}|}\int_0^\infty dr''\,r''\varrho(r'')\sin\left\{|\Delta\boldsymbol{k}|r''\right\}. \tag{11.43}$$

As the notation suggests $F(\theta)$ depends, through $|\Delta\boldsymbol{k}|$, only on the polar angle θ (and not on ϕ). Indeed

$$|\Delta \boldsymbol{k}| = |k\hat{r} - \boldsymbol{k}| = 2k \left|\sin\frac{\theta}{2}\right| . \tag{11.44}$$

Summarizing, one sees that the scattering amplitude due to a spherically symmetric electronic charge density is given by

$$f(\theta) = \frac{mq_e q}{8\pi\varepsilon_0 \hbar^2 k^2 \sin^2(\theta/2)} \{Z - F(\theta)\} . \tag{11.45}$$

The scattering amplitude only depends on θ. The differential cross section follows as

$$\sigma(\theta) = |f(\theta)|^2 = \left\{ \frac{mq_e q}{8\pi\varepsilon_0 \hbar^2 k^2 \sin^2(\theta/2)} [Z - F(\theta)] \right\}^2 . \tag{11.46}$$

The Born approximation, where valid, yields the differential scattering cross section in terms of the potential. But the investigator is more often interested in the inverse problem. He really wants to determine what potential produces the differential cross section he has measured. To see if this is possible, let us rewrite formula (11.37) for the scattering amplitude as

$$f_k(\theta, \phi) = f(\Delta \boldsymbol{k}) = -\frac{1}{4\pi} \int \mathrm{d}\boldsymbol{r}' \, \mathrm{e}^{-\mathrm{i}\Delta \boldsymbol{k}\cdot \boldsymbol{r}'} \frac{2m}{\hbar^2} V(\boldsymbol{r}') . \tag{11.47}$$

Here the k dependence of the scattering amplitude is shown explicitly. Let us write the inverse Fourier transform of (11.47). We find

$$V(\boldsymbol{r}) = -\frac{\hbar^2}{4\pi^2 m} \int \mathrm{d}(\Delta \boldsymbol{k}) \, \mathrm{e}^{\mathrm{i}\Delta \boldsymbol{k}\cdot \boldsymbol{r}} f(\Delta \boldsymbol{k}) . \tag{11.48}$$

To obtain $V(\boldsymbol{r})$, one should know $f_k(\theta, \phi)$ for all values of k (i.e. of E), θ and ϕ. But the differential cross section, which is the experimentally measurable quantity, is equal to the absolute square of the scattering amplitude. This means that f can be extracted from σ only if f is real. This happens if $V(-\boldsymbol{r}) = V(\boldsymbol{r})$ and, in particular, if the potential is spherically symmetric. Note that if $f(\Delta \boldsymbol{k})$ is poorly known at large energies (i.e. at large values of $|\Delta \boldsymbol{k}|$), one cannot determine finer details of $V(\boldsymbol{r})$. That remark is highly general. To determine finer details of a scattering center one must use particles with small wavelengths, that is, large energies.

11.6 Partial Waves and Phase Shifts

The formalism of the scattering integral equation, like the Born approximation, applies in particular to a spherically symmetric potential. Scattering by such a potential, however, can also be treated differently, directly from the differential equation. This approach is particularly useful when the Born approximation is not applicable. Furthermore, it is suitable to the use of the numerical methods developed in Chap. 8.

In general the scattering amplitude depends on θ and ϕ. For a spherically symmetric potential, however, the differential cross section should not depend on ϕ. Indeed from a physical point of view, nothing then should distinguish two setups obtained one from the other by a rotation about an axis parallel with the incoming beam and going through the origin. With no loss of generality the axis can be taken to coincide with the $\hat{z}$ direction.

Equation (11.4) can be written as

$$-\nabla^2\psi(\boldsymbol{r})+\frac{2m}{\hbar^2}V(r)\psi(\boldsymbol{r}) = k^2\psi(\boldsymbol{r}). \tag{11.49}$$

For large r, we look for solutions of the form

$$\psi(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}kz} + \frac{1}{r}f(\theta)\mathrm{e}^{\mathrm{i}kr}. \tag{11.50}$$

For this purpose we first write down the asymptotic form, for large r, of the most general solution of (11.49) that does not depend on ϕ. We will then compare it with the asymptotic form of (11.50).

In Sect. 7.4 we saw that for a spherically symmetric potential, the Hamiltonian's eigenvalue equation has solutions which are products of radial functions and spherical harmonics. Only the spherical harmonics $Y_{l0}(\theta,\phi)$ do not depend on ϕ. They are proportional to Legendre polynomials. For a given value of k, the most general solution of the eigenvalue equation that does not depend on ϕ can therefore be written as

$$\psi(\boldsymbol{r}) = \sum_{l=0}^{\infty} c_l(2l+1)\mathrm{i}^l\frac{\bar{u}_l(r)}{r}P_l(\cos\theta). \tag{11.51}$$

Here the c_l are constants and the factor $(2l+1)\mathrm{i}^l$ is introduced for later simplification. Substituting (11.51) in (11.49), one easily finds that $\bar{u}_l(r)$ must satisfy

$$\frac{\mathrm{d}^2\bar{u}_l}{\mathrm{d}r^2}+\left\{k^2-\frac{2m}{\hbar^2}V(r)-\frac{l(l+1)}{r^2}\right\}\bar{u}_l = 0. \tag{11.52}$$

This equation coincides with (8.38), except that variables with dimensions have been kept.

Insofar as the potential decreases fast enough where $r \to \infty$, the asymptotic behavior of $\bar{u}_l(r)$ is given by (8.50). Picking an appropriate positive value of the normalization factor one can write

$$\bar{u}_l(r) \to \frac{1}{k}\sin\left\{kr-\frac{\pi l}{2}+\delta_l(k)\right\}, \qquad r\to\infty. \tag{11.53}$$

The $\delta_l(k)$ are called *phase shifts*. Let us substitute (11.53) in (11.51). Writing the sine in terms of exponentials one finds that the asymptotic behavior of $\psi(\boldsymbol{r})$ is given by

$$\psi(\boldsymbol{r}) \to \left\{ \sum_{l=0}^{\infty} c_l (2l+1) \mathrm{i}^l \frac{1}{2\mathrm{i}} P_l(\cos\theta) \mathrm{e}^{-\mathrm{i}\pi l/2} \mathrm{e}^{\mathrm{i}\delta_l} \right\} \frac{\mathrm{e}^{\mathrm{i}kr}}{kr} + \left\{ \sum_{l=0}^{\infty} c_l (2l+1) \mathrm{i}^l \frac{(-1)}{2\mathrm{i}} P_l(\cos\theta) \mathrm{e}^{\mathrm{i}\pi l/2} \mathrm{e}^{-\mathrm{i}\delta_l} \right\} \frac{\mathrm{e}^{-\mathrm{i}kr}}{kr} . \tag{11.54}$$

To obtain the asymptotic form of (11.50) for large r, we must first expand the exponential of $\{\mathrm{i}kz\}$ by means of formula (7.201):

$$\mathrm{e}^{\mathrm{i}kz} = \mathrm{e}^{\mathrm{i}kr\cos\theta} = \sum_{l=0}^{\infty} (2l+1) \mathrm{i}^l j_l(kr) P_l(\cos\theta) .$$

Using the leading term in the asymptotic expansion (7.199) of the spherical Bessel function, one finds that the asymptotic form of (11.50) is given by

$$\begin{aligned} \psi(\boldsymbol{r}) &\to \sum_{l=0}^{\infty} (2l+1) \mathrm{i}^l \frac{1}{kr} \sin\left\{ kr - \frac{\pi l}{2} \right\} P_l(\cos\theta) + \frac{1}{r} f(\theta) \mathrm{e}^{\mathrm{i}kr} \\ &\to \left\{ \sum_{l=0}^{\infty} (2l+1) \mathrm{i}^l \frac{1}{2\mathrm{i}} P_l(\cos\theta) \mathrm{e}^{-\mathrm{i}\pi l/2} + k f(\theta) \right\} \frac{\mathrm{e}^{\mathrm{i}kr}}{kr} \\ &\quad + \left\{ \sum_{l=0}^{\infty} (2l+1) \mathrm{i}^l \frac{(-1)}{2\mathrm{i}} P_l(\cos\theta) \mathrm{e}^{\mathrm{i}\pi l/2} \right\} \frac{\mathrm{e}^{-\mathrm{i}kr}}{kr} . \end{aligned} \tag{11.55}$$

Equation (11.51) is the most general solution of the eigenvalue equation that does not depend on ϕ. For (11.50) to be a solution for large values of r, the two asymptotic forms must coincide. They will coincide if and only if

$$\sum_{l=0}^{\infty} c_l (2l+1) \frac{1}{2\mathrm{i}} \mathrm{e}^{\mathrm{i}\delta_l} P_l(\cos\theta) = \sum_{l=0}^{\infty} (2l+1) \frac{1}{2\mathrm{i}} P_l(\cos\theta) + k f(\theta) , \tag{11.56}$$

$$c_l \mathrm{e}^{-\mathrm{i}\delta_l} = 1 . \tag{11.57}$$

The scattering amplitude can therefore be written as

$$f(\theta) = \frac{1}{2\mathrm{i}k} \sum_{l=0}^{\infty} (2l+1) \left\{ \mathrm{e}^{2\mathrm{i}\delta_l} - 1 \right\} P_l(\cos\theta) . \tag{11.58}$$

The scattering amplitude is completely determined by the phase shifts. Recall that these vanish for a free particle. The phase shifts can seldom be obtained exactly. One can, however, compute them by the numerical methods developed in Sect. 8.5.[7]

The expression of the scattering amplitude in terms of phase shifts is especially useful when few phase shifts only are significantly different from zero. We pointed out that for a spherically symmetric potential the Born approximation yields a real scattering amplitude. Hence the Born approximation can only be valid if the phase shifts are very small or, perhaps, close to integral multiples of π.

The differential cross section is equal to $|f(\theta)|^2$. To obtain the total cross section one just integrates it on all solid angle. Using (7.170) as well as orthogonality properties of spherical harmonics one finds that

$$\begin{aligned}\sigma &= \int \mathrm{d}\Omega\, |f(\theta)|^2 \\ &= \frac{1}{4k^2} \sum_{l,l'=0}^{\infty} (2l+1)^{1/2}(2l'+1)^{1/2} \\ &\quad \times 4\pi \left\{\mathrm{e}^{-2\mathrm{i}\delta_l} - 1\right\} \left\{\mathrm{e}^{2\mathrm{i}\delta_{l'}} - 1\right\} \int \mathrm{d}\Omega\, Y_{l0}^*(\theta,\phi) Y_{l'0}(\theta,\phi) \\ &= \frac{4\pi}{4k^2} \sum_{l=0}^{\infty} (2l+1) \left\{\mathrm{e}^{-2\mathrm{i}\delta_l} - 1\right\} \left\{\mathrm{e}^{2\mathrm{i}\delta_l} - 1\right\} \\ &= \frac{4\pi}{k^2} \sum_{l=0}^{\infty} (2l+1) \sin^2 \delta_l \,. \end{aligned} \tag{11.59}$$

This is the expression of the total cross section in terms of the phase shifts.

The cross section σ is related to the total number of particles scattered per unit time. Obviously these particles are drawn form the incoming beam. The number of particles drawn from the incoming beam per unit time must somehow be related to the scattering amplitude at $\theta = 0$. Thus one can expect a relation between the total cross section and the scattering amplitude at $\theta = 0$. To establish this relation we note that

$$f(\theta = 0) = \frac{1}{2\mathrm{i}k} \sum_{l=0}^{\infty} (2l+1) \left\{\mathrm{e}^{2\mathrm{i}\delta_l} - 1\right\} P_l(1).$$

Making use of (7.156), one has

$$f(\theta = 0) = \frac{1}{2\mathrm{i}k} \sum_{l=0}^{\infty} (2l+1) \left\{\cos(2\delta_l) - 1 + \mathrm{i} \sin(2\delta_l)\right\}. \tag{11.60}$$

---

[7] To obtain the scattering amplitude it is enough to know the phase shifts modulo π. The numerical methods of Sect. 8.5 allow to compute them modulo 2π. In principle the phase shifts can be determined without ambiguity. For this purpose one replaces the potential $V(r)$ by $\kappa V(r)$, one computes the phase shifts as a function of κ and one lets κ vary from 0 to 1 continuously.

Comparing this equation with the expression for σ, one finds that

$$\sigma = \frac{4\pi}{k}\mathrm{Im}\left\{f(\theta = 0)\right\}. \tag{11.61}$$

This relation between the total cross section and the imaginary part of the scattering amplitude is very general, and is called the *optical theorem.*

It is interesting to rewrite the asymptotic form of (11.51) differently. Using (11.54) and (11.57) and letting $S_l = \exp\{2\mathrm{i}\delta_l\}$, one gets

$$\begin{aligned}\psi(\boldsymbol{r}) = \sum_{l=0}^{\infty}\psi_l(\boldsymbol{r}) \to \sum_{l=0}^{\infty}\frac{(2l+1)\mathrm{i}^{l+1}P_l(\cos\theta)}{2k}&\left\{\frac{1}{r}\exp\left[-\mathrm{i}\left(kr - \frac{\pi l}{2}\right)\right]\right.\\ &\left. - S_l\frac{1}{r}\exp\left[\mathrm{i}\left(kr - \frac{\pi l}{2}\right)\right]\right\}.\end{aligned} \tag{11.62}$$

Each term $\psi_l(\boldsymbol{r})$ is called a *partial wave.* Asymptotically the partial wave $\psi_l(\boldsymbol{r})$ decomposes into an (exponentially negative) converging wave and an (exponentially positive) diverging wave. Thus the scattering process can be interpreted as a conversion of converging into diverging waves, phase shifted by $2\delta_l$.[8] The phase shifts δ_l depend on l and on k (or, equivalently, on l and on E). Thus one can write

$$S_l(k) = \exp\left\{2\mathrm{i}\delta_l(k)\right\}. \tag{11.63}$$

The complex numbers $S_l(k)$ indicate the relative amplitude of diverging and converging partial waves. We will examine some of their properties in Sect. 11.8. We should point out here that the $S_l(k)$ are elements of a matrix called the *scattering matrix*, whose importance is central in space-time approaches to scattering theory.

11.7 Hard Sphere and Spherical Square Well

In general phase shifts must be calculated by numerical methods. There are, however, situations where they can be obtained exactly. Two examples will be given in this section. In both cases the potential $V(r)$ vanishes for $r > a$, where a is a specific length.

In a region where $V(r)$ vanishes the most general solution of (11.52) is given by

$$\bar{u}_l(r) = r\left\{A_l\, j_l(kr) + B_l\, n_l(kr)\right\}. \tag{11.64}$$

Here j_l and n_l are spherical Bessel and Neumann functions, while A_l and B_l are constants. The asymptotic form of this expression for large r is obtained from (7.199) and (7.200):

---

[8] Note that this interpretation is possible only if the plane wave $\exp\{ikz\}$ extends to infinity in the xy plane, that is, if the factor $F(u, v)$ of Sect. 11.2 is absent.

$$\bar{u}_l(r) \to A_l \frac{1}{k} \sin\left\{kr - \frac{\pi l}{2}\right\} + B_l \frac{1}{k} \cos\left\{kr - \frac{\pi l}{2}\right\}, \qquad r \to \infty. \tag{11.65}$$

This coincides (up to a factor) with the asymptotic form of $\bar{u}_l(r)$, given in (11.53) in terms of phase shifts, provided that $(0 \le \arctan < \pi)$

$$\delta_l = \begin{cases} \arctan\left(\frac{B_l}{A_l}\right) & \text{if } B_l > 0\,, \\ \arctan\left(\frac{B_l}{A_l}\right) + \pi & \text{if } B_l < 0\,. \end{cases} \tag{11.66}$$

i) Hard Sphere

A hard sphere of radius a corresponds to the following potential:[9]

$$V(r) = \begin{cases} \infty & \text{if } r < a\,, \\ 0 & \text{if } r > a\,. \end{cases} \tag{11.67}$$

Due to the potential's infinite discontinuity the wave function and, therefore, $\bar{u}_l$ must vanish at $r = a$. This means that

$$0 = A_l\, j_l(ka) + B_l\, n_l(ka)\,,$$

from which we get

$$\frac{B_l}{A_l} = -\frac{j_l(ka)}{n_l(ka)}\,. \tag{11.68}$$

Thus phase shifts are obtained by a very simple formula. From them one can get the exact form of the differential and total cross sections.

The cross section has a particularly simple form in the case of low energy, which means that $ka \ll 1$. To obtain it we use the approximations (7.197) and (7.198) for spherical Bessel functions with small arguments. We have

$$\tan\delta_l = \frac{B_l}{A_l} \approx \frac{-(ka)^{2l+1}}{(2l+1)\left[(2l-1)!!\right]^2}\,, \tag{11.69}$$

from which we get

$$\sigma(\theta) = |f(\theta)|^2 = a^2\left\{1 + O\left[(ka)^2\right]\right\}, \tag{11.70}$$

$$\sigma = 4\pi a^2\left\{1 + O\left[(ka)^2\right]\right\}. \tag{11.71}$$

Hence the low-energy differential cross section is approximately constant and the total cross section is approximately equal to four times the classical value.

---

[9] In Sect. 11.2 we assumed the potential has finite discontinuities only. The results we will obtain are nonetheless valid, as can be verified by an appropriate limiting process. Note, however, that phase shifts are obtained modulo π only.

ii) Spherical Square Well

The spherical square well of radius a and depth V_0 is defined in (7.114) and shown in Fig. 7.2 (p. 143). In the region where $r < a$, (11.52) can be written as

$$\frac{\mathrm{d}^2\bar{u}_l}{\mathrm{d}r^2} + \left\{ (k')^2 - \frac{l(l+1)}{r^2} \right\} \bar{u}_l = 0, \tag{11.72}$$

where $(k')^2 = k^2 + 2mV_0/\hbar^2$. The general solution of this equation is a linear combination of $rj_l(k'r)$ and $rn_l(k'r)$. The function $rn_l(k'r)$, however, must be rejected since it is singular at the origin. Thus one must have

$$\bar{u}_l(r) = \begin{cases} A_l' \, rj_l(k'r) & \text{if } r < a \, , \\ A_l \, rj_l(kr) + B_l \, rn_l(kr) & \text{if } r > a \, . \end{cases} \tag{11.73}$$

The radial function and its derivative should be continuous at $r = a$. This entails that (Exercise 11.8)

$$A_l = A_l' \left\{ (ka)(k'a) n_l(ka) j_l'(k'a) - (ka)^2 n_l'(ka) j_l(k'a) \right\}, \tag{11.74}$$

$$B_l = -A_l' \left\{ (ka)(k'a) j_l(ka) j_l'(k'a) - (ka)^2 j_l'(ka) j_l(k'a) \right\}. \tag{11.75}$$

Here j_l' and n_l' denote derivatives of j_l and n_l with respect to their argument. Once again phase shifts are obtained exactly.

11.8 Analytic Properties of Partial Waves

Let us now come back to a general potential and to (11.52). For small r physically acceptable solutions of (11.52) behave as r^{l+1} or, equivalently, as $rj_l(kr)$. It is convenient to normalize the function $\bar{u}_l$ so that

$$\bar{u}_{l,k}(r) \to krj_l(kr), \qquad r \to 0. \tag{11.76}$$

We have shown explicitly the k dependence of the solution.

One can show (Exercise 11.10) that $\bar{u}_{l,k}(r)$ satisfies the following integral equation:

$$\bar{u}_{l,k}(r) = kr \left\{ j_l(kr) + \int_0^r \mathrm{d}r' \, g_{l,k}(r,r') \frac{2m}{\hbar^2} V(r') \bar{u}_{l,k}(r') \right\}. \tag{11.77}$$

Here the kernel $g_{l,k}(r,r')$ can be written in terms of spherical Bessel functions as

$$\begin{aligned} g_{l,k}(r,r') &= r' \left\{ j_l(kr) n_l(kr') - n_l(kr) j_l(kr') \right\} \\ &= \frac{\mathrm{i}}{2} r' \left\{ h_l^-(kr) h_l^+(kr') - h_l^+(kr) h_l^-(kr') \right\}. \end{aligned} \tag{11.78}$$

From the integral equation it is easy to obtain a series representation of the function $\bar{u}_{l,k}$. Provided that the potential meets the conditions stated in Sect. 11.2, this series always converges.[10]

Any solution of (11.52) can, for large r, be written as a linear combination of functions rj_l and rn_l or, equivalently, of functions rh_l^- and rh_l^+. Besides, we know that $\bar{u}_{l,k}$ is real and that, for real arguments, $(h_l^-)^* = h_l^+$. Hence the asymptotic form of $\bar{u}_{l,k}(r)$ is given by

$$\bar{u}_{l,k}(r) \to \frac{1}{2}kr\left\{f_l(k)h_l^-(kr) + f_l^*(k)h_l^+(kr)\right\}, \qquad r \to \infty. \tag{11.79}$$

The coefficient $f_l(k)$ is called the *Jost function* and is directly related to the functions $S_l(k)$ introduced in Sect. 11.6. Indeed making use of the asymptotic expansion of h_l^- and h_l^+ and comparing (11.79) with (11.62), one easily sees that

$$S_l(k) = \frac{f_l^*(k)}{f_l(k)} \quad \text{and} \quad f_l(k) = |f_l(k)|\exp\left\{-\mathrm{i}\delta_l(k)\right\}. \tag{11.80}$$

Clearly, for every l and every k the Jost function can be computed by means of the numerical methods developed in Chap. 8. If, on the other hand, the integral equation is used for large values of r, then one can see that (Exercise 11.11)

$$f_l(k) = 1 + \mathrm{i}\int_0^\infty \mathrm{d}r'\, r' h_l^+(kr')\frac{2m}{\hbar^2}V(r')\bar{u}_{l,k}(r'). \tag{11.81}$$

This equation yields an important property of the Jost function. To see this, first note that

$$\bar{u}_{l,-k}(r) = (-1)^{l+1}\bar{u}_{l,k}(r)\,,$$
$$h_l^+(-kr) = (-1)^l\left\{h_l^+(kr)\right\}^*.$$

One then finds that

$$f_l(-k) = f_l^*(k). \tag{11.82}$$

The reader may have noticed that, in the last three equations, negative values of the variable k were introduced. This causes no problem because (11.52), our starting point, is well-defined at such values. But (11.52) is also defined for any complex value of k. Actually, it is very useful to take k as a complex variable. One can show that $\bar{u}_{l,k}(r)$, the solution of (11.52) or of the integral equation, is then an analytic function of k everywhere in the complex plane.

---

[10] The present section follows Taylor ([224], Chaps. 11–13), where proofs can be found for certain theorems that we simply state. A graphical display of analytic properties of partial waves can be found in [5].

Through its integral representation, the Jost function can also be defined for complex values of k. One finds that $f_l(k)$ is always analytic if $\mathrm{Im}(k) > 0$, and continuous if $\mathrm{Im}(k) \geq 0$. More restrictive conditions on the potential can extend the region of analyticity. If for large r the potential decreases like $\mathrm{e}^{-\alpha r}$ $(\alpha > 0)$, then $f_l(k)$ is analytic for $\mathrm{Im}(k) > -\alpha/2$. If the potential vanishes for $r > a > 0$, then $f_l(k)$ is everywhere analytic.

For k real one has

$$\bar{u}_{l,k}(r) \to \frac{1}{2} kr \left\{ f_l(k) h_l^-(kr) + f_l(-k) h_l^+(kr) \right\}, \qquad r \to \infty. \tag{11.83}$$

Assume for the moment that $f_l(k)$ is everywhere analytic. It is clear that relation (11.83) is then true for all k. Thus

$$S_l(k) = \frac{f_l(-k)}{f_l(k)}, \tag{11.84}$$

so that the amplitude $S_l(k)$ is everywhere analytic, except where the Jost function vanishes.

Points where the Jost function vanishes present great interest. We first point out that $f_l(k)$ cannot vanish if k is real and nonzero. If such were the case $\bar{u}_{l,k}(r)$ would asymptotically vanish, implying that $\bar{u}_{l,k}(r)$ would vanish for all r.[11] To begin with, let us investigate the possibility of a zero of $f_l(k)$ in the upper half-plane, that is,

$$f_l(\bar{k}) = 0 \quad \text{with} \quad \mathrm{Im}(\bar{k}) > 0. \tag{11.85}$$

In this case we have

$$\begin{aligned} \bar{u}_{l,k}(r) &\to \frac{1}{2} \bar{k} r f_l(-\bar{k}) h_l^+(\bar{k}r) \\ &\to -\frac{\mathrm{i}}{2} f_l(-\bar{k}) \exp\left\{ \mathrm{i} \left[\bar{k} r - \frac{l\pi}{2} \right] \right\}, \qquad r \to \infty. \end{aligned} \tag{11.86}$$

Clearly $\bar{u}_{l,\bar{k}}(r)$ decreases exponentially for large r, and is therefore a normalizable function. This means that $\bar{E} = \hbar^2 \bar{k}^2 / 2m$ is an eigenvalue of the Hamiltonian, corresponding to a bound state. Conversely, to each energy $\bar{E}$ in the discrete spectrum there corresponds a normalizable wave function, hence a zero of the Jost function. Note that since discrete spectrum energies are real and negative, values $\bar{k}$ are all on the imaginary axis.

The correspondence between zeros of the Jost function in the upper half-plane and discrete energies is valid even if $f_l(k)$ is analytic in the upper half-plane only. If $f_l(k)$ is also analytic at $-\bar{k}$, the bound state corresponds to a singularity of $S_l(k)$. One can show that this singularity is always a simple pole.

---

[11] If $k = 0, \bar{u}_{l,k}(r)$ vanishes identically, so that the Jost function can have a zero at the origin.

The Jost function can also have zeros in the lower half-plane. The most interesting are no doubt the ones that lie immediately below the positive real axis, that is,

$$\bar{k} = k_R - \mathrm{i}k_I \quad \text{with} \quad k_R \gg k_I > 0 \,. \tag{11.87}$$

Suppose $\bar{k}$ is a simple zero of $f_l(k)$. Near $\bar{k}$ one has approximately

$$f_l(k) = \left(\frac{\mathrm{d}f_l}{\mathrm{d}k}\right)_{\bar{k}} (k - \bar{k}) \,. \tag{11.88}$$

If k is real then

$$\delta_l(k) = -\arg\{f_l(k)\} = \delta_c - \arg\left\{k - \bar{k}\right\} = \delta_c - \arctan\left(\frac{k_I}{k - k_R}\right) . \tag{11.89}$$

Here we have let

$$\delta_c = -\arg\left\{\left(\frac{\mathrm{d}f_l}{\mathrm{d}k}\right)_{\bar{k}}\right\} . \tag{11.90}$$

Clearly, if k goes from values much lower that k_R to values much higher, the phase shift $\delta_l(k)$ increases by π (Fig. 11.3).

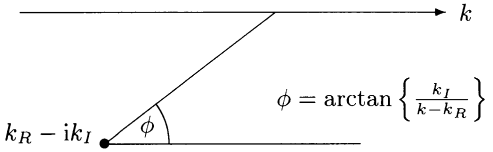

Fig. 11.3. Behavior of the phase shift near a zero of $f_l(k)$

The rapid increase of $\delta_l(k)$ by a value of π characterizes what is called a *resonance*. In the present case we will evaluate the contribution of the l-th partial wave to the total cross section, when k is near k_R. For simplicity we assume that $\delta_c = 0$. Furthermore, to keep up with standard notation we will write our results in terms of the variable E. Let

$$\bar{E} = \frac{\hbar^2}{2m}\bar{k}^2 = E_R - \frac{\mathrm{i}}{2}\Gamma \,. \tag{11.91}$$

It is easy to see that $E_R \gg \Gamma > 0$. For E real and $E \approx E_R$,

$$\frac{k_I}{k - k_R} \approx \frac{\Gamma}{2(E - E_R)} \,. \tag{11.92}$$

Hence

$$\delta_l(E) = -\arctan\left(\frac{\Gamma}{2(E - E_R)}\right). \tag{11.93}$$

By an elementary trigonometric argument one gets

$$\sin\delta_l(E) = -\frac{\Gamma/2}{\{(E - E_R)^2 + (\Gamma/2)^2\}^{1/2}}. \tag{11.94}$$

The contribution of the l-th partial wave to the total cross section is therefore given by

$$\sigma_l = \frac{4\pi}{k^2}(2l+1)\sin^2\delta_l(E) = \frac{2\pi\hbar^2}{mE_R}(2l+1)\frac{(\Gamma/2)^2}{(E - E_R)^2 + (\Gamma/2)^2}. \tag{11.95}$$

This is the *Breit–Wigner* formula. With the hypotheses made we also have

$$S_l(E) = \exp\{2\mathrm{i}\delta_l(E)\} = \frac{E - E_R - \mathrm{i}\Gamma/2}{E - E_R + \mathrm{i}\Gamma/2}. \tag{11.96}$$

It is important to note that the Breit–Wigner formula characterizes a specific type of resonance. It is not difficult to obtain the corresponding expression of σ_l in the case where $\delta_c \neq 0$. We should also point out that the correspondence between resonances and zeros of the Jost function is not universal. Indeed a zero far from the real axis will not produce a resonance, and it may happen that $\delta_l(k)$ increases rapidly by π without the Jost function vanishing. Nevertheless, the Breit–Wigner formula does characterize a large number of resonances actually occurring. Besides, in the neighborhood of a resonance σ_l often represents the most important contribution to the total cross section, so that the latter itself is well represented by the Breit–Wigner formula.

11.9 Scattering by Many Identical Centers

Throughout the foregoing discussion, the potential has been associated with one scattering center only. As a matter of fact, experimentalists work with a large number of centers. Under certain hypotheses, we pointed out in Sect. 11.1 how to extract the differential cross section from measurements made with many scattering centers. We now come back to this problem. Suppose N scattering centers, all identical, are situated at points $\boldsymbol{r}_i$ ($i = 1, 2, \ldots, N$). We will soon consider the case where scattering centers have changing and largely random positions, and the case where their positions are fixed and periodic.

We assume that the total potential, due to all scattering centers, can be represented as

$$V(\boldsymbol{r}) = \sum_{i=1}^{N} v(\boldsymbol{r} - \boldsymbol{r}_i). \tag{11.97}$$

This relation is clearly valid if scattering centers are far enough not to interact significantly with each other. Moreover $v(\boldsymbol{r})$ then represents the potential due to one center only. If, on the other hand, scattering centers interact, their structure can be more or less altered. Relation (11.97) can nonetheless remain valid, without $v(\boldsymbol{r})$ representing the potential due to an isolated center.

We will assume that the scattering amplitude is correctly given by the Born approximation. Thus one can write[12]

$$f(\theta, \phi) = -\frac{1}{4\pi} \int \mathrm{d}\boldsymbol{r}' \exp\{-\mathrm{i}\Delta\boldsymbol{k} \cdot \boldsymbol{r}'\} \frac{2m}{\hbar^2} \sum_{i=1}^{N} v(\boldsymbol{r}' - \boldsymbol{r}_i).$$

Taking the summation outside the integral and making the change of variables $\boldsymbol{\varrho} = \boldsymbol{r}' - \boldsymbol{r}_i$, one gets

$$\begin{aligned} f(\theta, \phi) &= -\frac{1}{4\pi} \sum_{i=1}^{N} \int \mathrm{d}\boldsymbol{\varrho} \exp\{-\mathrm{i}\Delta\boldsymbol{k} \cdot (\boldsymbol{\varrho} + \boldsymbol{r}_i)\} \frac{2m}{\hbar^2} v(\boldsymbol{\varrho}) \\ &= \left\{ -\frac{1}{4\pi} \int \mathrm{d}\boldsymbol{\varrho} \exp\{-\mathrm{i}\Delta\boldsymbol{k} \cdot \boldsymbol{\varrho}\} \frac{2m}{\hbar^2} v(\boldsymbol{\varrho}) \right\} \sum_{i=1}^{N} \exp\{-\mathrm{i}\Delta\boldsymbol{k} \cdot \boldsymbol{r}_i\}. \end{aligned} \tag{11.98}$$

Thus $f(\theta, \phi)$ is equal to the scattering amplitude due to the potential $v(\boldsymbol{\varrho})$, multiplied by the sum of exponentials of $-\mathrm{i}\Delta\boldsymbol{k} \cdot \boldsymbol{r}_i$.

The first case we investigate is a material where scattering centers move constantly. This happens in a gas or in a liquid. The differential cross section involves the absolute square of the sum of exponentials in (11.98). Thus one can write

$$\begin{aligned} \sigma(\theta, \phi) &\sim \left| \sum_{i=1}^{N} \exp\{-\mathrm{i}\Delta\boldsymbol{k} \cdot \boldsymbol{r}_i\} \right|^2 \\ &\sim \sum_{i,j=1}^{N} \exp\{-\mathrm{i}\Delta\boldsymbol{k} \cdot (\boldsymbol{r}_i - \boldsymbol{r}_j)\} \\ &\sim N + \sum_{i \neq j} \exp\{-\mathrm{i}\Delta\boldsymbol{k} \cdot (\boldsymbol{r}_i - \boldsymbol{r}_j)\}. \end{aligned} \tag{11.99}$$

---

[12] In the present case, (11.32) used to obtain Born's formula is valid if $L \ll r$ and $kL^2 \ll r$, where k is the wave number of the incoming beam, L represents linear dimensions of the bulk of scattering centers and r is the distance between sample and observation point. Note that both inequalities can be valid even if $kL \gg 1$. Recall finally that the incoming beam intensity should not be significantly attenuated in the sample.

Insofar as the number of scattering centers is very large, the environment of all centers not to close to the sample's boundary is, on the average, the same. Let $n(\boldsymbol{r})$ be the mean density of scattering centers at a point $\boldsymbol{r}$ with respect to a given center. Then one can write

$$\sum_{i \neq j} \exp\{-\mathrm{i}\Delta \boldsymbol{k} \cdot (\boldsymbol{r}_i - \boldsymbol{r}_j)\} = N \int \mathrm{d}\boldsymbol{r}\, n(\boldsymbol{r}) \exp\{-\mathrm{i}\Delta \boldsymbol{k} \cdot \boldsymbol{r}\}, \tag{11.100}$$

so that

$$\sigma(\theta, \phi) \sim N \left[1 + \int \mathrm{d}\boldsymbol{r}\, n(\boldsymbol{r}) \exp\{-\mathrm{i}\Delta \boldsymbol{k} \cdot \boldsymbol{r}\}\right]. \tag{11.101}$$

The differential cross section is therefore proportional to the total number of scattering centers, and is related to the position probability of one center relative to another.

The second case we investigate is a perfect crystal, i.e. a material where scattering centers are distributed on a periodic lattice. Let $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ be the fundamental lattice translation vectors. We assume the crystal is a parallelepiped. The scattering centers' positions are then given by (Fig. 11.4)[13]

$$\boldsymbol{r}_i = n_a \boldsymbol{a} + n_b \boldsymbol{b} + n_c \boldsymbol{c}. \tag{11.102}$$

Here n_a, n_b and n_c are three integers such that

$$0 \leq n_a \leq N_a - 1, \qquad 0 \leq n_b \leq N_b - 1, \qquad 0 \leq n_c \leq N_c - 1. \tag{11.103}$$

The number of scattering centers N is equal to$N_a N_b N_c$.

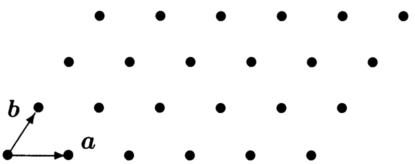

Fig. 11.4. Translation vectors in a two-dimensional lattice

It is easy to compute the sum of exponentials, which amounts to a product of geometric series. Indeed

$$\sum_{i=1}^{N} \exp\{-\mathrm{i}\Delta \boldsymbol{k} \cdot \boldsymbol{r}_i\} = \sum_{n_a=0}^{N_a-1} \sum_{n_b=0}^{N_b-1} \sum_{n_c=0}^{N_c-1} \exp\{-\mathrm{i}\Delta \boldsymbol{k} \cdot (n_a \boldsymbol{a} + n_b \boldsymbol{b} + n_c \boldsymbol{c})\}$$

---

[13] We do not assume that the number of scattering centers is equal to the number of atoms. There may be several atoms associated with each lattice node.

$$= \left[\sum_{n_a=0}^{N_a-1} \exp\{-\mathrm{i}n_a\Delta\boldsymbol{k}\cdot\boldsymbol{a}\}\right]\left[\sum_{n_b=0}^{N_b-1} \exp\{-\mathrm{i}n_b\Delta\boldsymbol{k}\cdot\boldsymbol{b}\}\right]$$
$$\times\left[\sum_{n_c=0}^{N_c-1} \exp\{-\mathrm{i}n_c\Delta\boldsymbol{k}\cdot\boldsymbol{c}\}\right]$$
$$= \left[\frac{1-\exp\{-\mathrm{i}N_a\Delta\boldsymbol{k}\cdot\boldsymbol{a}\}}{1-\exp\{-\mathrm{i}\Delta\boldsymbol{k}\cdot\boldsymbol{a}\}}\right]\left[\frac{1-\exp\{-\mathrm{i}N_b\Delta\boldsymbol{k}\cdot\boldsymbol{b}\}}{1-\exp\{-\mathrm{i}\Delta\boldsymbol{k}\cdot\boldsymbol{b}\}}\right]$$
$$\times\left[\frac{1-\exp\{-\mathrm{i}N_c\Delta\boldsymbol{k}\cdot\boldsymbol{c}\}}{1-\exp\{-\mathrm{i}\Delta\boldsymbol{k}\cdot\boldsymbol{c}\}}\right].$$

From this one gets

$$\left|\sum_{i=1}^{N}\exp\{-\mathrm{i}\Delta\boldsymbol{k}\cdot\boldsymbol{r}_i\}\right|^2 = F_a(\Delta\boldsymbol{k}\cdot\boldsymbol{a})\,F_b(\Delta\boldsymbol{k}\cdot\boldsymbol{b})\,F_c(\Delta\boldsymbol{k}\cdot\boldsymbol{c}), \tag{11.104}$$

where ($\alpha = a,\ b,\ c$; $\boldsymbol{\alpha} = \boldsymbol{a},\ \boldsymbol{b},\ \boldsymbol{c}$)

$$F_\alpha(\Delta\boldsymbol{k}\cdot\boldsymbol{\alpha}) = \frac{\sin^2\left\{\frac{1}{2}N_\alpha\Delta\boldsymbol{k}\cdot\boldsymbol{\alpha}\right\}}{\sin^2\left\{\frac{1}{2}\Delta\boldsymbol{k}\cdot\boldsymbol{\alpha}\right\}}. \tag{11.105}$$

The differential cross section is therefore given by

$$\sigma(\theta,\phi) = \left|\frac{1}{4\pi}\int \mathrm{d}\boldsymbol{\varrho}\,\exp\{-i\Delta\boldsymbol{k}\cdot\boldsymbol{\varrho}\}\,\frac{2m}{\hbar^2}v(\boldsymbol{\varrho})\right|^2$$
$$\times F_a(\Delta\boldsymbol{k}\cdot\boldsymbol{a})\,F_b(\Delta\boldsymbol{k}\cdot\boldsymbol{b})\,F_c(\Delta\boldsymbol{k}\cdot\boldsymbol{c}). \tag{11.106}$$

Given that N_α is large enough, the function $F_\alpha(\Delta\boldsymbol{k}\cdot\boldsymbol{\alpha})$ displays acute maxima at those points where $\Delta\boldsymbol{k}\cdot\boldsymbol{\alpha}$ is an integral multiple of 2π. The behavior of F_α around a maximum is shown in Fig. 11.5. The maximum's height is equal to N_α^2 and the interval between the nearest zeros on both sides of the central maximum is equal to $4\pi/N_\alpha$. The area under that maximum is proportional to N_α.

Thus for a periodic structure the differential cross section is appreciable only at values of $\Delta\boldsymbol{k}$ for which

$$\Delta\boldsymbol{k}\cdot\boldsymbol{a} = 2\pi q_a, \qquad \Delta\boldsymbol{k}\cdot\boldsymbol{b} = 2\pi q_b, \qquad \Delta\boldsymbol{k}\cdot\boldsymbol{c} = 2\pi q_c, \tag{11.107}$$

where q_a, q_b and q_c are integers. Observation of the position of maxima yields, with the help of relations (11.107), the lattice's fundamental translation vectors (Exercise 11.12). On the other hand, the relative intensity of the differential cross section's maxima depends on the first factor in the right-hand side of (11.106). Observation of the relative intensity thus yields information on the form of $v(\boldsymbol{\varrho})$. These notions are explained in detail in solid state physics books, for instance [133], Chap. 2.

Note that the integral of the differential cross section on a given maximum is proportional to $N_aN_bN_c = N$, that is, to the total number of scattering centers.

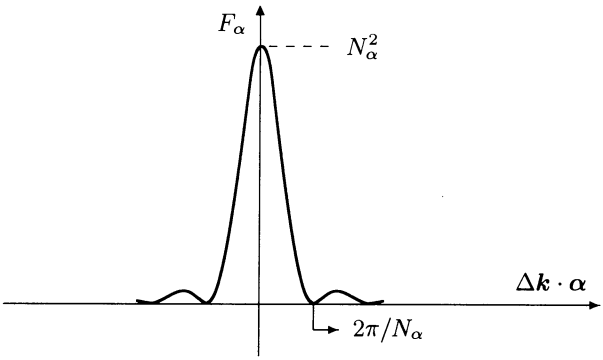

Fig. 11.5. The function $F_\alpha(\Delta \boldsymbol{k} \cdot \boldsymbol{\alpha})$ around a central maximum

Exercises

11.1. Compute the probability current associated with a wave function like

$$\psi(\boldsymbol{r}) = \frac{1}{r} \mathrm{e}^{\mathrm{i}kr} f(\theta, \phi),$$

where f is a complex function. [In spherical coordinates the gradient operator is given by

$$\boldsymbol{\nabla} = \hat{r}\frac{\partial}{\partial r} + \frac{\hat{\theta}}{r}\frac{\partial}{\partial \theta} + \frac{\hat{\phi}}{r\sin\theta}\frac{\partial}{\partial \phi},$$

where $\hat{r}$, $\hat{\theta}$ and $\hat{\phi}$ are orthonormal vectors.]

11.2. Compute the Green's function $\langle \boldsymbol{r}|G_0(E - \mathrm{i}0)|\boldsymbol{r}'\rangle$.

11.3. Show that the Green's function $\langle \boldsymbol{r}|G(\lambda)|\boldsymbol{r}'\rangle$ obeys the following differential equation:

$$\left\{\frac{\hbar^2}{2m}\nabla^2 - V(\boldsymbol{r}) + \lambda\right\} \langle \boldsymbol{r}|G(\lambda)|\boldsymbol{r}'\rangle = \delta(\boldsymbol{r}-\boldsymbol{r}').$$

11.4. Compute the following integral, for $\alpha > 0$:

$$\int \mathrm{d}\boldsymbol{s}\, \frac{1}{s} \exp\left\{-\alpha s - \mathrm{i}\Delta \boldsymbol{k} \cdot \boldsymbol{s}\right\}.$$

11.5. Compute the atomic form factor $F(\theta)$ in the situation where the electronic charge density is constant inside a sphere of radius a, and vanishes everywhere else.

11.6. Consider an electronic charge distribution $\varrho(r)$ proportional to $\mathrm{e}^{-r/a}$, where a is a positive constant.

a) Normalize the distribution so that $\int \mathrm{d}\boldsymbol{r}\, \varrho(r) = Zq_e$. Next compute the atomic form factor $F(\theta)$.
b) Compute the differential scattering cross section for the case where $ka \gg 1$ and $\theta \gg 0$, and for the case where $\theta \approx 0$.

11.7. Consider a spherically symmetric potential given by

$$V(r) = \frac{V_0 a}{r} \mathrm{e}^{-r/a} .$$

a) In the Born approximation, compute the differential cross section associated with $V(r)$.
b) Discuss the approximation's conditions of validity for this potential.

11.8. Use the result of Exercise (7.17) to obtain (11.74) and (11.75).

11.9. Let $A > 0$ and consider the potential

$$V(r) = \frac{\hbar^2}{2m} \frac{A}{r^2} .$$

a) Write down the eigenvalue equation of the Hamiltonian associated with this potential, and find a solution in terms of spherical Bessel functions.
b) Obtain the phase shifts exactly. Show that for $A \ll 1$, one approximately has

$$\delta_l = -\frac{\pi}{2} \frac{A}{2l+1} .$$

What is the value of the total cross section? Comment on this result ([138], Sect. 124).

11.10. Use the result of Exercise (7.17) to check that the function $\bar{u}_{l,k}(r)$, defined in (11.77), satisfies (11.52). Next compute $g_{l,k}(r, r')$ for small r and r', and show that (11.77) is consistent with condition (11.76).

11.11. Obtain the integral representation (11.81) of the Jost function.

11.12. Let $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ be three linearly independent vectors. It is not difficult to see that the following equations define three vectors $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ uniquely:

$$\begin{aligned} &\boldsymbol{A} \cdot \boldsymbol{a} = 2\pi , \qquad && \boldsymbol{B} \cdot \boldsymbol{a} = 0 , \qquad && \boldsymbol{C} \cdot \boldsymbol{a} = 0 , \\ &\boldsymbol{A} \cdot \boldsymbol{b} = 0 , && \boldsymbol{B} \cdot \boldsymbol{b} = 2\pi , && \boldsymbol{C} \cdot \boldsymbol{b} = 0 , \\ &\boldsymbol{A} \cdot \boldsymbol{c} = 0 , && \boldsymbol{B} \cdot \boldsymbol{c} = 0 , && \boldsymbol{C} \cdot \boldsymbol{c} = 2\pi . \end{aligned}$$

a) Show that

$$\boldsymbol{a} = 2\pi \frac{\boldsymbol{B} \times \boldsymbol{C}}{\boldsymbol{A} \cdot (\boldsymbol{B} \times \boldsymbol{C})}, \quad \boldsymbol{b} = 2\pi \frac{\boldsymbol{C} \times \boldsymbol{A}}{\boldsymbol{B} \cdot (\boldsymbol{C} \times \boldsymbol{A})}, \quad \boldsymbol{c} = 2\pi \frac{\boldsymbol{A} \times \boldsymbol{B}}{\boldsymbol{C} \cdot (\boldsymbol{A} \times \boldsymbol{B})} .$$

b) Show that (11.107) is satisfied if and only if

$$\Delta \boldsymbol{k} = q_a \boldsymbol{A} + q_b \boldsymbol{B} + q_c \boldsymbol{C} .$$

A measurement of $\Delta \boldsymbol{k}$ thus yields $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ and, consequently, $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$.

12 The Density Operator

The state of an isolated quantum system has hitherto been represented by a vector in the state space. We shall see that it can also be represented by a Hermitian operator called the density operator. The usefulness of that operator comes from the fact that it can represent not only the state of an isolated system, but also the state of a system that genuinely interacts with the environment, or even the state of an ensemble of systems prepared in different ways. We will introduce the notion of tensor product of vector spaces, a mathematical tool well adapted to the description of a system made of two or more subsystems. The density operator finds a particularly important application in statistical physics. It will also be used to introduce the notion of history, useful for the clarification of interpretation questions.

12.1 Pure State

The eigenvalue equation associated with a dynamical variable A can be written as

$$A|a_i^\alpha\rangle = a_i|a_i^\alpha\rangle ,$$

where the a_i are eigenvalues and the $|a_i^\alpha\rangle$ orthonormal eigenvectors of A. As in Chap. 10 the exponent α keeps track of eigenvalues' degeneracy. In Sect. 3.2 we saw that the probability that a measurement of A in a normalized state $|\psi(t)\rangle$ yields value a_i is equal to $\sum_\alpha |\langle a_i^\alpha|\psi(t)\rangle|^2$.

It is easy to see that the expression for the probability can also be written as

$$P(A \to a_i; t) = \sum_\alpha \left|\langle a_i^\alpha|\psi(t)\rangle\right|^2 = \mathrm{Tr}\,\{\varrho(t)P_{a_i}\} , \tag{12.1}$$

where

$$\varrho(t) = |\psi(t)\rangle\langle\psi(t)| \tag{12.2}$$

and

$$P_{a_i} = \sum_\alpha |a_i^\alpha\rangle\langle a_i^\alpha| . \tag{12.3}$$

L. Marchildon, *Quantum Mechanics*

Here P_{a_i} is simply the projector on the subspace of eigenvectors of A associated with the eigenvalue a_i. The symbol $\varrho(t)$ denotes a Hermitian operator called the *density* (or *statistical*) *operator.*

As with the probability of a measurement result, the mean value of dynamical variable A in state $|\psi(t)\rangle$ can be written in terms of the density operator. Indeed

$$\langle A\rangle_{\psi(t)} = \langle\psi(t)|A|\psi(t)\rangle = \mathrm{Tr}\,\{\varrho(t)A\}\,. \tag{12.4}$$

One easily checks that the operator $\varrho(t)$, defined in (12.2), has the following properties:

$$\varrho = \varrho^\dagger\,, \tag{12.5}$$

$$\mathrm{Tr}(\varrho) = 1\,, \tag{12.6}$$

$$\varrho\varrho = \varrho\,. \tag{12.7}$$

Equations (12.5) and (12.7) mean that $\varrho(t)$ is a (time-dependent) projector, while (12.6) entails that the subspace on which ϱ projects has dimension one. It is, of course, the subspace corresponding to $|\psi(t)\rangle$.

Thus computation of measurement result probabilities and mean values can be carried out from the density operator as well as from the state vector. Both are equivalent ways to describe a physical system. The second quantum-mechanical interpretation rule, stated on p. 32, can thus be formulated in the following way: At a given time, the state of an isolated quantum system[1] is described by an operator satisfying (12.5)–(12.7). One can easily see that knowledge of $|\psi(t)\rangle$ yields $\varrho(t)$, while knowledge of $\varrho(t)$ determines $|\psi(t)\rangle$ up to a multiplicative factor.

The time evolution of the density operator is governed by an equation that follows from Schrödinger's equation. Indeed

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\varrho(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\{|\psi(t)\rangle\langle\psi(t)|\}\\ &= \left\{\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle\right\}\langle\psi(t)| + |\psi(t)\rangle\left\{\frac{\mathrm{d}}{\mathrm{d}t}\langle\psi(t)|\right\}\\ &= \left\{\frac{1}{\mathrm{i}\hbar}H|\psi(t)\rangle\right\}\langle\psi(t)| + |\psi(t)\rangle\left\{\frac{-1}{\mathrm{i}\hbar}\langle\psi(t)|H\right\}.\end{aligned}$$

Thus one finds that

$$\frac{\mathrm{d}}{\mathrm{d}t}\varrho(t) = \frac{1}{\mathrm{i}\hbar}\,[H,\varrho(t)]\,. \tag{12.8}$$

This equation can be solved formally as

---

[1] Recall that an isolated quantum system interacts with its environment only through classical fields.

$$\varrho(t) = U(t, t_0)\varrho(t_0)U^\dagger(t, t_0), \tag{12.9}$$

where U is the system's evolution operator and t_0 is an arbitrary initial time. Equation (12.9) directly follows from (3.13).

It is shown in Exercise (12.1) that Schrödinger's equation for the state vector can be recovered from (12.8).

12.2 Statistical Mixture

Let us now consider an ensemble of identical and isolated quantum systems, which however are not all similarly prepared. Specifically, we assume that for different values of n a fraction w_n of systems is described by the state vector $|\psi_n(t)\rangle$. Clearly $w_n \geq 0$ and $\sum_n w_n = 1$. State vectors are normalized but not necessarily orthogonal. The probability that a measurement of A yields value a_i is given by a sum of expressions like the left-hand side of (12.1), weighed by the w_n, that is,

$$\begin{aligned} P(A \to a_i; t) &= \sum_n w_n \left\{ \sum_\alpha |\langle a_i^\alpha | \psi_n(t)\rangle|^2 \right\} \\ &= \sum_n w_n \mathrm{Tr}\left\{ |\psi_n(t)\rangle\langle\psi_n(t)| P_{a_i} \right\} \\ &= \mathrm{Tr}\left\{ \varrho(t) P_{a_i} \right\}. \end{aligned} \tag{12.10}$$

Here P_{a_i} is still given by (12.3) but now

$$\varrho(t) = \sum_n w_n |\psi_n(t)\rangle\langle\psi_n(t)|. \tag{12.11}$$

In the present case $\varrho(t)$ is still referred to as the density operator, but it now corresponds to a *statistical mixture.* It is not difficult to show that $\varrho(t)$, defined in (12.11), has the following properties:

$$\varrho = \varrho^\dagger, \tag{12.12}$$

$$\mathrm{Tr}(\varrho) = 1. \tag{12.13}$$

We point out, however, that $\varrho^2 \neq \varrho$.

Fractions w_n can depend on time or not. Whenever they are time independent, one easily checks (Exercise 12.2) that $\varrho(t)$ satisfies the evolution equation (12.8).

The density operator for a statistical mixture was defined in (12.11). Is it possible, however, that a statistical mixture be equivalent to a pure state? In other words, can a vector $|\psi(t)\rangle$ exist so that the right-hand side of (12.11) is equal to $|\psi(t)\rangle\langle\psi(t)|$? The answer to this question is negative. The proof is contained in the following theorem.

Theorem Let $\varrho(t) = \sum_n w_n|\psi_n\rangle\langle\psi_n|$, where $\sum_n w_n = 1$, $w_n > 0$, $|\psi_n\rangle$ is normalized (for every n) and $|\psi_n\rangle$ is not a multiple of $|\psi_m\rangle$ if $m \neq n$. Then the following statements are equivalent:

i) There exists a normalized $|\psi\rangle$ such that $\varrho(t) = |\psi\rangle\langle\psi|$.

ii) $\varrho^2 = \varrho$.

iii) $\mathrm{Tr}\left\{\varrho^2\right\} = 1$.

iv) There is only one w_n in the expression for ϱ.

Proof It is obvious that (i) implies (ii), that (ii) implies (iii) and that (iv) implies (i). Hence it is enough to show that (iii) implies (iv). But

$$\begin{aligned}
\mathrm{Tr}\left\{\varrho^2\right\} &= \mathrm{Tr}\left\{\left[\sum_n w_n|\psi_n\rangle\langle\psi_n|\right]\left[\sum_m w_m|\psi_m\rangle\langle\psi_m|\right]\right\} \\
&= \sum_{n,m} w_n w_m \mathrm{Tr}\left\{|\psi_n\rangle\left[\langle\psi_n|\psi_m\rangle\right]\langle\psi_m|\right\} \\
&= \sum_n w_n \left\{\sum_m w_m \left|\langle\psi_n|\psi_m\rangle\right|^2\right\}.
\end{aligned}$$

From the Schwartz inequality (2.6) and Exercise (2.4), we immediately see that $|\langle\psi_n|\psi_m\rangle|^2 < 1$ if $m \neq n$. If there is more than one value of the index, i.e. more than one nonzero w_n, we have

$$\sum_m w_m \left|\langle\psi_n|\psi_m\rangle\right|^2 < \sum_m w_m = 1\,.$$

Thus

$$\mathrm{Tr}\left\{\varrho^2\right\} < \sum_n w_n = 1\,.$$

Hence (iii) implies (iv). ♣

As a corollary of the previous theorem, we point out that the statistical mixture $\varrho(t) = \sum_n w_n|\psi_n\rangle\langle\psi_n|$ is completely different from the pure state described by the vector $|\psi\rangle = \sum_n \sqrt{w_n}\mathrm{e}^{\mathrm{i}\alpha_n}|\psi_n\rangle$. This can be checked directly (Exercise 12.5) by computing mean values of various dynamical variables in the pure state and in the mixture.

One can show that the density operator is the most general way of attributing additive probabilities to subspaces of the state space. This result is known as *Gleason's theorem* ([124], Chap. 7; [173], Chap. 3).

It is clear that the state preparation procedure described in Sect. 6.3 yields a pure state. Indeed suppose that a procedure yields an ensemble of microobjects described by the density operator $\varrho = w_1|\psi_1\rangle\langle\psi_1| + w_2|\psi_2\rangle\langle\psi_2|$, where $w_1 \neq 0 \neq w_2$ and $\langle\psi_1|\psi_2\rangle = 0$. One can then, in principle, separate microobjects associated with $|\psi_1\rangle$ and $|\psi_2\rangle$ in two subensembles S_1 and S_2. Obviously a subsequent merging of S_1 and S_2 yields an ensemble that has the same statistical properties as the initial mixture.

12.3 Tensor Product of State Spaces

Consider two independent quantum systems labeled (1) et (2). We denote the corresponding Hamiltonians by $H(1)$ and $H(2)$. Let $\{|u_i(1)\rangle\}$ be an orthonormal basis of the first system's state space $\mathcal{V}^1$ and let $\{|v_k(2)\rangle\}$ be an orthonormal basis of the second system's state space $\mathcal{V}^2$.[2]

The two independent systems can be viewed as one large system. We will develop a way to describe this composite system. For this purpose we first introduce the notion of tensor product of vector spaces.

By definition, the *tensor product* of two vector spaces $\mathcal{V}^1$ and $\mathcal{V}^2$ is the vector space, denoted by $\mathcal{V}^1 \otimes \mathcal{V}^2$, which is specified in the following way: to each pair $|u_i(1)\rangle$ and $|v_k(2)\rangle$ of basis vectors in $\mathcal{V}^1$ and $\mathcal{V}^2$ there corresponds a basis vector $|u_i(1)\rangle \otimes |v_k(2)\rangle$ in $\mathcal{V}^1 \otimes \mathcal{V}^2$; and to linear combination like

$$|\phi(1)\rangle = \sum_i c_i |u_i(1)\rangle \quad \text{and} \quad |\psi(2)\rangle = \sum_k d_k |v_k(2)\rangle \tag{12.14}$$

there corresponds the linear combination

$$|\phi(1)\rangle \otimes |\psi(2)\rangle = \sum_{i,k} c_i d_k |u_i(1)\rangle \otimes |v_k(2)\rangle \tag{12.15}$$

in the product space. Thus to each pair of vectors $|\phi(1)\rangle$ in $\mathcal{V}^1$ and $|\psi(2)\rangle$ in $\mathcal{V}^2$ there corresponds a vector in the product space. But it is not true that to each vector in $\mathcal{V}^1 \otimes \mathcal{V}^2$ there corresponds a pair $|\phi(1)\rangle$, $|\psi(2)\rangle$. Indeed for any complex matrix M_{ik} the linear combination

$$\sum_{i,k} M_{ik} |u_i(1)\rangle \otimes |v_k(2)\rangle \tag{12.16}$$

belongs to $\mathcal{V}^1 \otimes \mathcal{V}^2$. That vector, however, can be written in the form $|\phi(1)\rangle \otimes |\psi(2)\rangle$ only if matrix M_{ik} factorizes into a product $c_i d_k$.

It is easy to see that if $\mathcal{V}^1$ and $\mathcal{V}^2$ have finite dimensions equal to N_1 and N_2, then $\mathcal{V}^1 \otimes \mathcal{V}^2$ has dimension equal to $N_1 N_2$.

In the product space a scalar product can be defined in a natural way. For two basis vectors one lets

$$\begin{aligned}
&(|u_i(1)\rangle \otimes |v_k(2)\rangle, |u_j(1)\rangle \otimes |v_l(2)\rangle) \\
&\quad \equiv (\langle u_i(1)| \otimes \langle v_k(2)|)\,(|u_j(1)\rangle \otimes |v_l(2)\rangle) \\
&\quad = \langle u_i(1)|u_j(1)\rangle \langle v_k(2)|v_l(2)\rangle \,.
\end{aligned} \tag{12.17}$$

For arbitrary vectors the scalar product is defined so as to meet the conditions of complex conjugation and linearity.

---

[2] Note that in general indices i and k do not have the same range. The notation is adapted to the case where indices are discrete, but they may be continuous.

One can show that the tensor product of two Hilbert spaces is itself a Hilbert space.

Let $A(1)$ and $B(2)$ be two linear operators defined in $\mathcal{V}^1$ and $\mathcal{V}^2$ respectively. The tensor product of $A(1)$ and $B(2)$, denoted by $A(1) \otimes B(2)$, is defined as the linear operator which, in the product space, acts on basis vectors in the following way:

$$\{A(1) \otimes B(2)\} \{|u_i(1)\rangle \otimes |v_k(2)\rangle\} = \{A(1)|u_i(1)\rangle\} \otimes \{B(2)|v_k(2)\rangle\} . \tag{12.18}$$

The action of $A(1) \otimes B(2)$ on an arbitrary vector is obtained by linearity.

Example The representation of bras and kets by means of rows and columns helps to understand the meaning of the tensor product. Thus

$$\begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix} , \qquad (a^* b^*) \otimes (c^* d^*) = (a^* c^* \ a^* d^* \ b^* c^* \ b^* d^*).$$

Likewise, if σ_x and σ_y denote Pauli matrices,

$$\sigma_x \otimes \sigma_y = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & -\mathrm{i} \\ 0 & 0 & \mathrm{i} & 0 \\ 0 & -\mathrm{i} & 0 & 0 \\ \mathrm{i} & 0 & 0 & 0 \end{pmatrix} . \qquad \clubsuit$$

Let us come back to the two independent quantum systems introduced at the beginning of this section. With the product space formalism, we can describe the composite (or total) system. To the first system's normalized state vector $|\phi(1;t)\rangle$ and the second system's normalized state vector $|\psi(2;t)\rangle$, there corresponds the total system's state vector $|\phi(1,t)\rangle \otimes |\psi(2,t)\rangle$. It is easy to show (Exercise 12.6) that the derivative of a tensor product of vectors obeys the same law as the derivative of an ordinary product of functions. The time evolution of the total state vector is thus given by

$$\begin{aligned} \mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t} & \{|\phi(1;t)\rangle \otimes |\psi(2;t)\rangle\} \\ &= \mathrm{i}\hbar \left\{ \frac{\mathrm{d}}{\mathrm{d}t} |\phi(1;t)\rangle \right\} \otimes |\psi(2;t)\rangle + \mathrm{i}\hbar |\phi(1;t)\rangle \otimes \left\{ \frac{\mathrm{d}}{\mathrm{d}t} |\psi(2;t)\rangle \right\} \\ &= (H(1)|\phi(1;t)\rangle) \otimes |\psi(2;t)\rangle + |\phi(1;t)\rangle \otimes (H(2)|\psi(2;t)\rangle) \\ &= \{H(1) \otimes I(2) + I(1) \otimes H(2)\} \{|\phi(1;t)\rangle \otimes |\psi(2;t)\rangle\} , \end{aligned} \tag{12.19}$$

where $I(1)$ and $I(2)$ are identity operators in spaces $\mathcal{V}^1$ and $\mathcal{V}^2$. Hence the total system's state vector satisfies the Schrödinger equation and the total system's Hamiltonian is given by

$$H(1,2) = H(1) \otimes I(2) + I(1) \otimes H(2) . \tag{12.20}$$

Description by the product space formalism is equivalent to separate descriptions of the two systems. For example, the mean value of dynamical variable $A(1)$ in state $|\phi(1,t)\rangle$, given by

$$\langle A(1)\rangle_{\phi(1)} = \langle\phi(1)|A(1)|\phi(1)\rangle\,, \tag{12.21}$$

can just as well be interpreted as the mean value of $A(1) \otimes I(2)$ in state $|\phi(1;t)\rangle \otimes |\psi(2;t)\rangle$. Indeed one sees at once that

$$\langle A(1)\rangle_{\phi(1)} = (\langle\phi(1)| \otimes \langle\psi(2)|)\, A(1)\otimes I(2)\, (|\phi(1)\rangle \otimes |\psi(2)\rangle)\,. \tag{12.22}$$

One can also show (Exercise 12.7) that the separate systems' Schrödinger equations can be recovered from the total system's Schrödinger equation.

Up to now we have considered independent quantum systems. The product space formalism is particularly useful in the more general case where there is interaction between the two systems. The total system's Hamiltonian can then be written as

$$H(1,2) = H(1)\otimes I(2) + I(1)\otimes H(2) + H_I(1,2)\,, \tag{12.23}$$

where H_I is an interaction term that cannot be expressed as a sum of two operators associated with systems (1) and (2), respectively. One can define dynamical variables $A(1,2)$ that are not tensor products of operators associated with subsystems. The state vector $|\chi(1,2;t)\rangle$ is no longer the product of two vectors like $|\phi(1)\rangle$ and $|\psi(2)\rangle$, but rather a general vector in the product space, like

$$|\chi(1,2;t)\rangle = \sum_{i,k} c_{ik}|u_i(1)\rangle\otimes|v_k(2)\rangle\,. \tag{12.24}$$

12.4 Description of a Subsystem

To make equations simpler, it is convenient to simplify the notations we have introduced. The symbols (1) and (2) will be suppressed where no confusion if possible. Furthermore, we will write $|u_i(1)\rangle \otimes |v_k(2)\rangle$ as $|u_i\rangle|v_k\rangle$ or as $|u_i v_k\rangle$; $A(1) \otimes I(2)$ as $A(1)$; and $A(1) \otimes B(2)$ as $A(1)B(2)$. We will also assume that $|u_i\rangle$ and $|v_k\rangle$ are orthonormal bases of $\mathcal{V}^1$ and $\mathcal{V}^2$, so that $|u_i v_k\rangle$ is an orthonormal basis of $\mathcal{V}^1 \otimes \mathcal{V}^2$.

Consider a quantum system made up of two interacting subsystems. The Hamiltonian is given by

$$H(1,2) = H(1) + H(2) + H_I(1,2)\,. \tag{12.25}$$

The total system's normalized state vector can be written as

$$|\chi(1,2)\rangle = \sum_{i,k} c_{ik}(t)|u_i v_k\rangle . \tag{12.26}$$

Equivalently, the total system is described by the density operator

$$\varrho(1,2) = |\chi\rangle\langle\chi| = \sum_{i,j,k,l} c_{ik}^*(t)c_{jl}(t)|u_j v_l\rangle\langle u_i v_k| . \tag{12.27}$$

Consider now a dynamical variable $B(2)$ associated with the second subsystem only. As in the case of independent systems, this should coincide with the total system's dynamical variable $I(1) \otimes B(2)$. The mean value of $B(2)$ in state $|\chi(1,2)\rangle$ is therefore given by

$$\begin{aligned} \langle B(2)\rangle_{\chi(1,2)} &= \langle\chi(1,2)|B(2)|\chi(1,2)\rangle \\ &= \left\{\sum_{i,k} c_{ik}^*(t)\langle u_i v_k|\right\} B(2) \left\{\sum_{j,l} c_{jl}(t)|u_j v_l\rangle\right\} \\ &= \sum_{i,j,k,l} c_{ik}^*(t)c_{jl}(t)\langle u_i|u_j\rangle\langle v_k|B(2)|v_l\rangle \\ &= \mathrm{Tr}\left\{\varrho(2)B(2)\right\} . \end{aligned} \tag{12.28}$$

Here $\varrho(2)$ was introduced as

$$\varrho(2) = \sum_{k,l}\left\{\sum_i c_{ik}^*(t)c_{il}(t)\right\}|v_l\rangle\langle v_k| . \tag{12.29}$$

The operator $\varrho(2)$ acts in the space of the second subsystem. Formally, it is obtained from $\varrho(1,2)$ by replacing $|u_j\rangle\langle u_i|$ by $\langle u_i|u_j\rangle = \delta_{ij}$. For this reason $\varrho(2)$ is called the *partial trace* of $\varrho(1,2)$ on the variables (1), and it is denoted as

$$\varrho(2) = \mathrm{Tr}_1\left\{\varrho(1,2)\right\} . \tag{12.30}$$

It is easy to check that $\varrho(2) = \varrho(2)^\dagger$ and that $\mathrm{Tr}\{\varrho(2)\} = 1$. Since $\varrho(2)$ is Hermitian, the spectral decomposition theorem entails there exists an orthonormal basis $|\psi_k(2)\rangle$ such that[3]

$$\varrho(2) = \sum_k w_k|\psi_k(2)\rangle\langle\psi_k(2)| , \tag{12.31}$$

where the w_k are eigenvalues of $\varrho(2)$. To prove that $\varrho(2)$ is a density operator (in the space of the second subsystem), there remains to show that $w_k \geq 0$ for every k. To do this note that, for any $|v\rangle$,

---

[3] We assume that this result, familiar if $\varrho(2)$ acts in a finite-dimensional space, can be adapted to infinite-dimensional spaces.

$$\langle v|\varrho(2)|v\rangle = \sum_m \langle u_m v|\varrho(1,2)|u_m v\rangle = \sum_m |\langle u_m v|\chi(1,2)\rangle|^2 \geq 0.$$

Since $w_k = \langle\psi_k|\varrho(2)|\psi_k\rangle$, the result follows.

Thus $\varrho(2)$ is a density operator. Due to the term $H_I(1,2)$ in the total Hamiltonian, however, the time evolution of $\varrho(2)$ cannot be obtained through a Hamiltonian involving only operators pertaining to the second subsystem.[4] Nevertheless all the information relative to the first subsystem needed to compute mean values of operators like $B(2)$ is contained in the density operator $\varrho(2)$.

Subsystem (2) is not an isolated quantum system in the sense of the definition given in Sect. 3.1. Indeed it interacts with the environment [here subsystem (1)] through the interaction Hamiltonian $H_I(1,2)$, i.e. not through a classical field. The density operator (12.29) is said to represent an *improper mixture*, in contrast with the operator (12.11) which represents a *proper mixture* ([67], Chap. 7). In a proper mixture each quantum system is in a well-defined, albeit unknown, state. In an improper mixture, on the other hand, the influence of the environment on the subsystem has the effect that the latter is not in a well-defined state.

To close this section we will prove two important results that will be needed in Chap. 21. The first one concerns the probability of getting a specific value upon measurement of a dynamical variable associated with a subsystem. The second one pertains to the state in which the first subsystem finds itself when a measurement made on the second subsystem has yielded a specific value.

We have seen how to compute the mean value of a dynamical variable $B(2)$ in state $|\chi(1,2)\rangle$. Consider the dynamical variable $P_b(2)$, that is, the projector on the space of eigenvectors of $B(2)$ with eigenvalue b. The mean value of $P_b(2)$ is equal to the probability that a measurement of $B(2)$ yields value b, and it is given by

$$\mathrm{Tr}\,\{\varrho(2)P_b(2)\} = \sum_{k,l}\left\{\sum_i c_{ik}^*(t)c_{il}(t)\right\}\langle v_k|P_b(2)|v_l\rangle. \tag{12.32}$$

Let $P_a(1)$ be the projector on the space of eigenvectors of a dynamical variable $A(1)$ that have eigenvalue a. The mean value of $P_a(1)\otimes P_b(2)$ is equal to the probability that a measurement of $A(1)$ yields value a and that a measurement of $B(2)$ yields value b, and is given by

$$\begin{aligned}&\langle\chi(1,2)|P_a(1)\otimes P_b(2)|\chi(1,2)\rangle\\ &\quad = \sum_{i,j,k,l} c_{ik}^*(t)c_{jl}(t)\langle u_i|P_a(1)|u_j\rangle\langle v_k|P_b(2)|v_l\rangle\,. \end{aligned}\tag{12.33}$$

Comparing (12.32) and (12.33), one easily checks the following result.

---

[4] The time evolution of subsystem (2) is not governed by (12.8) because the w_k depend on time.

Theorem 1 The probability that a measurement of $B(2)$ yields value b when nothing is done with system (1) is equal to the sum, on all values of a, of the joint probabilities that a measurement of $A(1)$ yields value a and that a measurement of $B(2)$ yields value b.

Now let $\{|b_k\rangle\}$ be an orthonormal basis of eigenvectors of $B(2)$, and suppose that eigenvalues b_k are not degenerate. At a given time the total system's state vector $|\chi(1,2)\rangle$ can be written as

$$|\chi(1,2)\rangle = \sum_k c_k |u_k\rangle \otimes |b_k\rangle \,, \tag{12.34}$$

where the $|u_k\rangle$ are normalized but not necessarily orthogonal. There are no restrictions on coefficients c_k other than the normalization of $|\chi\rangle$.

Theorem 2 Let (12.34) be the total system's state. If without disturbing system (1) one measures dynamical variable $B(2)$ and obtains value b_m, then immediately after measurement, the state of system (1) is necessarily given by $|u_m\rangle$.

Proof Let P_{u_m} be the projector on $|u_m\rangle$. Denote by $P(p|q)$ the probability of proposition p conditional on proposition q, and by $P(p,q)$ the joint probability of p and q. To prove the theorem it is enough to show that the probability of obtaining value 1 when measuring P_{u_m} conditional on obtaining value b_m when measuring $B(2)$, is equal to unity. But

$$\begin{aligned} P[P_{u_m} \to 1 | B(2) \to b_m] &\equiv \frac{P[P_{u_m} \to 1, B(2) \to b_m]}{P[B(2) \to b_m]} \\ &= \frac{\langle\chi| P_{u_m} \otimes P_{b_m} |\chi\rangle}{\langle\chi| I \otimes P_{b_m} |\chi\rangle} = \frac{|c_m|^2}{|c_m|^2} = 1 \,. \end{aligned}$$

♣

12.5 Application to Statistical Physics

The formalism of Sect. 12.4 finds an important application in statistical physics. Subsystem (2) is then associated with the statistical system investigated, which is macroscopic and thus made of a large number of particles. Subsystem (1) consists in the environment, that is, the rest of the world. In this section we assume that terms $H(1)$, $H(2)$ and $H_I(1,2)$ are all time independent.

Let $|\chi\rangle$ be the total system's state vector, given in (12.26). We pick the $|v_k(2)\rangle$ as eigenvectors of $H(2)$, i.e. $|v_k(2)\rangle = |E_k(2)\rangle$. It is not difficult to show (Exercise 12.10) that the coefficients $\sum_i c_{ik}^*(t) c_{il}(t)$ do not depend on the choice of orthonormal basis vectors $|u_i(1)\rangle$. Without loss of generality, we can thus pick the $|u_i(1)\rangle$ as eigenvectors of $H(1)$.

Clearly, if there were no interaction term $H_I(1,2)$ in the Hamiltonian one would have

$$c_{ik}(t) = c_{ik}(0) \exp\left\{ -\frac{it}{\hbar} \left[E_i(1) + E_k(2) \right] \right\}, \tag{12.35}$$

hence

$$\sum_i c_{ik}^* c_{il} \sim \exp\left\{ -\frac{it}{\hbar} \left[E_l(2) - E_k(2) \right] \right\}. \tag{12.36}$$

In general $H_I(1,2)$ does not vanish, but its mean value in a typical state $|\chi\rangle$ is usually much smaller than the mean value of $H(2)$. Thus one can think that (12.36), with small corrections, remains valid in situations where $H_I(1,2)$ does not vanish.

Energies $E_k(2)$ and $E_l(2)$ are macroscopic. The argument of the exponential in (12.36) changes very rapidly with time, unless E_k and E_l are virtually equal. Let us postulate that the average of $\sum_i c_{ik}^*(t) c_{il}(t)$, taken on a time interval large on a molecular scale and small on a macroscopic scale, vanishes if $k \neq l$. Operator $\varrho(2)$ then is effectively given by

$$\varrho(2) = \sum_l w_l |E_l(2)\rangle\langle E_l(2)|, \tag{12.37}$$

where

$$w_l = \sum_i |c_{il}(t)|^2. \tag{12.38}$$

The choice of the w_l completely determines the statistics of subsystem (2). At equilibrium the w_l do not depend on time. We will briefly examine two possible choices of w_l, corresponding to the so-called microcanonical and canonical distributions.

First consider the case where subsystem (2), after interacting with the environment, becomes completely isolated. Then its energy is very well defined, so that one can let

$$\bar{E} < E(2) < \bar{E} + \varDelta, \tag{12.39}$$

where $\varDelta \ll \bar{E}$. The *microcanonical distribution* consists in assigning the same nonzero value to every w_l for which $E_l(2)$ satisfies (12.39), and zero to all others. Explicitly

$$w_l = \begin{cases} 0 & \text{if } E_l(2) < \bar{E}\,, \\ \left[N(\bar{E})\right]^{-1} & \text{if } \bar{E} < E_l(2) < \bar{E} + \varDelta\,, \\ 0 & \text{if } E_l(2) > \bar{E} + \varDelta\,. \end{cases} \tag{12.40}$$

Obviously $N(\bar{E})$ is the number of quantum states with energy between $\bar{E}$ and $\bar{E} + \varDelta$.

Consider next the case where subsystem (2) is immersed in a constant-temperature heat reservoir. One can show that the w_l are then given by the *canonical distribution*, that is,

$$w_l = Z^{-1} \exp\left(-\frac{E_l(2)}{kT}\right), \tag{12.41}$$

where k is Boltzmann's constant. The coefficient Z, called the *partition function*, is determined by the condition $\mathrm{Tr}\{\varrho(2)\} = 1$, so that

$$Z = \sum_l \exp\left(-\frac{E_l(2)}{kT}\right). \tag{12.42}$$

To establish the connection with macroscopic variables, one uses the definition of *entropy* in terms of density operator coefficients ([140], Sect. 7):

$$S = -k \sum_l w_l \ln w_l\,. \tag{12.43}$$

For the microcanonical distribution one gets

$$S = -k \sum \frac{1}{N(\bar{E})} \ln\left(\frac{1}{N(\bar{E})}\right) = -k \ln\left(\frac{1}{N(\bar{E})}\right) = k \ln N(\bar{E}). \tag{12.44}$$

For the canonical distribution one finds

$$\begin{aligned} S &= -k \sum_l w_l \ln\left[\frac{1}{Z}\exp\left(-\frac{E_l(2)}{kT}\right)\right] \\ &= -k \sum_l w_l \left(-\ln Z - \frac{E_l(2)}{kT}\right) \\ &= k \ln Z + \frac{1}{T}\langle H(2)\rangle\,. \end{aligned} \tag{12.45}$$

Thus one can set

$$F = -kT \ln Z\,, \tag{12.46}$$

where F is the *Helmholtz free energy.*

Applications of statistical physics are found everywhere in modern physics. Let us just point out that the problems of blackbody radiation and specific heat of solids, mentioned in Chap. 1, are solved in that context. The interested reader can refer to a number of books in this area, in particular [140], Sects. 63–66.

In closing this section we come back to the density operator in the canonical distribution, and note that one can write

$$\varrho = Z^{-1} \sum_l \exp\left(-\frac{E_l}{kT}\right) |E_l\rangle\langle E_l| = Z^{-1} \exp\left(-\frac{H}{kT}\right). \tag{12.47}$$

Recall that if the Hamiltonian does not depend on time, the evolution operator is given by

$$U(t, t_0) = U(\Delta t) = \exp\left(-\frac{\mathrm{i}}{\hbar} H \Delta t\right). \tag{12.48}$$

Comparing the last two expressions we see that

$$Z\varrho(kT) = U\left(-\frac{\mathrm{i}\hbar}{kT}\right). \tag{12.49}$$

Up to a multiplicative factor, the density operator in the canonical distribution thus coincides with the evolution operator at an imaginary time.

12.6 Consistent Histories

The density operator characterizes a quantum state in a completely general way. Based on it we will introduce several concepts that are very useful in questions of interpretation of quantum mechanics.

Recall that a property of a quantum system is associated with a projector. Indeed a property consists in specifying that the value of a dynamical variable A is equal to a given real number or, more generally, that it belongs to a set of values $\mathcal{E}$. This uniquely defines the projector $P_{\mathcal{E}}$ on the subspace associated with eigenvalues of A belonging to $\mathcal{E}$.

Let $\varrho(t)$ be the density operator of an isolated quantum system. By definition, the probability at time t of the property associated with projector $P_{\mathcal{E}}$ is equal to the probability that a measurement of A at time t yields a result included in $\mathcal{E}$, that is,

$$P\left(P_{\mathcal{E}}\right) = \mathrm{Tr}\left\{\varrho(t) P_{\mathcal{E}}\right\}. \tag{12.50}$$

Because the system is isolated, the density operator satisfies (12.9). Letting $\varrho_0 = \varrho(t_0)$ and making use of the cyclic invariance of the trace, one can rewrite (12.50) as

$$P\left(P_{\mathcal{E}}\right) = \mathrm{Tr}\left\{P_{\mathcal{E}} U(t, t_0) \varrho_0 U^{\dagger}(t, t_0) P_{\mathcal{E}}\right\}. \tag{12.51}$$

The probability at time t of property $P_{\mathcal{E}}$ is thus expressed in terms of the density operator at an arbitrary initial time t_0.

We can now introduce the concept of history. Let t_1, t_2, $\ldots$, t_N be a sequence of instants such that $t_1 < t_2 < \cdots < t_N$. To each time t_i we associate a property represented by a projector P_i. A *history* is defined as the sequence P_1, P_2, …, P_N. In other words, a history is a sequence of properties associated with chronologically ordered instants.

A property is a specific case of history, namely a history involving one instant only. Generalizing from formula (12.51), we will take the probability of a history $h = (P_1, P_2, \ldots, P_N)$ to be equal to

$$\begin{aligned} P(h) = \mathrm{Tr}\,\{ & P_N U(t_N, t_{N-1}) P_{N-1} U(t_{N-1}, t_{N-2}) \ldots P_1 U(t_1, t_0) \varrho_0 \\ & \cdot U^\dagger(t_1, t_0) P_1 \ldots U^\dagger(t_{N-1}, t_{N-2}) P_{N-1} U^\dagger(t_N, t_{N-1}) P_N \} . \end{aligned} \tag{12.52}$$

Here ϱ_0 is the quantum system's density operator at time $t_0 \leq t_1$.

We will soon see that definition (12.52) cannot be maintained consistently in all generality. In other words, one can associate a probability with certain histories only. Nevertheless one can show under rather general hypotheses that this definition is the only one possible ([173], Chap. 4). Note (Exercise 12.11) that for any history h, $P(h) \geq 0$, and that $P(h) = 1$ if all projectors coincide with the identity.

We now introduce the notion of complete family of histories. For this purpose consider $N + 1$ instants $t_0 \leq t_1 < \cdots < t_N$ and an initial state $\varrho_0 = \varrho(t_0)$. At every time t_i ($1 \leq i \leq N$), let us write the identity operator as

$$I = \sum_{\alpha_i} P_i^{\alpha_i} , \tag{12.53}$$

where the $P_i^{\alpha_i}$ are projectors. Here the exponent α_i labels different projectors associated with time t_i. These in general differ from projectors associated with time t_j. A *complete family of histories* is then given by all histories like $h = (P_1^{\alpha_1}, P_2^{\alpha_2}, \ldots, P_N^{\alpha_N})$, where the α_i independently assume all their possible values.[5]

As an example we consider a spin 1/2 particle. We associate the dynamical variable S_x with time t_1 and the variable S_z with time t_2. Let

$$\begin{aligned} P_1^1 &= |+; \hat{x}\rangle\langle +; \hat{x}| , & P_2^1 &= |+; \hat{z}\rangle\langle +; \hat{z}| , \\ P_1^2 &= |-; \hat{x}\rangle\langle -; \hat{x}| , & P_2^2 &= |-; \hat{z}\rangle\langle -; \hat{z}| . \end{aligned} \tag{12.54}$$

Then (P_1^1, P_2^1), (P_1^1, P_2^2), (P_1^2, P_2^1) and (P_1^2, P_2^2) make up a complete family of histories. Note that this example is highly restrictive. In general projectors involve subspaces of dimension higher than 1, and the range of indices α_i and α_j is not necessarily the same.

Consider a complete family of histories $\{(P_1^{\alpha_1}, P_2^{\alpha_2}, \ldots, P_N^{\alpha_N})\}$. One can define an *enlarged* family by adding all histories like

$$(\bar{P}_1, \bar{P}_2, \ldots, \bar{P}_N) = \left(\sum_{\alpha_1} P_1^{\alpha_1}, \sum_{\alpha_2} P_2^{\alpha_2}, \ldots, \sum_{\alpha_N} P_N^{\alpha_N} \right) , \tag{12.55}$$

---

[5] This definition, although not the most general one, is adequate for our purposes.

where $\sum_{\alpha_i}$ denotes a partial sum of projectors. In the previous example the enlarged family includes, in addition to members of the complete family, histories (I,P_2^1), (I,P_2^2), (P_1^1,I), (P_1^2,I) and (I,I).

Let $h' = (\bar{P}_1,\ldots,\bar{P}_i',\ldots,\bar{P}_N)$ and $h'' = (\bar{P}_1,\ldots,\bar{P}_i'',\ldots,\bar{P}_N)$ be two members of the enlarged family. Then $h = (\bar{P}_1,\ldots,\bar{P}_i'+\bar{P}_i'',\ldots,\bar{P}_N)$ is also a member of the enlarged family. The complete family of histories is *consistent* if for any h' and h''

$$P(h) = P(h') + P(h''). \tag{12.56}$$

These conditions, due to Griffiths ([102], [104]), are very restrictive. They determine families of histories on which one can define additive probabilities. They must hold if one is to endow histories with some kind of reality.

To keep on with the example of a spin 1/2 particle, suppose that the evolution operators $U(t_2,t_1)$ and $U(t_1,t_0)$ do not act on the spin. In other words they commute with the projectors and density operator. One then shows that the family of histories is consistent if $\varrho_0 = |+;\hat{x}\rangle\langle+;\hat{x}|$, and is not consistent if $\varrho_0 = |+;\hat{z}\rangle\langle+;\hat{z}|$. More generally, the following is shown in Exercise (12.13). Suppose that projectors at times t_1 and t_2 correspond to spins along the $\hat{n}_1$ and $\hat{n}_2$ axes respectively, and that the initial state corresponds to a spin along the $\hat{n}_0$ axis. Then the family of histories is consistent if and only if $(\hat{n}_2 \times \hat{n}_1)\cdot(\hat{n}_0 \times \hat{n}_1) = 0$.

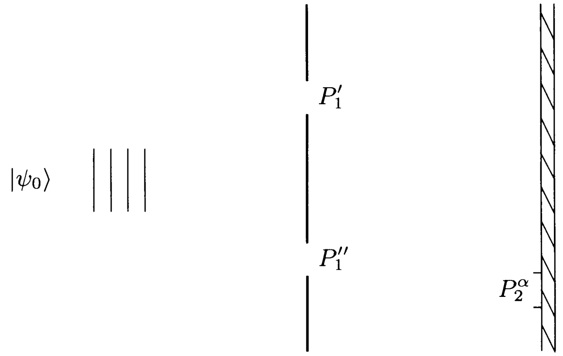

Fig. 12.1. Two-slit interference setup

To understand better the physical meaning of the notions introduced above, we will examine the interference phenomenon shown in Fig. 12.1. Particles prepared in state $|\psi_0\rangle$ move toward a screen where they arrive in state $U(t_1,t_0)|\psi_0\rangle$. The screen has two slits in it. Let us first suppose that the lower slit is obstructed, and let P_1' be the projector associated with the upper slit. Particles emerge from that slit in state $P_1'U(t_1,t_0)|\psi_0\rangle$. To the right of the slits is a second screen equipped with detectors that correspond to projectors P_2^α (the exponent α being associated with a detector's vertical position). Particles arrive at the second screen in state $U(t_2, t_1)P_1'U(t_1, t_0)|\psi_0\rangle$. The final state of a particle reaching detector α is therefore given by

$$|\psi_f'\rangle = P_2^\alpha U(t_2, t_1)P_1'U(t_1, t_0)|\psi_0\rangle. \tag{12.57}$$

If, on the other hand, the upper slit is obstructed, the final state is given by

$$|\psi_f''\rangle = P_2^\alpha U(t_2, t_1)P_1''U(t_1, t_0)|\psi_0\rangle, \tag{12.58}$$

where P_1'' is the projector associated with the lower slit.

When no slit is obstructed, the projector associated with the first screen is equal to $P_1' + P_1''$. The final state is then given by

$$|\psi_f\rangle = P_2^\alpha U(t_2, t_1)\,[P_1' + P_1'']\,U(t_1, t_0)|\psi_0\rangle = |\psi_f'\rangle + |\psi_f''\rangle, \tag{12.59}$$

where the linearity of projectors and evolution operators was used. Equation (12.59) expresses the property of superposition of amplitudes at the root of interference phenomena.

Equation (12.59) is true for any value of α. Hence the norm of $|\psi_f\rangle$ will not, in general, be equal to the sum of the norms of $|\psi_f'\rangle$ and $|\psi_f''\rangle$. In other words,

$$\langle\psi_f|\psi_f\rangle \neq \langle\psi_f'|\psi_f'\rangle + \langle\psi_f''|\psi_f''\rangle. \tag{12.60}$$

It is easy to see that $\langle\psi_f|\psi_f\rangle$ is equal to the probability of the history associated with projectors $P_1' + P_1''$ and P_2^α and density operator $|\psi_0\rangle\langle\psi_0|$. A similar interpretation holds for $\langle\psi_f'|\psi_f'\rangle$ and $\langle\psi_f''|\psi_f''\rangle$. Because of (12.60), (12.56) is not obeyed. That is, the family of histories associated with the interference setup is not consistent. Thus one cannot claim that a particle reaching the second screen went through one and only one slit.

Let us now resume the general argument and consider a complete family of consistent histories. It is possible to associate a *logical structure* with the family. For this purpose let us first associate with each history $(P_1^{\alpha_1}, P_2^{\alpha_2}, \ldots, P_N^{\alpha_N})$ the proposition that consists in asserting property $P_1^{\alpha_1}$ at time t_1, property $P_2^{\alpha_2}$ at time t_2, and so on. From these basic propositions, introduce next more complex propositions by means of logical connectives $\wedge$ (and), $\vee$ (or) and $\neg$ (not), used in a natural way. In fact each complex proposition can be obtained by a disjunction of basic propositions, that is, as $p_1 \vee p_2 \vee \ldots \vee p_n$.

To every proposition one can associate a *probability* in the following way. The probability of a basic proposition is defined as the probability of the corresponding history, and the probability of a complex proposition is defined as the sum of probabilities of basic propositions of which it is the disjunction.

We can now define logical implication. A proposition p *implies* a proposition q if (i) the probability of p does not vanish and (ii) the probability of $p \wedge q$ is equal to the probability of p.

One can show ([173], Chap. 5) that the structure so defined obeys all the axioms of logic, provided that the family of histories on which it is based is consistent. Note, however, that we have not yet given a truth criterion for propositions. We shall come back to this in Chap. 21.

At any rate we can now state the following rule for the interpretation of the formalism of quantum mechanics ([173], p. 163).

Interpretation Rule The description of an isolated physical system must be made by means of propositions belonging to a logic based on a complete family of consistent histories. Any reasoning about the system's properties must be the result of valid implications in this logic.

Applied to the interference setup analyzed above, this rule entails the conclusion we then reached. Together with a more restrictive specification of consistent families of histories, it may also entail the interpretation rules that pertain to the measurement of a dynamical variable.

Exercises

12.1.

a) Show that if $\varrho(t)$ satisfies (12.8) and $\varrho(t) = |\psi(t)\rangle\langle\psi(t)|$, then $|\psi(t)\rangle$ satisfies the equation

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = \{H - \lambda(t)\}\,|\psi(t)\rangle\,,$$

where $\lambda(t)$ is a real function of t.

b) Show that if $|\psi(t)\rangle$ satisfies the previous equation, then

$$|\psi(t)\rangle = \exp\left\{\frac{\mathrm{i}}{\hbar}\int^{t}\mathrm{d}t'\,\lambda(t')\right\}|\bar{\psi}(t)\rangle\,,$$

where $|\bar{\psi}(t)\rangle$ satisfies Schrödinger's equation for Hamiltonian H. From this derive that the function $\lambda(t)$ has no physical significance, and can be suppressed.

12.2. Let $\{|\psi_n(t)\rangle\}$ be a set of solutions of the Schrödinger equation associated with a Hamiltonian H, and let

$$\varrho(t) = \sum_n w_n|\psi_n(t)\rangle\langle\psi_n(t)|$$

be the density operator of a statistical mixture. What is the time evolution of $\varrho(t)$ in the case where the w_n depend on time? Show that, in the case where the w_n do not depend on time, the evolution equation reduces to (12.8).

12.3. Let $\varrho(t)$ be the density operator given in Exercise (12.2), where each w_n is constant. Show that the trace of ϱ^2 is constant.

12.4. Let $|+;\hat{z}\rangle \equiv |+\frac{1}{2};\hat{z}\rangle$ and $|-;\hat{z}\rangle \equiv |-\frac{1}{2};\hat{z}\rangle$ be state vectors of a spin 1/2 particle. In some basis

$$|+;\hat{z}\rangle = \begin{pmatrix}1\\0\end{pmatrix}, \qquad |-;\hat{z}\rangle = \begin{pmatrix}0\\1\end{pmatrix}.$$

Let $\hat{u}$ be an axis in the xz plane that makes an angle θ with the z axis. In Chap. 4 we found the components of $|+;\hat{u}\rangle$ and $|-;\hat{u}\rangle$ as

$$U\begin{pmatrix}1\\0\end{pmatrix} \quad \text{and} \quad U\begin{pmatrix}0\\1\end{pmatrix},$$

where (4.61)

$$U = \begin{pmatrix}\cos(\theta/2) & -\sin(\theta/2)\\ \sin(\theta/2) & \cos(\theta/2)\end{pmatrix}.$$

a) Consider an ensemble of particles half with their spin along the positive z axis direction and half along the negative direction. Write the matrix elements of the density operator in the above basis.
b) Answer the same question if half of the particles have their spin along the positive, and half along the negative $\hat{u}$ axis direction. Interpret the result.

12.5. In the basis specified in Exercise (12.4), the Hermitian operators associated with the z and y spin components of a spin 1/2 particle are proportional to

$$\sigma_z = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix} \quad \text{and} \quad \sigma_y = \begin{pmatrix}0 & -\mathrm{i}\\ \mathrm{i} & 0\end{pmatrix}.$$

The vector $|\phi\rangle$ given by

$$|\phi\rangle = \frac{1}{\sqrt{2}}\left\{|+;\hat{z}\rangle + \mathrm{i}|-;\hat{z}\rangle\right\},$$

represents a state where the y spin component is equal to $+\hbar/2$.
a) What is, in matrix notation, the density operator ϱ_1 associated with the pure state $|\phi\rangle$?
b) Compute the mean values of σ_y and σ_z in the two situations described by ϱ_1 and ϱ_2, where ϱ_2 is the density operator found in Exercise (12.4a).

12.6. Show that if $|\phi(1)\rangle$ and $|\psi(2)\rangle$ both depend on a parameter t, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left\{|\phi(1)\rangle \otimes |\psi(2)\rangle\right\} = \left\{\frac{\mathrm{d}}{\mathrm{d}t}|\phi(1)\rangle\right\}\otimes|\psi(2)\rangle + |\phi(1)\rangle\otimes\left\{\frac{\mathrm{d}}{\mathrm{d}t}|\psi(2)\rangle\right\}.$$

12.7. A quantum system is made of two noninteracting subsystems. Show that the total system's Schrödinger equation implies that each subsystem itself obeys Schrödinger's equation.

12.8. Let $|\chi(1,2)\rangle$ be a total system's state vector and let $\varrho(2)$ be the density operator of the second subsystem in the presence of the first. Show that if

$$H(1,2) = H(1) \otimes I(2) + I(1) \otimes H(2),$$

then the evolution equation of $\varrho(2)$ is given by

$$\mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t}\varrho(2) = [H(2), \varrho(2)] .$$

Here we assume that $H(1)$ and $H(2)$ are time independent.

12.9. Let $\hat{u}$ be a unit vector. The Hermitian operator associated with the $\hat{u}$ spin component of a spin 1/2 particle is proportional to

$$\boldsymbol{\sigma} \cdot \hat{u} = \begin{pmatrix} u_z & u_x - \mathrm{i}u_y \\ u_x + \mathrm{i}u_y & -u_z \end{pmatrix} .$$

a) Consider a system of two spin 1/2 particles in a state $|\chi\rangle$ given by

$$|\chi(1,2)\rangle = \frac{1}{\sqrt{2}} \{|+;\hat{z}\rangle \otimes |-;\hat{z}\rangle - |-;\hat{z}\rangle \otimes |+;\hat{z}\rangle\} .$$

Show that the mean value of operator $\boldsymbol{\sigma} \cdot \hat{u} \otimes \boldsymbol{\sigma} \cdot \hat{v}$ in state $|\chi\rangle$ is equal to $-\hat{u} \cdot \hat{v}$.

b) Deduce that if a measurement of the first particle's spin along z yields value $+\hbar/2$, then a measurement of the second particle's spin along z will necessarily yield value $-\hbar/2$.

12.10. Let $|\chi\rangle$ be defined as in (12.26). Show that

$$\sum_i c_{ik}^*(t) c_{il}(t) = \langle\chi| \left[\left(\sum_i |u_i\rangle\langle u_i| \right) \otimes |v_k\rangle\langle v_l| \right] |\chi\rangle$$

and thus that the left-hand side does not depend on the choice of orthonormal basis vectors $|u_i\rangle$.

12.11. Let $P(h)$ be given by (12.52).

a) Write $P(h)$ as $\mathrm{Tr}\{A\varrho_0 A^\dagger\}$, and show that $P(h) \geq 0$.

b) Show that if $P_1 = P_2 = \cdots = P_N = I$, then $P(h) = 1$.

12.12. Consider a complete family of histories involving two instants t_1 and t_2.

a) Show that condition (12.56) is always met if $h' = (\bar{P}_1, \bar{P}_2')$ and $h'' = (\bar{P}_1, \bar{P}_2'')$.

b) Show that condition (12.56) becomes, if $h' = (\bar{P}_1', \bar{P}_2)$ and $h'' = (\bar{P}_1'', \bar{P}_2)$,

$$\mathrm{Re}\left[\mathrm{Tr}\left\{\bar{P}_2 U(t_2,t_1) \bar{P}_1' U(t_1,t_0) \varrho_0 U^\dagger(t_1,t_0) \bar{P}_1'' U^\dagger(t_2,t_1)\right\}\right] = 0 .$$

12.13. Consider a spin 1/2 particle and a family of histories associated with instants t_1 and t_2. Suppose that the projectors at t_1 and t_2 correspond to spins along the $\hat{n}_1$ and $\hat{n}_2$ axes respectively. Let $\varrho_0 = |+;\hat{n}_0\rangle\langle+;\hat{n}_0|$ and assume evolution operators commute with ϱ_0 and projectors. Use results of Exercise (12.12) to show that the family of histories is consistent if and only if

$$(\hat{n}_2 \times \hat{n}_1) \cdot (\hat{n}_0 \times \hat{n}_1) \equiv \hat{n}_2 \cdot \hat{n}_0 - (\hat{n}_2 \cdot \hat{n}_1)(\hat{n}_0 \cdot \hat{n}_1) = 0.$$

13 Symmetry of the Hamiltonian

The notion of symmetry is familiar from daily experience. We say that an object displays a symmetry if it is invariant under a transformation. This means that after the transformation, the object's configuration is identical with the one it had before the transformation. Thus a sphere is symmetric because it is invariant under rotations. In quantum mechanics the concept of symmetry is particularly important. The importance comes from the nature of the formalism and from the fact that microobjects are simpler, i.e. have less structure, than human-size objects. From a mathematical point of view, the notion of symmetry is intimately related to the algebraic structure called a group. We shall see that knowledge of a quantum system's symmetry group reveals a number of the system's properties, without its Hamiltonian being completely known. These properties are shared by all quantum systems whose Hamiltonian has the same symmetry group.[1]

13.1 Symmetry Group

Let H be the Hamiltonian of a quantum system. The time evolution of the state vector $|\psi(t)\rangle$ is governed by the Schrödinger equation:

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi\rangle = H|\psi\rangle \,. \tag{13.1}$$

Let U be an invertible time-independent operator. Suppose U transforms vectors in the state space as

$$|\psi'\rangle = U|\psi\rangle \,. \tag{13.2}$$

We will investigate the conditions under which the quantum system's description by means of vectors $|\psi'\rangle$ is identical with the description by means of vectors $|\psi\rangle$.

In Sect. 3.2 we have shown that both descriptions are equivalent if the operator U is unitary. To each dynamical variable A in the original description

---

[1] Group theory and its applications in quantum mechanics are treated in many books. See for instance [19], [57], [90], [100], [105], [113], [183], [226], [234] and [238].

L. Marchildon, *Quantum Mechanics*

there corresponds then a variable $A' = UAU^{-1}$ in the primed description, where A' is Hermitian and has the same spectrum as A. The Schrödinger equation becomes

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi'\rangle = H'|\psi'\rangle\,, \tag{13.3}$$

where

$$H' = UHU^{-1}\,. \tag{13.4}$$

For descriptions in terms of $|\psi\rangle$ and $|\psi'\rangle$ to be *identical*, they must first be equivalent, that is, U must be unitary. But in addition (13.1) and (13.3) must coincide, i.e. the Schrödinger equation must be exactly the same in both descriptions. Such will be the case if $H' = H$ which, given (13.4), is equivalent to

$$[H, U] = 0\,. \tag{13.5}$$

By definition, a *symmetry transformation* is the effect of a unitary operator that commutes with the Hamiltonian. The term also applies to the unitary operator itself. A Hamiltonian is *invariant* under a transformation if it commutes with the corresponding unitary operator. Two descriptions of a quantum system are identical if they are related by a symmetry transformation.[2]

The notion of symmetry just introduced coincides with the one we have intuitively. The Hamiltonian governs the whole dynamics of a quantum system. If a transformation, represented by operator U in (13.2), is made on a system, and if the transformed system's Hamiltonian coincides with the original system's, then both systems have the same dynamics. Hence they cannot be distinguished.

Consider the set $\mathcal{U}$ of all symmetry transformations of a Hamiltonian H. This set can be finite or infinite. Its elements are unitary operators that commute with H. The set $\mathcal{U}$ always has the following properties:

---

[2] The notion of equivalent description (like that of identical description) can be enlarged by forsaking the linearity hypothesis on operator U. Let V be a mapping of the state space on itself such that $|\psi\rangle \to |\psi'\rangle = V|\psi\rangle$. Descriptions by means of vectors $|\psi\rangle$ and $|\psi'\rangle$ are called *equivalent* if for every $|\psi\rangle$ and $|\phi\rangle$, $|\langle\phi'|\psi'\rangle| = |\langle\phi|\psi\rangle|$. One can show ([218], App. D; [238], Chap. 20) that descriptions are equivalent if and only if V is unitary or antiunitary. An *antiunitary* operator is *antilinear*, which means that

$$V\left\{\alpha_1|\phi_1\rangle + \alpha_2|\phi_2\rangle\right\} = \alpha_1^* V\left\{|\phi_1\rangle\right\} + \alpha_2^* V\left\{|\phi_2\rangle\right\}.$$

Moreover

$$(V|\phi\rangle, V|\psi\rangle) = (|\psi\rangle, |\phi\rangle) = \langle\phi|\psi\rangle^*\,.$$

In a given basis an antiunitary operator can always be written as the product of a unitary operator and the complex conjugation operator.

i) $\mathcal{U}$ is closed under the product of operators, i.e. the product of two symmetry transformations is a symmetry transformation. Indeed U_1U_2 is unitary if U_1 and U_2 are. Furthermore if U_1 and U_2 commute with H then

$$[U_1U_2, H] = U_1 [U_2, H] + [U_1, H] U_2 = 0 .$$

ii) The product of operators belonging to $\mathcal{U}$ is associative. Actually the product of linear operators is always associative. Note that associativity was used in the proof of property (i).

iii) The idendity operator belongs to $\mathcal{U}$.

iv) $\mathcal{U}$ contains inverses of all its elements. Indeed U^{-1} is unitary if U is. If U and H commute then $H = UHU^{-1}$, and

$$\left[U^{-1}, H\right] = U^{-1}H - HU^{-1} = U^{-1}\left\{UHU^{-1}\right\} - HU^{-1} = 0 .$$

In algebra a *group* is a set $\mathcal{U}$ and a binary product law, defined on elements of $\mathcal{U}$, that satisfy the four properties just stated.[3] We have thus shown that a Hamiltonian's symmetry transformations always form a group.

From an algebraic point of view, a group is completely defined in terms of its elements' multiplication table. The nature of its elements does not matter. Sets of widely different objects can thus make up the same abstract group.

Let G be a group and let K be a subset of G. The set K is called a *subgroup* of G if K is itself a group under the product law defined in G. In this case K is closed under the product law, and it contains the identity and the inverse of each of its elements. Any set of symmetry transformations of a Hamiltonian which is a group with respect to the operator product is called a *symmetry group* of the Hamiltonian.

13.2 Space Inversion, Rotations and Translations

We now define three kinds of very important operators that act in the state space of a particle in three dimensions (or, more generally, in the state space of several particles in three dimensions). These operators are space inversion (also called parity), rotations and translations. They will be defined independently of any reference to a Hamiltonian. Hence they may or may not be symmetry transformations of a Hamiltonian.

i) The *space inversion* (or *parity*) operator, denoted by P, is defined by the following transformation, for every ket $|\boldsymbol{r}\rangle$:

$$|\boldsymbol{r}'\rangle \equiv |-\boldsymbol{r}\rangle = P|\boldsymbol{r}\rangle . \tag{13.6}$$

---

[3] Strictly speaking, the group is made up of the set *and* product law. When the latter is implicit, the group is often identified with the set.

ii) Let $\boldsymbol{r}' = O(\hat{n};\theta)\boldsymbol{r}$ be the three-dimensional vector obtained by rotating $\boldsymbol{r}$ by an angle θ about the $\hat{n}$ axis. The *rotation* operator $\mathcal{R}(\hat{n};\theta)$ is then defined by the following transformation:

$$|\boldsymbol{r}'\rangle \equiv |O(\hat{n};\theta)\boldsymbol{r}\rangle = \mathcal{R}(\hat{n};\theta)|\boldsymbol{r}\rangle\,. \tag{13.7}$$

One should not confuse $\mathcal{R}(\hat{n};\theta)$, acting in an infinite-dimensional space, with $O(\hat{n};\theta)$, acting in a three-dimensional space.

iii) Let $\boldsymbol{r}' = \boldsymbol{r} + \boldsymbol{r}_0$, where $\boldsymbol{r}_0$ is a constant vector. The *translation* operator $\mathcal{T}(\boldsymbol{r}_0)$ is defined by the following transformation:

$$|\boldsymbol{r}'\rangle \equiv |\boldsymbol{r}+\boldsymbol{r}_0\rangle = \mathcal{T}(\boldsymbol{r}_0)|\boldsymbol{r}\rangle\,. \tag{13.8}$$

It is easy to see that all operators defined in (i), (ii) and (iii) are unitary. Indeed they transform into itself the orthonormal basis of kets $|\boldsymbol{r}\rangle$. The space-inversion operator and the identity make up a group of two elements. The set of all rotations and the set of all translations each make up an infinite group.

We now examine the case where space inversion, rotations or translations are symmetry transformations of a Hamiltonian. Let U be one of the operators defined in (i), (ii) or (iii). We know that U is unitary and that it transforms kets $|\boldsymbol{r}\rangle$ as

$$|\boldsymbol{r}'\rangle = U|\boldsymbol{r}\rangle\,.$$

If U is a symmetry transformation,

$$\langle\boldsymbol{r}_1'|H|\boldsymbol{r}_2'\rangle = \langle\boldsymbol{r}_1|U^\dagger HU|\boldsymbol{r}_2\rangle = \langle\boldsymbol{r}_1|U^{-1}HU|\boldsymbol{r}_2\rangle = \langle\boldsymbol{r}_1|H|\boldsymbol{r}_2\rangle. \tag{13.9}$$

Thus the $(\boldsymbol{r}_1\boldsymbol{r}_2)$ matrix element of H is equal to the $(\boldsymbol{r}_1'\boldsymbol{r}_2')$ matrix element, kets $|\boldsymbol{r}_1'\rangle$ and $|\boldsymbol{r}_2'\rangle$ being obtained from kets $|\boldsymbol{r}_1\rangle$ and $|\boldsymbol{r}_2\rangle$ by the symmetry transformation.

Let us focus on a physical system made of a particle in a potential $V(\boldsymbol{r})$. Matrix elements of H in the coordinate representation are then given by (5.16):

$$\langle\boldsymbol{r}_1|H|\boldsymbol{r}_2\rangle = \left\{-\frac{\hbar^2}{2m}\nabla^2_{\boldsymbol{r}_1} + V(\boldsymbol{r}_1)\right\}\delta(\boldsymbol{r}_1 - \boldsymbol{r}_2), \tag{13.10}$$

$$\langle\boldsymbol{r}_1'|H|\boldsymbol{r}_2'\rangle = \left\{-\frac{\hbar^2}{2m}\nabla^2_{\boldsymbol{r}_1'} + V(\boldsymbol{r}_1')\right\}\delta(\boldsymbol{r}_1' - \boldsymbol{r}_2')\,. \tag{13.11}$$

It is not difficult to show (Exercise 13.1) that the Laplacian and Dirac delta function are invariant under space inversion, rotations and translations. Hence for all these transformations

$$\langle\boldsymbol{r}_1'|H|\boldsymbol{r}_2'\rangle = \left\{-\frac{\hbar^2}{2m}\nabla^2_{\boldsymbol{r}_1} + V(\boldsymbol{r}_1')\right\}\delta(\boldsymbol{r}_1-\boldsymbol{r}_2). \tag{13.12}$$

Comparing (13.10) and (13.12) one sees that (13.9) is equivalent to

$$V(\boldsymbol{r}_1') = V(\boldsymbol{r}_1)\,. \tag{13.13}$$

Let us now examine each case specifically. From (13.13) and (13.6) one sees that invariance of H under space inversion means that

$$V(-\boldsymbol{r}) = V(\boldsymbol{r})\,. \tag{13.14}$$

Invariance of H under a rotation $(\hat{n};\theta)$ means that

$$V\left\{O(\hat{n};\theta)\boldsymbol{r}\right\} = V(\boldsymbol{r})\,, \tag{13.15}$$

where (13.7) was used. We know that any vector $\boldsymbol{r}'$ with the same length as a vector $\boldsymbol{r}$ can be obtained from $\boldsymbol{r}$ through a rotation. Hence H will be invariant under the whole group of rotations if and only if

$$V(\boldsymbol{r}) = V(r)\,, \tag{13.16}$$

that is, if the potential is spherically symmetric. Finally, invariance of H under a translation means that

$$V(\boldsymbol{r}+\boldsymbol{r}_0) = V(\boldsymbol{r})\,. \tag{13.17}$$

Thus H will be invariant under the whole group of translations if and only if $V(\boldsymbol{r})$ is independent of $\boldsymbol{r}$, that is, if the potential is a constant.

Space inversion, rotations and translations act on kets $|\boldsymbol{r}\rangle$ in a simple way. Other symmetry transformations of a given Hamiltonian may, however, act in more complicated ways.

13.3 General Properties of Groups

We now give a number of definitions about groups and outline some of their general properties.

Elements of a group G will normally be denoted by g, g_1, g_2, $\ldots$, no assumption being made here that their number be finite or even denumerable. The unit element (i.e. the identity) will be denoted by e (or I) while g^{-1} will denote the inverse of g. One easily sees that $(g_1g_2)^{-1} = g_2^{-1}g_1^{-1}$. Trivially, e and G are two subgroups of G.

Let G and $\bar{G}$ be two groups. We say that $\bar{G}$ is *homomorphic* to G if there exists between G and $\bar{G}$ a correspondence such that (i) to any element g of G there corresponds one and only one element $\bar{g}$ of $\bar{G}$; and (ii) to a product g_1g_2 of two elements of G there corresponds the product $\bar{g}_1\bar{g}_2$ of the two corresponding elements of $\bar{G}$. Note that the relation "homomorphic to" is not symmetric. A group $\bar{G}$ homomorphic to G is, from the point of view of the multiplication table, a subgroup of G.

If $\bar{G}$ is homomorphic to G and if the correspondence $g \to \bar{g}$ is one-to-one (i.e. to an element of $\bar{G}$ there corresponds one and only one element of G), then G and $\bar{G}$ are *isomorphic.* Isomorphism is a symmetric relation. From the point of view of their multiplication table, two isomorphic groups cannot be distinguished.

A group G is *Abelian* if the product of elements is commutative, i.e. if $g_1 g_2 = g_2 g_1$ for any g_1 and g_2 in G.

Let g be an element of G and K a subgroup of G. Denote by $\{gK\}$ the set of elements of the form gk where k belongs to K. The set $\{gK\}$ is called the *left coset of K with respect to g.* Clearly if g belongs to K, $\{gK\}$ then coincides with K. Take two left cosets $\{g_1 K\}$ and $\{g_2 K\}$. Either they coincide, or they are disjoint. Indeed suppose they are not disjoint. This means there exist a k_1 and a k_2 in K such that $g_1 k_1 = g_2 k_2$. But then $g_1 = g_2 k_2 k_1^{-1}$, so that

$$\{g_1 K\} = \left\{(g_2 k_2 k_1^{-1})K\right\} = \left\{g_2 \left\{k_2 k_1^{-1} K\right\}\right\} = \{g_2 K\} .$$

Thus elements of G can be distributed exhaustively in disjoint subsets $\{g_i K\}$ of G, elements of each of these subsets being in one-to-one correspondence with elements of K.

Similarly, one can divide G in *right cosets* $\{Kg\}$. A subgroup K of G is called *normal* if its left and right cosets coincide. Equivalently, K is a normal subgroup of G if for all g in G the set $\{gKg^{-1}\}$ coincides with K.

The set of left cosets is denoted by G/K. If K is a normal subgroup of G, one can endow G/K with a group structure by defining the product of two cosets as

$$\{g_1 K\} \cdot \{g_2 K\} = \{(g_1 g_2)K\} . \tag{13.18}$$

One is easily convinced that this product does not depend on elements g_1 and g_2 picked to generate cosets $\{g_1 K\}$ and $\{g_2 K\}$.

If G is a finite group the number of elements of G is called the *order* of G. The distribution of elements of G in cosets associated with a subgroup K implies that the order of a subgroup K is always a divisor of the order of G.

We say G is *simple* if it has no normal subgroups, while G is *semisimple* if it has no Abelian normal subgroups.

Let G and $\bar{G}$ be two groups. Consider the set of ordered pairs of elements of G and $\bar{G}$, denoted by $(g, \bar{g})$. A product can be defined on the set of all pairs in the following way:

$$(g_1, \bar{g}_1) \cdot (g_2, \bar{g}_2) \equiv (g_1 g_2, \bar{g}_1 \bar{g}_2) . \tag{13.19}$$

One easily checks that this product gives rise to a group structure. The group thus built is called the *direct product* of groups G and $\bar{G}$, and is denoted by $G \otimes \bar{G}$.[4] It is easy to show that G and $\bar{G}$ are two normal subgroups of $G \otimes \bar{G}$. Moreover G is isomorphic to $(G \otimes \bar{G})/\bar{G}$ and $\bar{G}$ is isomorphic to $(G \otimes \bar{G})/G$.

---

[4] Although they are denoted similarly, the direct product of groups and the tensor product of vector spaces are two different structures.

Let g_1 be an element of a group G. The *class* of g_1 is the set of all elements like gg_1g^{-1}, where g is an arbitrary element of G. One can check that the relation "g_2 belongs to the class of g_1" is an equivalence relation, that is, it is symmetric, reflexive and transitive. Thus elements of G can be exhaustively distributed in disjoint classes. This, in general, is different from a distribution in cosets.

13.4 Matrix Representation of a Group

A group of nonsingular $N \times N$ matrices that is homomorphic to a group G is called a *matrix representation* of G.[5] That is, a matrix representation of G is the association, to each element g of G, of a nonsingular $N \times N$ matrix $M(g)$ such that, for all g_1 and g_2 in G,

$$M(g_1)M(g_2) = M(g_1g_2). \tag{13.20}$$

The *dimension* of a matrix representation is the dimension of matrices in it. It can be infinite.

Let e be the unit element of the group. Letting $g_2 = e$ in (13.20) one finds that

$$M(g_1)M(e) = M(g_1). \tag{13.21}$$

Since $[M(g_1)]^{-1}$ exists, one immediately obtains that $M(e)$ is the identity matrix. Furthermore, letting $g_2 = g_1^{-1}$ in (13.20) yields

$$M(g_1)M(g_1^{-1}) = M(e), \tag{13.22}$$

so that for every g_1

$$M(g_1^{-1}) = [M(g_1)]^{-1} \equiv M^{-1}(g_1). \tag{13.23}$$

The matrix of the inverse is therefore the inverse of the matrix.

A matrix representation of G is *faithful* if the correspondence between elements of G and matrices is one-to-one. That is, a matrix representation of G is faithful if G and the group of matrices are isomorphic. A matrix representation is *unitary* if all matrices in it are unitary.

Let $\{M(g)\}$ and $\{\bar{M}(g)\}$ be two matrix representations of a group G with the same dimension. These representations are *equivalent* if there exists a matrix F such that for any g in G

$$\bar{M}(g) = F\,M(g)\,F^{-1}. \tag{13.24}$$

Two equivalent representations are isomorphic.

---

[5] More generally, a *representation* of G is a group of operators (with inverses) homomorphic to G.

Matrix representations, and in particular unitary representations, are very important in group theory. In quantum mechanics they are associated with spaces of eigenvectors corresponding to degenerate eigenvalues of the Hamiltonian. Let H be the Hamiltonian of a quantum system, $|E\rangle$ an eigenvector of H (with eigenvalue E) and let U be a symmetry transformation of the Hamiltonian. Then

$$H\{U|E\rangle\} = HU|E\rangle = UH|E\rangle = UE|E\rangle = E\{U|E\rangle\}. \tag{13.25}$$

Thus $U|E\rangle$ is also an eigenvector of H with eigenvalue E. We now prove the following theorem.

Theorem Let G be a group of symmetry transformations of a Hamiltonian H. Let $\mathcal{V}^N$ be the vector space (with finite dimension N) made up of all eigenvectors of H with eigenvalue E_i. Let $\{|E_i^\alpha\rangle,\ \alpha = 1, 2, \ldots, N\}$ be an orthonormal basis of $\mathcal{V}^N$. To each element U of G, associate a matrix $M(U)$, with dimension N, whose elements are given by

$$M_{\alpha\beta}(U) = \langle E_i^\alpha|U|E_i^\beta\rangle. \tag{13.26}$$

Then this correspondence is a unitary matrix representation of G.

Proof We must show that matrices $M(U)$ satisfy (13.20) and are unitary. If U_1 and U_2 belong to G, then

$$\begin{aligned} M_{\alpha\beta}(U_1U_2) &= \langle E_i^\alpha|U_1U_2|E_i^\beta\rangle \\ &= \sum_j \sum_\gamma \langle E_i^\alpha|U_1|E_j^\gamma\rangle\langle E_j^\gamma|U_2|E_i^\beta\rangle \\ &= \sum_\gamma \langle E_i^\alpha|U_1|E_i^\gamma\rangle\langle E_i^\gamma|U_2|E_i^\beta\rangle , \end{aligned}$$

the last equality following from the fact that $\langle E_i^\alpha|U_1|E_j^\gamma\rangle$ vanishes if $E_j \neq E_i$. Thus one gets

$$M_{\alpha\beta}(U_1U_2) = \sum_\gamma M_{\alpha\gamma}(U_1)M_{\gamma\beta}(U_2) = [M(U_1)M(U_2)]_{\alpha\beta} ,$$

so that (13.20) is satisfied. There remains to prove the unitarity of $M(U)$. We know that U is a unitary operator. Hence

$$\begin{aligned} \left[M^\dagger(U)M(U)\right]_{\alpha\beta} &= \sum_{\gamma=1}^N M_{\gamma\alpha}^*(U)M_{\gamma\beta}(U) \\ &= \sum_{\gamma=1}^N \langle E_i^\gamma|U|E_i^\alpha\rangle^* \langle E_i^\gamma|U|E_i^\beta\rangle \end{aligned}$$

$$= \sum_j \sum_\gamma \langle E_i^\alpha | U^\dagger | E_j^\gamma \rangle \langle E_j^\gamma | U | E_i^\beta \rangle$$

$$= \langle E_i^\alpha | U^\dagger U | E_i^\beta \rangle = \langle E_i^\alpha | E_i^\beta \rangle = \delta_{\alpha\beta} \, . \qquad \clubsuit$$

Unitary matrix representations of a group G are completely determined by the group's abstract properties, with no reference to any Hamiltonian whatsoever. The importance of the theorem comes from the fact that information can be obtained on the structure of spaces of eigenvectors of H, corresponding to a degenerate eigenvalue, from the mere knowledge of its symmetry group, rather than its explicit form.

We now introduce the important concept of irreducibility of a representation. Let $\{M(g)\}$ be an N-dimensional matrix representation of a group G. Matrices $M(g)$ can be considered as operators acting in a vector space $\mathcal{V}^N$. The representation is *irreducible* if there exists no trivial subspace of $\mathcal{V}^N$ (i.e. other than $\mathcal{V}^N$ itself and the space with dimension zero) which is invariant under the action of every matrix $M(g)$. A representation that is not irreducible is *reducible.* A reducible representation has the property that all matrices $M(g)$ leave a proper subspace $\mathcal{V}^{N_1}$ of $\mathcal{V}^N$ invariant. A reducible representation is therefore equivalent to a representation $\{\bar{M}(g)\}$ where every matrix $\bar{M}(g)$ has the form

$$\bar{M}(g) = \begin{pmatrix} A(g) & B(g) \\ 0 & C(g) \end{pmatrix} . \tag{13.27}$$

Here $A(g)$ is a $N_1 \times N_1$ matrix, $B(g)$ is a $N_1 \times (N - N_1)$ matrix, 0 is the null matrix with dimension $(N - N_1) \times N_1$ and $C(g)$ is a $(N - N_1) \times (N - N_1)$ matrix.

Reducibility is different from complete reducibility. A matrix representation $\{M(g)\}$ is *completely reducible* if it is equivalent to a block-diagonal representation $\{\bar{M}(g)\}$, each block being an irreducible representation of G. In other words there exist then irreducible representations $\{M_1(g)\}$, $\{M_2(g)\}$, ... such that all matrices $\bar{M}(g)$ have the form

$$\bar{M}(g) = \begin{pmatrix} M_1(g) & & \\ & M_2(g) & \\ & & \ddots \end{pmatrix} . \tag{13.28}$$

All elements of $\bar{M}(g)$ outside blocks vanish.

A completely reducible representation is obviously reducible. The converse is not always true. One can, however, prove the following theorem (Exercise 13.4).

Theorem Let $\{M(g)\}$ be a finite-dimensional unitary matrix representation of a group G. If such a representation is reducible, it is completely reducible.

Thus a finite-dimensional unitary matrix representation is either irreducible or completely reducible.

From a mathematical point of view the importance of irreducible representations is clear. With their help one can build all finite-dimensional unitary representations. But irreducible representations also have a particular importance in quantum mechanics. Let us go back to the theorem on p. 282. If G is the group of all symmetry transformations of a given Hamiltonian, then the correspondence (13.26) usually makes up an irreducible representation. In other words, eigenvectors associated with a degenerate eigenvalue of the Hamiltonian make up a vector space where, usually, G is represented irreducibly.[6]

It is easy to see that a matrix representation of a group G, restricted to elements of a subgroup K of G, is a matrix representation of K. Most of the time the restriction of an irreducible representation of G to a subgroup K is a reducible representation of K. It happens that only a subgroup K of all symmetry transformations of a given Hamiltonian is known. Then one should expect that eigenvectors associated with a degenerate eigenvalue of the Hamiltonian make up the space of a reducible representation of K.

Let $\{M^I(g)\}$ and $\{M^{II}(g)\}$ be two matrix representations of a group G, with dimensions N_I and N_{II}. Consider the tensor product of the two vector spaces in which these representations are defined. In this product space one can define a representation of G, with dimension $N_I N_{II}$, by associating the matrix $M^I(g) \otimes M^{II}(g)$ to each element g of G. One checks that this association is indeed a representation. It is called the *tensor product* of the two original representations.

The tensor product of two unitary representations is itself a unitary representation. Suppose that $\{M^I(g)\}$ and $\{M^{II}(g)\}$ are unitary and irreducible. Then $\{M^I(g) \otimes M^{II}(g)\}$ is unitary but, in general, reducible. A fundamental problem of the theory of matrix representations of groups consists in finding a representation like (13.28) that is equivalent to the tensor product. When that equivalent representation has been found, we say that the tensor product is expressed as a *direct sum* of irreducible representations.

13.5 Finite Groups

Two kinds of finite groups are particularly important in quantum mechanics: point groups and permutation groups. *Point groups* are finite subgroups of the group of spatial rotations (and space inversion). They are symmetry groups of noncollinear molecules.[7] The *permutation group* S_n, on the other hand, is the group of all permutations of n distinct objects. Thus it is a symmetry group of a system of n identical particles.

Before examining these groups in detail we will state a certain number of general properties of matrix representations of finite groups. Proofs of these

---

[6] This statement will be made more precise in Sect. 13.6.

[7] Collinear molecules have axial symmetry. Hence all rotations about the axis belong to their symmetry group.

statements can be found in books on group theory, in particular [105], Chap. 3 and [183], Chap. 3.

Let us first point out that any matrix representation of a finite group is equivalent to a unitary representation. The argument in Sect. 13.4 thus entails that every finite-dimensional reducible representation of a finite group is completely reducible.

Matrices of irreducible representations of a group have two important properties summarized in the *first and second Schur's lemmas.*

Lemma 1 Let $\{M(g)\}$ be matrices of an irreducible representation of a group G. Any matrix that commutes with every $M(g)$ is necessarily a multiple of the identity.

Lemma 2 Let $\{M^{(1)}(g)\}$ and $\{M^{(2)}(g)\}$ be matrices of two nonequivalent irreducible representations of a group G, with dimensions N_1 and N_2. Let F be a $N_2 \times N_1$ matrix such that, for any g in G,

$$F\,M^{(1)}(g) = M^{(2)}(g)\,F\,.$$

Then F is a null matrix.

Schur's lemmas are used to prove important orthogonality relations between matrix elements of irreducible representations. These relations are contained in the following theorem.

Theorem Let $\{M^{(i)}(g)\}$ and $\{M^{(j)}(g)\}$ be matrices of two nonequivalent irreducible unitary representations of a group G. Let N_i and N_j be the dimensions of these representations and let m be the order of G. Then

$$\sum_g M^{(i)}_{\alpha\beta}(g)\left\{M^{(j)}_{\gamma\delta}(g)\right\}^* = 0\,, \tag{13.29}$$

$$\sum_g M^{(i)}_{\alpha\beta}(g)\left\{M^{(i)}_{\gamma\delta}(g)\right\}^* = \frac{m}{N_i}\delta_{\alpha\gamma}\delta_{\beta\delta}\,, \tag{13.30}$$

the sums being carried out on all elements of G.

Let $\{M(g)\}$ be matrices of a representation of a group G. The *character* of the representation is a function $\chi(g)$ of group elements defined as

$$\chi(g) = \mathrm{Tr}\,\{M(g)\}\,. \tag{13.31}$$

Obviously, equivalent representations have the same character. For a given representation, character values are the same on all elements of a class.

Orthogonality relations (13.29) and (13.30) for matrix elements yield the following relation, where $\chi^{(i)}$ and $\chi^{(j)}$ are the characters of representations $M^{(i)}$ and $M^{(j)}$:

$$\sum_g \chi^{(i)}(g) \left\{\chi^{(j)}(g)\right\}^* = m\,\delta_{ij}. \tag{13.32}$$

Let $\{M(g)\}$ be a reducible representation of a finite group G. That representation can be written in the form (13.28) as a direct sum of irreducible representation $\{M^{(i)}(g)\}$. Suppose the representation $M^{(i)}$ occurs r_i times in decomposition (13.28). Then one can write

$$\chi(g) = \sum_j r_j\, \chi^{(j)}(g). \tag{13.33}$$

Making use of the orthogonality relation (13.32) one easily obtains that

$$r_i = \frac{1}{m} \sum_g \{\chi(g)\}^*\, \chi^{(i)}(g). \tag{13.34}$$

Equation (13.34) implies that the decomposition of a reducible representation in irreducible ones is unique (up to the position of matrices $M^{(i)}$). It also provides this decomposition explicitly, if the characters of irreducible representations are known.

The number of nonequivalent irreducible representations of a finite group is equal to the number of classes. If $\{N_j\}$ denotes dimensions of nonequivalent irreducible representations, one can show that

$$\sum_j N_j^2 = m, \tag{13.35}$$

where m is the order of the group.

Let us now investigate more specifically the finite groups often met in quantum mechanics. We first consider point groups, the symmetry groups of noncollinear molecules.

Let us introduce some notations. We will denote by O_n a rotation by an angle $2\pi/n$ about a given axis. A reflection in a plane perpendicular to the axis is denoted by σ_h. This reflection is equal to the product of space inversion and a rotation by π about the axis, that is,

$$\sigma_h = O_2 P = P O_2. \tag{13.36}$$

The product of σ_h and a rotation (about the same axis) is called a *rotary reflection* and is denoted by R_n. One has

$$R_n = \sigma_h O_n = O_n \sigma_h. \tag{13.37}$$

Finally, we denote by σ_v a reflection in a plane that contains a given axis. If two planes meet on an axis, then

$$\sigma_v \sigma_{v'} = O(2\phi), \tag{13.38}$$

where ϕ is the angle between the planes and $O(2\phi)$ denotes a rotation by 2ϕ about the axis.

We now enumerate all point groups. The characters and irreducible representations of the ones most often used are given in [138], Sect. 95, [183], Chap. 6 and [233], Chap. 5.

i) C_n, C_{nh} and C_{nv}. The group C_n contains n elements, rotations by an integral multiple of $2\pi/n$ about a given axis. The group C_{nh} contains $2n$ elements generated by those of C_n and by σ_h. The group C_{nv} also contains $2n$ elements, those of C_n and n reflections σ_v in planes meeting on the axis and making between them angles that are integral multiples of π/n. Note that C_n and C_{nh} are Abelian while C_{nv} is not $(n > 2)$.

ii) S_{2n}. The group S_{2n} contains $2n$ elements generated by the rotary reflection R_{2n}. It is Abelian.

iii) D_n, D_{nh} and D_{nd}. The group D_n contains the n rotations of C_n about a given axis. It also contains n rotations by π about n perpendicular axes, that make between them angles which are integral multiples of π/n. The group D_{nh} is generated by the elements of D_n and by σ_h, a reflection in the plane that contains the n axes. D_{nh} contains $4n$ elements. Because σ_h commutes with every element of D_{nh}, one can write $\mathrm{D}_{nh} = \mathrm{D}_n \otimes C_{1h}$. The group D_{nd} also contains $4n$ elements. They are generated by those of D_n and reflections in n planes, the latter containing the n-fold axis of D_n and bisecting the binary axes.

iv) T, T_d and T_h. These groups are associated with a tetrahedron. The group T contains rotations by π about axes linking centers of opposite sides of the tetrahedron, and rotations by $2\pi/3$ about axes linking vertices with centers of opposite faces. This adds up to 12 elements. The group T_d contains 24 elements, generated by those of T and reflections in planes containing one side and the center of the opposite side. T_d is the group of all symmetry transformations of the tetrahedron. The group T_h is generated by T and space inversion relative to the tetrahedron's center.

v) O and O_h. The group O is made of the 24 rotations that leave a cube invariant. The group O_h, generated by O and space inversion, is the group of all symmetry transformations of the cube.

vi) Y and Y_h. The group Y is made of the 60 rotations that leave an icosahedron (or a dodecahedron) invariant. Y_h is generated by Y and space inversion and is the group of all symmetry transformations of the icosahedron (or dodecahedron).

The group of all rotations about an axis is denoted by C_∞. It is an infinite group which, in a way, is obtained from C_n if $n \to \infty$. One similarly defines $\mathrm{C}_{\infty h}$, $\mathrm{C}_{\infty v}$, D_∞ and $\mathrm{D}_{\infty h}$. The group $\mathrm{D}_{\infty d}$ is isomorphic to $\mathrm{D}_{\infty h}$.

To illustrate concepts developed in this section, we will obtain selection rules for matrix elements of some operators. *Selection rules* are necessary conditions for matrix elements to differ from zero. Let $\{|E_i^\alpha\rangle, \alpha = 1, \ldots, N_i\}$ and $\{|E_j^\beta\rangle, \beta = 1, \ldots, N_j\}$ be orthonormal eigenvectors of the Hamiltonian, corresponding to eigenvalues E_i and E_j. Suppose the group of all symmetry transformations of H is a finite group, for instance one of those we have just enumerated. Let U_g be the unitary operator acting in the quantum system's state space and corresponding to element g of G. Let O be a linear operator such that, for any g in G

$$[O, U_g] = 0 . \tag{13.39}$$

We will consider matrix elements $\langle E_i^\alpha | O | E_j^\beta \rangle$. Let m be the order of G. Since U_g is unitary and commutes with O, one can write

$$\begin{aligned} \langle E_i^\alpha | O | E_j^\beta \rangle &= \frac{1}{m} \sum_g \langle E_i^\alpha | U_g^\dagger O U_g | E_j^\beta \rangle \\ &= \frac{1}{m} \sum_g \left\{ \sum_{\alpha'} \sum_{\beta'} \left[M_{\alpha'\alpha}^{(i)}(g) \right]^* M_{\beta'\beta}^{(j)}(g) \langle E_i^{\alpha'} | O | E_j^{\beta'} \rangle \right\} . \end{aligned} \tag{13.40}$$

Here matrices $M^{(i)}(g)$ and $M^{(j)}(g)$ are defined as in (13.26). If G is the group of all symmetry transformations of H, these matrices are usually irreducible. In this case one can use orthogonality relations (13.29) and (13.30). Carrying out first the sum over g in (13.40), one easily finds that

$$\langle E_i^\alpha | O | E_j^\beta \rangle = 0$$

if representations $\{M^{(i)}\}$ and $\{M^{(j)}\}$ are not equivalent. If, instead, they coincide, one finds that

$$\begin{aligned} \langle E_i^\alpha | O | E_j^\beta \rangle &= \frac{1}{m} \sum_{\alpha'} \sum_{\beta'} \frac{m}{N_i} \delta_{\alpha'\beta'} \delta_{\alpha\beta} \langle E_i^{\alpha'} | O | E_j^{\beta'} \rangle \\ &= \frac{1}{N_i} \delta_{\alpha\beta} \sum_{\beta'} \langle E_i^{\beta'} | O | E_j^{\beta'} \rangle \\ &= C \delta_{\alpha\beta} , \end{aligned}$$

where C is a constant independent of α and β. To conclude, we have shown that

$$\langle E_i^\alpha | O | E_j^\beta \rangle = \begin{cases} 0 & \text{if } \{M^{(i)}\} \not\sim \{M^{(j)}\} , \\ C\delta_{\alpha\beta} & \text{if } \{M^{(i)}\} = \{M^{(j)}\} . \end{cases} \tag{13.41}$$

To sum up, matrix elements of an operator O that commutes with the symmetry transformations of a Hamiltonian vanish, except if they are restricted to eigenvectors associated with the same irreducible representation. Relations like (13.41) are selection rules. Selection rules for operators O_s that themselves transform as an irreducible representation of G, that is,

$$U_g^\dagger O_\gamma U_g = \sum_{\gamma'} M_{\gamma'\gamma}^{(k)}(g) O_{\gamma'}, \tag{13.42}$$

are developed for example in [138], Sect. 97 and [183], Chap. 21.

Point groups are symmetry groups of noncollinear molecules. Also used in quantum mechanics are permutation groups. Let $H(1,2,\dots,n)$ be the Hamiltonian of n identical particles. Here the numbers 1, 2, ..., n stand symbolically for the set of (spatial, spin, internal) coordinates of each particle. Since particles are identical one must have

$$H(1,2\dots,n) = H\left\{P(1,2,\dots,n)\right\}, \tag{13.43}$$

where $P(1,2,\dots,n)$ denotes any permutation of 1, 2, ..., n. This means that the Hamiltonian is invariant under the permutation group of n objects. This group is denoted by S_n,[8] and it contains $n!$ elements.

The eigenvalue equation of $H(1,2,\dots,n)$ can be written as

$$H(1,2,\dots,n)|\Psi(1,2,\dots,n)\rangle = E|\Psi(1,2,\dots,n)\rangle. \tag{13.44}$$

Eigenvectors $|\Psi\rangle$ that correspond to a given eigenvalue make up a vector space where the group S_n is represented.

From a fundamental point of view, irreducible representations of S_n found in nature are extremely simple. As we will see in Chap. 17, the Pauli exclusion principle states that the state vector of a system of n identical particles is completely symmetric or completely antisymmetric under the exchange of coordinates of two particles, according as they have integral or half-integral spin. This means that

$$|\Psi\left\{P(1,2,\dots,n)\right\}\rangle = \begin{cases} +|\Psi(1,2,\dots,n)\rangle & \text{integral spin,} \\ \pm|\Psi(1,2,\dots,n)\rangle & \text{half-integral spin.} \end{cases} \tag{13.45}$$

Thus the state vector of a system of n identical particles always makes up a basis for an irreducible one-dimensional representation of the group S_n. For integral-spin particles it is the trivial representation, which associates 1 with each permutation. For half-integral-spin particles, it is the representation that associates +1 with each even permutation and −1 with each odd permutation.[9]

In practice, however, one also works with other representations of the group S_n in quantum mechanics. This comes from the fact that it is often very convenient to replace (13.44) by an approximation where spin (and other internal degrees of freedom) are ignored. Equation (13.44) can then be written as (see, for example, Chaps. 9 and 17)

---

[8] One should not confuse S_n with the point group S_{2n}.

[9] An even permutation is the product of an even number of permutations of two particles, while an odd permutation is the product of an odd number of such permutations.

$$H\{\boldsymbol{r}_1,\ldots,\boldsymbol{r}_n;-\mathrm{i}\hbar\boldsymbol{\nabla}_1,\ldots,-\mathrm{i}\hbar\boldsymbol{\nabla}_n\}\Psi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_n)=E\Psi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_n). \tag{13.46}$$

The Hamiltonian is still invariant under the group S_n. Wave functions corresponding to a value of E therefore make up a vector space where S_n is represented. Because certain wave function coordinates are suppressed, one cannot apply Pauli's principle. Hence possible representations of S_n are no longer restricted to those of (13.45). Completely antisymmetric wave functions are then built as sums of products of spatial functions $\Psi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_n)$ and spin functions ([183], Chap. 17).

13.6 Time Reversal

Consider a spinless particle with no degrees of freedom, in three dimensions. The state space has a basis of kets $|\boldsymbol{r}\rangle$. We assume the Hamiltonian does not depend on time.

In classical mechanics the time-reversal operation consists in changing the sign of the time coordinate. This transforms vectors $\boldsymbol{r}$ and $\boldsymbol{p}$ associated with a particle into vectors $\boldsymbol{r}$ and $-\boldsymbol{p}$. To define time reversal in quantum mechanics, we will look for an operator that transforms kets $|\boldsymbol{r}\rangle$ into $|\boldsymbol{r}\rangle$ and kets $|\boldsymbol{p}\rangle$ into $|-\boldsymbol{p}\rangle$.

Obviously, this operator cannot be linear. Indeed the only linear operator that transforms all vectors of a basis (here kets $|\boldsymbol{r}\rangle$) into themselves is the identity. But the identity cannot transform $|\boldsymbol{p}\rangle$ into $|-\boldsymbol{p}\rangle$.

Let K be the operator defined so that for any vector $|\psi\rangle$,

$$\langle\boldsymbol{r}|\psi'\rangle\equiv\langle\boldsymbol{r}|K|\psi\rangle=\langle\boldsymbol{r}|\psi\rangle^*. \tag{13.47}$$

In other words, K performs the complex conjugation of the components of a vector in the $|\boldsymbol{r}\rangle$ basis. One easily checks that operator K is antiunitary. Furthermore

$$\begin{aligned}\langle\boldsymbol{r}|K|\boldsymbol{r}_0\rangle&=\langle\boldsymbol{r}|\boldsymbol{r}_0\rangle^*=[\delta(\boldsymbol{r}-\boldsymbol{r}_0)]^*=\delta(\boldsymbol{r}-\boldsymbol{r}_0)=\langle\boldsymbol{r}|\boldsymbol{r}_0\rangle\,,\\ \langle\boldsymbol{r}|K|\boldsymbol{p}\rangle&=\langle\boldsymbol{r}|\boldsymbol{p}\rangle^*=\left[(2\pi\hbar)^{-3/2}\exp\left(\frac{\mathrm{i}}{\hbar}\boldsymbol{p}\cdot\boldsymbol{r}\right)\right]^*\\ &=(2\pi\hbar)^{-3/2}\exp\left(-\frac{\mathrm{i}}{\hbar}\boldsymbol{p}\cdot\boldsymbol{r}\right)=\langle\boldsymbol{r}|-\boldsymbol{p}\rangle\,.\end{aligned}$$

Hence K transforms $|\boldsymbol{r}_0\rangle$ into $|\boldsymbol{r}_0\rangle$ and $|\boldsymbol{p}\rangle$ into $|-\boldsymbol{p}\rangle$. This entails that

$$K\boldsymbol{R}K^{-1}=\boldsymbol{R},\qquad K\boldsymbol{P}K^{-1}=-\boldsymbol{P}. \tag{13.48}$$

From this one gets

$$K\boldsymbol{L}K^{-1}=-\boldsymbol{L}. \tag{13.49}$$

Using properties of spherical harmonics, one can check directly that K transforms a vector $|l, m\rangle$ into a vector $|l, -m\rangle$.

To show that K coincides with the *time-reversal operator*, it remains to check that vector $K|\psi\rangle$ evolves backward in time. Let $\psi(\boldsymbol{r}, t)$ be a solution of the Schrödinger equation, that is,

$$\mathrm{i}\hbar \frac{\partial \psi(\boldsymbol{r}, t)}{\partial t} = H\psi(\boldsymbol{r}, t). \tag{13.50}$$

Let us take the complex conjugate of (13.50). One finds that

$$\mathrm{i}\hbar \frac{\partial \psi^*(\boldsymbol{r}, t)}{\partial(-t)} = H^*\psi^*(\boldsymbol{r}, t), \tag{13.51}$$

so that $\psi^*(\boldsymbol{r}, t)$ satisfies the Schrödinger equation in terms of the variable $-t$.[10]

Although the Hamiltonian is always a Hermitian operator it does not, in general, coincide with its complex conjugate. Equations (13.50) and (13.51) therefore are not usually identical. They are if $H = H^*$, i.e. if the Hamiltonian is real in the coordinate representation. We say then that the quantum system is time-reversal invariant.

As an example, the Hamiltonian of a particle in a magnetic field is given by

$$H = \frac{1}{2m}\left[-\mathrm{i}\hbar\boldsymbol{\nabla} - q\boldsymbol{A}\right]^2 + V(\boldsymbol{r}). \tag{13.52}$$

This Hamiltonian is real if $\boldsymbol{A} = 0$. This means that a particle in an external magnetic field is not a time-reversal-invariant system.

Because it is represented by a nonunitary operator, time reversal is not a symmetry transformation like the ones we have hitherto considered. Anyway, suppose a system is time-reversal invariant. Let $\psi(\boldsymbol{r})$ be an eigenfunction of H. Since H is real in the coordinate representation, it is clear that $\psi^*(\boldsymbol{r})$ is also an eigenfucntion of H, with the same eigenvalue. This may entail that the set of eigenfunctions of H corresponding to a given eigenvalue is associated with a reducible (rather than irreducible) representation of G, where G is understood as the group of all unitary transformations that commute with H. Indeed suppose an irreducible representation of G is not equivalent to the representation obtained by taking the complex conjugate of every matrix. Necessarily, the space of eigenfunctions of H corresponding to a degenerate eigenvalue will always contain both representations if it contains ones. This additional degeneracy does not always occur, however, if an irreducible representation is equivalent to the conjugate representation. Functions Ψ and Ψ^* can then belong to only one irreducible representation.

---

[10] H^* is Hermitian if H is and it has the same eigenvalues as H. Operator H^* is therefore an admissible Hamiltonian.

The discussion in this section easily generalizes to a system of N spinless particles with no internal degrees of freedom. The time-reversal operator for a particle with nonzero spin is developed in Exercise (13.7).[11]

13.7 Lie Groups

Among groups that have an infinite number of elements, the most interesting are the *Lie groups.* In a Lie group, elements are labeled by (real or complex) parameters that vary continuously in some internal. Thus an arbitrary element of the group can be denoted as $U(\alpha_1, \alpha_2, \ldots, \alpha_D)$, where each α_n takes values in a given interval. The number of independent parameters necessary to specify an arbitrary element of the group, which we denoted by D, is the *dimension* of the group. It is important to realize that the dimension of a Lie group has nothing to do with the dimension of its matrix representations.

The product of two group elements can be represented as

$$U(\alpha_1, \ldots, \alpha_D) U(\beta_1, \ldots, \beta_D) = U(\gamma_1, \ldots, \gamma_D), \tag{13.53}$$

where

$$\gamma_n = \gamma_n(\alpha_1, \ldots, \alpha_D; \beta_1, \ldots, \beta_D). \tag{13.54}$$

In a Lie group all γ_n are analytic functions of their arguments. Hence they can always be expanded in power series.

The group of rotations and the group of spatial translations are both three-dimensional Lie groups. In the latter, the vector $\boldsymbol{r}_0$ is a set of three appropriate real parameters. One easily checks that the product law is analytic. For the rotation group, Euler angles are one possible set of three real parameters.[12] Although this may not be obvious, the product law is once again analytic.

The investigation of Lie groups is made easier by a remarkable characteristic. Most of their properties simply follow from the consideration of group elements close to the unit element. We first point out that parameters can always be picked so that the unit element corresponds to null values for all α_n, that is,

$$U(0, 0, \ldots, 0) = I. \tag{13.55}$$

Elements close to the unit element are then the $U(\alpha_1, \alpha_2, \ldots, \alpha_D)$ for which every α_n is small.

Consider a $U(\alpha_1, \ldots, \alpha_D)$ close to the unit element. Expanding U in a series of powers of α_n one can write

---

[11] Time reversal is thoroughly treated in [13], Sect. 13.3, [160], Sects. 15.15–15.22 and [198], Sect. 4.4.

[12] The vector $\hat{n}\theta$ is also a set of three appropriate parameters.

$$U(\alpha_1, \dots, \alpha_D) = I - \mathrm{i} \sum_{n=1}^{D} \alpha_n T_n + O(\alpha^2). \tag{13.56}$$

Here $O(\alpha^2)$ stands for all terms of order two or more in α_n. The T_n are defined as

$$T_n = \mathrm{i} \left(\frac{\partial U}{\partial \alpha_n} \right)_0 , \tag{13.57}$$

where the index zero means that derivatives should be evaluated at $\alpha_j = 0$ for all j. Operators T_n act in the same space as group elements U. In our case this is the Hilbert space of a quantum system. Operators T_n are called group *generators*. We emphasize that generators are defined solely in terms of group elements close to the identity.

Nevertheless, generators are used in a much wider context that the one in which they were defined. In fact they can generate most group elements, even those that are very far from the identity. To see this consider the expression

$$\left[I - \mathrm{i} \sum_{n=1}^{D} \left(\frac{\alpha_n}{L} \right) T_n \right]^L , \tag{13.58}$$

where the α_n are not small and where L is a positive and very large integer. Assume L is large enough for all α_n / L to be very small. For all practical purposes, the operator in square brackets is then a group element close to the identity. Expression (13.58) is the product of L elements of the group, and should therefore also belong to the group. Insofar as the argument remains valid as $L \to \infty$, the element

$$\begin{aligned} \lim_{L\to\infty} \left[I - \mathrm{i} \sum_{n=1}^{D} \left(\frac{\alpha_n}{L} \right) T_n \right]^L &= \lim_{L\to\infty} \left[I + \frac{1}{L} \left(-\mathrm{i} \sum_{n=1}^{D} \alpha_n T_n \right) \right]^L \\ &= \exp \left[-\mathrm{i} \sum_{n=1}^{D} \alpha_n T_n \right] \end{aligned} \tag{13.59}$$

belongs to the group. Obviously the right-hand side of (13.59) is not, in general, close to the identity.

The very schematic argument just outlined is certainly not rigorous. Nonetheless the final result can be proved rigorously. If $\{T_n, n = 1, 2, \dots, D\}$ is the set of generators of a finite-dimensional Lie group, then

$$U(\alpha_1, \dots, \alpha_D) \equiv \exp \left[-\mathrm{i} \sum_{n=1}^{D} \alpha_n T_n \right] \tag{13.60}$$

is a group element. The representation of group elements as in (13.60) is called the *exponential parametrization.*

We should point out that although the right-hand side of (13.60) is always a group element, the converse is not true in general. It may happen that some elements of a Lie group cannot be obtained by exponentiating generators. We shall come back to this later.

In quantum mechanics, elements $U(\alpha_1,\dots,\alpha_D)$ dealt with are in general unitary operators acting in the state space. The unitarity of group elements has an important consequence on generators. To derive it consider once more the case where each α_n is very small. Making use of (13.56), one can write

$$U(\alpha_1,\dots,\alpha_D)U^\dagger(\alpha_1,\dots,\alpha_D) = I - \mathrm{i}\sum_{n=1}^{D} \alpha_n T_n + \mathrm{i}\sum_{n=1}^{D} \alpha_n^* T_n^\dagger + O(\alpha^2). \tag{13.61}$$

The right-hand side of (13.61) coincides with the identity for all infinitesimal parameter values only if, for every n,

$$\alpha_n T_n - \alpha_n^* T_n^\dagger = 0\,. \tag{13.62}$$

It is not difficult to see that this equation cannot be satisfied if α_n is a complex variable. It can be satisfied, however, if α_n has the form $\bar{\alpha}_n \mathrm{e}^{\mathrm{i}\delta_n}$, where $\bar{\alpha}_n$ is a real variable and δ_n is a real constant. But then the factor $\mathrm{e}^{\mathrm{i}\delta_n}$ can be absorbed in T_n. Hence with no loss of generality one can set α_n real, in which case (13.62) entails that T_n is Hermitian. Summarizing we have the following result.

Theorem In a real parametrization, the generators of a unitary representation of a Lie group are Hermitian.

Lie group generators satisfy important relations that we will now derive. Let $\{T_n,\ n = 1, 2, \dots, D\}$ be the set of generators of a Lie group G. If α_m and β_n are two parameters, then $\exp(-\mathrm{i}\alpha_m T_m)$ and $\exp(-\mathrm{i}\beta_n T_n)$ belong to G. Therefore

$$U \equiv \exp(-\mathrm{i}\alpha_m T_m)\exp(-\mathrm{i}\beta_n T_n)\exp(\mathrm{i}\alpha_m T_m)\exp(\mathrm{i}\beta_n T_n) \tag{13.63}$$

also belongs to G, for all α_m and β_n. In the particular case where α_m and β_n are infinitesimal one can show that[13]

$$U = \exp\left\{-\alpha_m\beta_n\left[T_m, T_n\right] + O^3\right\}, \tag{13.64}$$

where O^3 stands for terms of order three or more in α_m and β_n. Obviously U is close to the identity. For it to belong to G, the exponential's argument must be a linear combination of generators. Thus one can write

---

[13] Actually this result was obtained in Exercise (4.2). Clearly the rotation generators introduced in Sect. 4.2 are instances of the operators T_n defined in this section.

$$[T_m, T_n] = \mathrm{i} \sum_{l=1}^{D} C^l_{mn} T_l . \tag{13.65}$$

Coefficients C^l_{mn} are constants. Equation (13.65) holds for every m and n. The generators of a Lie group together with their commutation relations form a *Lie algebra.* The coefficients C^l_{mn} are called *structure constants* of the Lie algebra. It is easy to check that if the generators are Hermitian, the structure constants are real.

It is important to note that the foregoing discussion applies mutatis mutandis to any matrix representation of a Lie group. Let $\{M(\alpha_1, \ldots, \alpha_D)\}$ be a matrix representation of G. Its generators (also denoted by T_n) are defined as

$$T_n = \mathrm{i} \left(\frac{\partial M}{\partial \alpha_n} \right)_0 .$$

In the exponential parametrization

$$M(\alpha_1, \ldots, \alpha_D) = \exp \left[-\mathrm{i} \sum_{n=1}^{D} \alpha_n T_n \right] .$$

If matrices M are unitary, the generators are Hermitian. The structure constants which, in a given parametrization, are completely determined by the group's product law, are the same in every matrix representation of the group and its generators.

As examples we now give a list of many Lie groups, each one defined by a matrix representation. The fact that the sets of matrices defined below make up groups follows from elementary properties of matrices and determinants. In each case we also give generators. Commutation relations can be obtained easily. Some of these examples are further developed in exercises.

i) $\mathrm{GL}(n;C)$ and $\mathrm{SL}(n;C)$. The group $\mathrm{GL}(n;C)$ is the set of all complex nonsingular (i.e. with nonzero determinant) $n \times n$ matrices. $\mathrm{SL}(n;C)$ is the subgroup of $\mathrm{GL}(n;C)$ made of complex $n \times n$ matrices with determinant 1. $\mathrm{GL}(n;C)$ and $\mathrm{SL}(n;C)$ have complex dimensions n^2 and $n^2 - 1$ respectively. Generators of $\mathrm{GL}(n;C)$ are arbitrary complex matrices while those of $\mathrm{SL}(n;C)$ are zero-trace complex matrices.

ii) $\mathrm{GL}(n;R)$ and $\mathrm{SL}(n;R)$. They are defined exactly as the previous ones, except that their parameters are real.

iii) $\mathrm{U}(n)$, $\mathrm{SU}(n)$ and $\mathrm{SU}(p,q)$. The group $\mathrm{U}(n)$ is the set of all unitary $n \times n$ matrices. Its subgroup $\mathrm{SU}(n)$ is made of all unitary $n \times n$ matrices with determinant 1. Groups $\mathrm{U}(n)$ and $\mathrm{SU}(n)$ have real dimensions n^2 and $n^2 - 1$ respectively. Generators of $\mathrm{U}(n)$ are arbitrary Hermitian matrices while those of $\mathrm{SU}(n)$ are zero-trace Hermitian matrices.

Let $I(p,q)$ be the diagonal matrix with p occurrences of -1 and q occurrences of $+1$ on its diagonal. The group $\mathrm{SU}(p,q)$ is the set of all $n \times n$ matrices with determinant 1 that satisfy

$$U^\dagger I(p,q)U = I(p,q)\,. \tag{13.66}$$

This group has real dimension $n^2 - 1$. Its generators are matrices H with zero trace that satisfy

$$H^\dagger I(p,q) = I(p,q)H\,. \tag{13.67}$$

iv) O(n), SO(n), O(p,q) and SO(p,q). The group O(n) is the set of all orthogonal real $n \times n$ matrices. Its subgroup SO(n) is made of all orthogonal real $n \times n$ matrices with determinant 1. Both groups have real dimension $n(n-1)/2$. Their generators are antisymmetric imaginary matrices. The group O(p,q) is the set of all real $n \times n$ matrices that satisfy

$$O^T I(p,q)O = I(p,q)\,. \tag{13.68}$$

This group has real dimension $n(n-1)/2$. Its generators are imaginary matrices Σ that satisfy

$$\Sigma^T I(p,q) = -I(p,q)\Sigma\,. \tag{13.69}$$

The group SO(p,q) is the subgroup of O(p,q) made of matrices with determinant 1.

To close this section, we will obtain generators of the tensor product of two representations of a Lie group. In Sect. 13.4 we saw that if $\{M^I(\alpha_1,\ldots,\alpha_D)\}$ and $\{M^{II}(\alpha_1,\ldots,\alpha_D)\}$ are two representations of a group, the set of matrices

$$M(\alpha_1,\ldots,\alpha_D) = M^I(\alpha_1,\ldots,\alpha_D) \otimes M^{II}(\alpha_1,\ldots,\alpha_D) \tag{13.70}$$

is also a representation. When, as the notation suggests, one deals with a Lie group, generators T_n of the representation $\{M\}$ can be written in terms of generators T_n^I and T_n^{II} of the representations $\{M^I\}$ and $\{M^{II}\}$ as

$$\begin{aligned} T_n &= \mathrm{i}\left(\frac{\partial M}{\partial \alpha_n}\right)_0 = \mathrm{i}\left[\frac{\partial}{\partial \alpha_n}\left[M^I(\alpha_1,\ldots,\alpha_D) \otimes M^{II}(\alpha_1,\ldots,\alpha_D)\right]\right]_0 \\ &= \mathrm{i}\left[\frac{\partial M^I}{\partial \alpha_n} \otimes M^{II} + M^I \otimes \frac{\partial M^{II}}{\partial \alpha_n}\right]_0 \\ &= \mathrm{i}\left[\left(\frac{\partial M^I}{\partial \alpha_n}\right)_0 \otimes I^{II} + I^I \otimes \left(\frac{\partial M^{II}}{\partial \alpha_n}\right)_0\right] \\ &= T_n^I \otimes I^{II} + I^I \otimes T_n^{II}\,. \end{aligned} \tag{13.71}$$

13.8 Spatial Translations

The group of spatial translations was introduced in Sect. 13.2 by specifying the action of an arbitrary element of the group $\mathcal{T}(\boldsymbol{r}_0)$ on basis vectors $|\boldsymbol{r}\rangle$:

$$\mathcal{T}(\boldsymbol{r}_0)|\boldsymbol{r}\rangle = |\boldsymbol{r}+\boldsymbol{r}_0\rangle .$$

The group of spatial translations is a Lie group. We will obtain its generators.

Consider an infinitesimal translation parametrized by $\delta\boldsymbol{r}$. From the definition of generators,

$$|\boldsymbol{r}+\delta\boldsymbol{r}\rangle = \mathcal{T}(\delta\boldsymbol{r})|\boldsymbol{r}\rangle = \left[I - \mathrm{i}\sum_{n=1}^{3}\delta x_n T_n\right]|\boldsymbol{r}\rangle . \tag{13.72}$$

Let X_l be a component of the position operator. We know that

$$X_l|\boldsymbol{r}+\delta\boldsymbol{r}\rangle = (x_l+\delta x_l)|\boldsymbol{r}+\delta\boldsymbol{r}\rangle . \tag{13.73}$$

On the other hand,

$$\begin{aligned} X_l|\boldsymbol{r}+\delta\boldsymbol{r}\rangle &= X_l\left[I - \mathrm{i}\sum_{n=1}^{3}\delta x_n T_n\right]|\boldsymbol{r}\rangle \\ &= \left[I - \mathrm{i}\sum_{n=1}^{3}\delta x_n T_n\right]X_l|\boldsymbol{r}\rangle + \left[X_l, I - \mathrm{i}\sum_{n=1}^{3}\delta x_n T_n\right]|\boldsymbol{r}\rangle \\ &= x_l\left[I - \mathrm{i}\sum_{n=1}^{3}\delta x_n T_n\right]|\boldsymbol{r}\rangle - \mathrm{i}\sum_{n=1}^{3}\delta x_n\,[X_l, T_n]\,|\boldsymbol{r}\rangle \\ &= x_l|\boldsymbol{r}+\delta\boldsymbol{r}\rangle - \mathrm{i}\sum_{n=1}^{3}\delta x_n\,[X_l, T_n]\,|\boldsymbol{r}\rangle . \end{aligned} \tag{13.74}$$

Comparing (13.73) and (13.74) we see that

$$\delta x_l|\boldsymbol{r}+\delta\boldsymbol{r}\rangle = -\mathrm{i}\sum_{n=1}^{3}\delta x_n\,[X_l, T_n]\,|\boldsymbol{r}\rangle . \tag{13.75}$$

To first order in $\delta\boldsymbol{r}$ this equation is equivalent to

$$\sum_{n=1}^{3}\delta x_n\,\{\delta_{ln}I + \mathrm{i}\,[X_l, T_n]\}\,|\boldsymbol{r}\rangle = 0 . \tag{13.76}$$

Equation (13.76) is satisfied for all δx_n and all $|\boldsymbol{r}\rangle$ if and only if

$$[X_l, T_n] = \mathrm{i}\delta_{ln}I . \tag{13.77}$$

Thus T_n can be identified with $\hbar^{-1}P_n$, where P_n is the n-th component of the momentum operator. Generators of the translation group are therefore proportional to components of the momentum operator.[14]

---

[14] Strictly speaking (13.77) entails that $T_n = \hbar^{-1}P_n + F_n(\boldsymbol{R})$. Since the translation group is Abelian, all its generators commute, which implies that $\boldsymbol{\nabla}\times\boldsymbol{F} = 0$. The function $\boldsymbol{F}$ can be eliminated by a suitable choice of the phase of kets $|\boldsymbol{r}\rangle$.

Following Sect. 13.7, the operator $\mathcal{T}(\boldsymbol{r}_0)$ for a finite translation has the form

$$\mathcal{T}(\boldsymbol{r}_0) = \exp\left[-\mathrm{i}\sum_{n=1}^{3}(x_0)_n T_n\right] = \exp\left[-\mathrm{i}\frac{\boldsymbol{r}_0 \cdot \boldsymbol{P}}{\hbar}\right]. \tag{13.78}$$

This operator is unitary. Moreover, it is diagonal in the basis of eigenvectors of $\boldsymbol{P}$. Indeed

$$\exp\left[-\mathrm{i}\frac{\boldsymbol{r}_0 \cdot \boldsymbol{P}}{\hbar}\right]|\boldsymbol{p}\rangle = \exp\left[-\mathrm{i}\frac{\boldsymbol{r}_0 \cdot \boldsymbol{p}}{\hbar}\right]|\boldsymbol{p}\rangle. \tag{13.79}$$

Thus every vector $|\boldsymbol{p}\rangle$ corresponds to the space of a one-dimensional representation of the translation group. That representation is characterized by the vector $\boldsymbol{p}$. All unitary irreducible representations of the translation group are one-dimensional. So is it with every Abelian group.

Note that if $|\psi'\rangle$ is the vector obtained from $|\psi\rangle$ through a translation by $\boldsymbol{r}_0$, that is,

$$|\psi'\rangle = \mathcal{T}(\boldsymbol{r}_0)|\psi\rangle, \tag{13.80}$$

then one simply has

$$\psi'(\boldsymbol{r}) \equiv \langle\boldsymbol{r}|\psi'\rangle = \langle\boldsymbol{r}|\mathcal{T}(\boldsymbol{r}_0)|\psi\rangle = \langle\boldsymbol{r}-\boldsymbol{r}_0|\psi\rangle \equiv \psi(\boldsymbol{r}-\boldsymbol{r}_0). \tag{13.81}$$

Consider a set of particles that interact through potentials that depend on coordinate differences only. The Hamiltonian is invariant under the translation group. Correlatively, the electronic Hamiltonian of an infinite crystal is invariant under a discrete infinite subgroup of the translation group. Indeed a crystal is made up of identical atomic aggregates, with each aggregate located at a node of the lattice. Let $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ be the fundamental lattice translation vectors. If n_a, n_b and n_c are integers, every translation parametrized by $n_a\boldsymbol{a}+n_b\boldsymbol{b}+n_c\boldsymbol{c}$ leaves the lattice invariant (Fig. 11.4, p. 250). Clearly the set or all such translations is a subgroup of the translation group. This subgroup is isomorphic to $Z \otimes Z \otimes Z$, where Z is the group of integers.

In general the group of discrete translations is not the largest symmetry group of a given crystal lattice. The lattice can also be invariant under a point group. We will now see, however, that the presence of translations severely limits the choice of point group ([183], Sect. 8.2).

We first point out that every crystal lattice is invariant under space inversion with respect to a node. Indeed the position of each node, relative to a given node, is specified by a triplet of integers (n_a, n_b, n_c). Space inversion brings a node (n_a, n_b, n_c) to the position $(-n_a, -n_b, -n_c)$, where there was a node before.

Suppose now that a rotation or a rotary reflection O by an angle θ leaves the crystal lattice invariant. With the help of three-dimensional rotation matrices it is easy to show that

$$\mathrm{Tr}(O) = 2\cos\theta \pm 1 . \tag{13.82}$$

The + and − signs correspond to rotation and rotary reflection, respectively. Equation (13.82) remains true if matrix O is expressed in the (nonorthonormal) basis of vectors $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$. But in this basis elements of O can only be integers, since O must transform any triplet of integers into a triplet of integers. Hence $\mathrm{Tr}(O)$ must be an integer. From (13.82) one finds that

$$\cos\theta = 0\,, \pm\frac{1}{2}\,, \pm 1 . \tag{13.83}$$

Thus only elementary rotations or rotary reflection by 2π, $2\pi/2$, $2\pi/3$, $2\pi/4$ and $2\pi/6$ can belong to the symmetry group of a crystal lattice.

One can show that if C_n $(n > 2)$ is a symmetry group of a crystal lattice, so is C_{nv}. This means that if the lattice is invariant under a rotation by $2\pi/n$, it is invariant under reflections in n planes that contain the axis and make with each other angles of π/n. This remark, together with restrictions on rotations and rotary reflections and the presence of space inversion, implies that the only point groups that can leave a crystal lattice invariant are the following:

$$\mathrm{S}_2\,, \mathrm{C}_{2h}\,, \mathrm{D}_{2h}\,, \mathrm{D}_{3h}\,, \mathrm{D}_{4h}\,, \mathrm{D}_{6h}\,, \mathrm{O}_h . \tag{13.84}$$

These groups correspond to the triclinic, monoclinic, orthorhombic, trigonal, tetragonal, hexagonal and cubic lattices, respectively.

The group of all symmetry transformations of an infinite crystal lattice is thus produced by discrete translations and one of the seven point groups (13.84). One should note, however, that the symmetry group of the crystal lattice does not in general coincide with the symmetry group of the crystal. To determine the latter (called a *space group*), one must consider the specific atomic patterns that are repeated at every node. On this the reader is referred, for instance, to [57], Chap. 9, [138], Chap. 13, [183], Chap. 8 and [226], Chaps. 6 and 8. Let us simply mention that spatial translations leaving a crystal lattice invariant always leave the associated crystal invariant, and that the crystal's point group of symmetry is a subgroup of one of the groups (13.84).

Exercises

13.1. Show that if x', y' and z' are related to x, y, and z by a rotation, a translation or a space inversion, then

$$\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} = \frac{\partial^2}{\partial x'^2} + \frac{\partial^2}{\partial y'^2} + \frac{\partial^2}{\partial z'^2}\,,$$

$$\delta(x - x_0)\delta(y - y_0)\delta(z - z_0) = \delta(x' - x_0')\delta(y' - y_0')\delta(z' - z_0') .$$

13.2. The group of permutations of three distinct objects has six elements that can be denoted by

$$\begin{aligned}
g_0 &: \quad (123) \to (123)\,; \\
g_1 &: \quad (123) \to (231)\,; \\
g_2 &: \quad (123) \to (312)\,; \\
g_3 &: \quad (123) \to (321)\,; \\
g_4 &: \quad (123) \to (213)\,; \\
g_5 &: \quad (123) \to (132)\,.
\end{aligned}$$

Write down the multiplication table of group elements.

13.3.

a) Show that the set of (positive, negative, and null) integers is a group with respect to the operation of addition.

b) Let n be an arbitrary integer. Show that the correspondence

$$n \to \begin{pmatrix} \mathrm{e}^n & 0 \\ 0 & 10^{n/2} \end{pmatrix}$$

is a faithful representation of that group. Is this representation irreducible? Please explain.

13.4. Show that every finite-dimensional unitary reducible matrix representation of a group G is completely reducible.

13.5. Give all symmetry transformations of the H_2O, CO_2 and NH_3 molecules (Fig. 13.1). What is, in each case, the symmetry group?

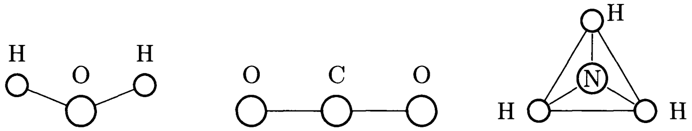

Fig. 13.1. The H_2O, CO_2 and NH_3 molecules; the nitrogen atom is above the plane

13.6. Four identical spheres are laid out on a square and rigidly attached one to the others (Fig. 13.2).

a) Find the eight rotations that leave this pattern invariant.

b) Show that these rotations are all obtained from two transformations: a rotation by $\pi/2$ about an axis perpendicular to the square and going through its center; and a rotation by π about the broken line. What is the resulting point group?

c) What is the point group made by the set of all symmetry transformations of the pattern?

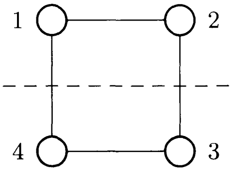

Fig. 13.2. Four identical spheres laid out on a square

13.7. Let S_x, S_y and S_z be spin operators of a spin s particle. In the basis developed in Sect. 4.3, S_x and S_z are real while S_y is imaginary. Let T be the operator defined as

$$T = \exp\left\{-\frac{\mathrm{i}}{\hbar}\pi S_y\right\} K\,,$$

where K is the complex conjugation defined in Sect. 13.6. Show that ($i = 1$, 2, 3)

$$T S_i T^{-1} = -S_i\,.$$

Why is T the time-reversal operator of a spin s particle in three dimensions?

13.8. Consider the set of nonsingular real orthogonal $n \times n$ matrices.

a) Show that this set is a group with respect to matrix product. The group is called $\mathrm{O}(n)$.

b) Show that the subset of $\mathrm{O}(n)$ made of matrices with determinant equal to 1 is a subgroup of $\mathrm{O}(n)$. This subgroup is called $\mathrm{SO}(n)$.

c) Show that the dimension of $\mathrm{SO}(n)$ is equal to $n(n-1)/2$.

13.9. Let O be a matrix close to the identity, i.e. $O = I - \mathrm{i}\Sigma$ where Σ is infinitesimal.

a) Show that O is real and orthogonal if and only if Σ is imaginary and antisymmetric. All terms of order Σ^2 can be neglected.

b) Any imaginary antisymmetric 3×3 matrix can be written as

$$\begin{pmatrix} 0 & \mathrm{i}\alpha_1 & \mathrm{i}\alpha_2 \\ -\mathrm{i}\alpha_1 & 0 & \mathrm{i}\alpha_3 \\ -\mathrm{i}\alpha_2 & -\mathrm{i}\alpha_3 & 0 \end{pmatrix},$$

where every α_i is real. Use this result to obtain the SO(3) generators and its Lie algebra.

13.10. Consider the set of nonsingular complex unitary $n \times n$ matrices.

a) Show that this set is a group with respect to matrix product. The group is called $\mathrm{U}(n)$.

b) Show that the subset of $\mathrm{U}(n)$ made of matrices with determinant equal to 1 is a subgroup of $\mathrm{U}(n)$. This subgroup is called $\mathrm{SU}(n)$.

c) Show that the dimension of $\mathrm{SU}(n)$ is equal to $n^2 - 1$.

13.11. Let U be a matrix close to the identity, i.e. $U = I - \mathrm{i}H$ where H is infinitesimal.

a) Show that U is unitary if and only if H is Hermitian.

b) Any Hermitian 2×2 matrix can be written as

$$\begin{pmatrix} \alpha_1 + \alpha_2 & \alpha_3 + \mathrm{i}\alpha_4 \\ \alpha_3 - \mathrm{i}\alpha_4 & \alpha_1 - \alpha_2 \end{pmatrix},$$

where every α_i is real. Use this result to obtain the U(2) generators and its Lie algebra.

13.12. Let M be a real 3×3 matrix and let O be an element of SO(3). The set of all M matrices can be viewed as a nine-dimensional vector space denoted by $\mathcal{V}^9$. One can define the action of an element of SO(3) in $\mathcal{V}^9$ as

$$O(M) = M' = OMO^T . \tag{13.85}$$

a) Show that these operators, acting in $\mathcal{V}^9$, make up a representation of SO(3).

b) Show that $\mathcal{V}^9$ is the direct sum of $\mathcal{V}^5$, $\mathcal{V}^3$ and $\mathcal{V}^1$, where

$\mathcal{V}^5$ is the set of real symmetric 3×3 matrices with zero trace;

$\mathcal{V}^3$ is the set of real antisymmetric 3×3 matrices;

$\mathcal{V}^1$ is the set of real multiples of the identity.

c) Show that $\mathcal{V}^5$, $\mathcal{V}^3$ and $\mathcal{V}^1$ are invariant under transformation (13.85).

13.13. The group SL(2; R) is made up of the set of real 2×2 matrices with determinant 1. An element of the group can be parametrized as

$$\begin{pmatrix} 1 + \alpha_1 & \alpha_2 \\ \alpha_3 & 1 + \frac{\alpha_2\alpha_3 - \alpha_1}{1+\alpha_1} \end{pmatrix}.$$

a) Compute generators T_1, T_2 and T_3 of SL(2; R).

b) Compute commutators $[T_1, T_2]$, $[T_2, T_3]$ and $[T_3, T_1]$, and write each commutator as a linear combination of generators.

13.14. The set of all nonsingular real 2×2 matrices makes up the group GL(2; R). Parametrize group elements close to the identity. Find four linearly independent generators, and give their commutation relations.

13.15. The Hamiltonian of n independent identical harmonic oscillators is given by

$$H = \sum_{i=1}^{n} \left\{ \frac{1}{2m} P_i^2 + \frac{1}{2} m\omega^2 X_i^2 \right\},$$

where P_i and X_i are the momentum and position operators.
a) Show that H can be written as

$$H = \hbar\omega \left[\sum_{i=1}^{n} a_i^\dagger a_i + \frac{n}{2} I \right] ,$$

where a_i and $a_i^\dagger$ (the annihilation and creation operators) satisfy

$$[a_i, a_j] = 0 = \left[a_i^\dagger, a_j^\dagger\right] , \qquad \left[a_i, a_j^\dagger\right] = \delta_{ij} .$$

b) Let

$$\mathcal{L} = - \sum_{k,l=1}^{n} L_{kl} a_k a_l^\dagger \quad \text{and} \quad \bar{\mathcal{L}} = - \sum_{k,l=1}^{n} \bar{L}_{kl} a_k a_l^\dagger ,$$

where $L_{kl} = L_{lk}^*$ and $\bar{L}_{kl} = \bar{L}_{lk}^*$. Check that $\mathcal{L}$ and $\bar{\mathcal{L}}$ are Hermitian and that

$$[\mathcal{L}, \bar{\mathcal{L}}] = - \sum_{k,l=1}^{n} \left[L, \bar{L}\right]_{kl} a_k a_l^\dagger ,$$

thus showing that the Lie algebra of U(n) can be represented by means of annihilation and creation operators.
c) Check that $[\mathcal{L}, H] = 0$, thus showing that U(n) is a symmetry group of the Hamiltonian. [Actually the really significant group is SU(n), since $\mathcal{L}$ essentially reduces to H if $L_{kl} = \delta_{kl}$.]

13.16. Let T_m, T_n and T_l be three generators of a Lie algebra.
a) Prove the following identity:

$$[[T_m, T_n], T_l] + [[T_n, T_l], T_m] + [[T_l, T_m], T_n] = 0 .$$

b) Show that the structure constants C_{mn}^l defined from relation (13.65) satisfy

$$\sum_k \left\{ C_{mn}^k C_{kl}^i + C_{nl}^k C_{km}^i + C_{lm}^k C_{kn}^i \right\} = 0 .$$

13.17. Let T_n^I and T_n^{II} be generators of two representations of a Lie group. Let T_n be a generator of the tensor-product representation, that is,

$$T_n = T_n^I \otimes I^{II} + I^I \otimes T_n^{II} .$$

Show directly that the commutation relations of the T_n are the same as those of the T_n^I and T_n^{II}.

13.18. Let $\boldsymbol{a}$ be a real three-dimensional vector and let O be an orthogonal 3×3 matrix. Denote by $(O, \boldsymbol{a})$ the operation that consists in making first a rotation O and then a translation $\boldsymbol{a}$.

a) Show that $(O_2, \boldsymbol{a}_2) \cdot (O_1, \boldsymbol{a}_1) = (O_2 O_1, \boldsymbol{a}_2 + O_2 \boldsymbol{a}_1)$ and that the set of all $(O, \boldsymbol{a})$ is a group with respect to this product.

b) Show that the following matrices make up a faithful representation of this group (here $\boldsymbol{a}$ is a column vector and 0 a row of three zeros):

$$\begin{pmatrix} O & \boldsymbol{a} \\ 0 & 1 \end{pmatrix} .$$

14 Rotations and Angular Momentum

Spatial rotations have been encountered repeatedly and in different contexts. Spin spaces, brought to light by the Stern–Gerlach experiment, were treated in Chap. 4. Orbital angular momentum operators were introduced in Chap. 7. In Chap. 13 the rotation group was defined. In the present chapter we will first show that group-theoretical concepts afford a synthesis of all results on rotations that have hitherto been obtained. In so doing we will obtain rotation matrices for arbitrary spin. We will next develop the formalism for the addition of two angular momenta. From a group-theoretical point of view it is the decomposition, into irreducible representations, of a reducible representation of the rotation group. We will introduce the concept of irreducible tensor operator and prove an important theorem about matrix elements of such operators. Finally we will examine two important quantum systems whose symmetry groups are larger than the rotation group.[1]

14.1 The Rotation Group

In Chap. 13 the Lie group O(3) was defined as the set of all real orthogonal 3×3 matrices, i.e. real matrices that obey the constraint $O^T O = I$. A real 3×3 matrix can be considered as a linear operator transforming a vector $\boldsymbol{r}$ with components x_i into a vector $\boldsymbol{r}'$ with components x_i' given by

$$x_i' = \sum_j O_{ij} x_j \, . \tag{14.1}$$

Clearly the length of $\boldsymbol{r}$ is invariant under this transformation. Indeed

$$\begin{aligned}\boldsymbol{r}' \cdot \boldsymbol{r}' &= \sum_i x_i' x_i' = \sum_i \left\{ \sum_j O_{ij} x_j \right\} \left\{ \sum_k O_{ik} x_k \right\} \\ &= \sum_{j,k} x_j x_k \sum_i O_{ji}^T O_{ik} = \sum_{j,k} x_j x_k \delta_{jk}\end{aligned}$$

---

[1] The rotation group and angular momentum are treated in detail in [27], [28], [73], [149] and [245].

L. Marchildon, *Quantum Mechanics*

$$= \sum_j x_j x_j = \boldsymbol{r} \cdot \boldsymbol{r} \,. \tag{14.2}$$

Conversely, it is not difficult to show that a linear operator preserves the length of an arbitrary vector $\boldsymbol{r}$ only if the corresponding matrix is orthogonal.

The group O(3) is thus intimately related to spatial rotations. Note, however, that O(3) also contains space inversion, i.e. the transformation $\boldsymbol{r}' = -\boldsymbol{r}$. We know that every matrix O belonging to O(3) has its determinant equal to ± 1. The set of elements of O(3) with determinant equal to $+1$ makes up the subgroup SO(3) of O(3). Space inversion does not belong to SO(3). Actually SO(3) is what is called the group of *proper rotations*. It is easy to show that every element of O(3) with determinant equal to -1 can be represented as a product of space inversion and an element of SO(3).

Any proper rotation can be parametrized by the Euler angles α, β and γ. We define $O(\alpha,\beta,\gamma)$ as the product of three rotations: one by an angle α about the z axis, followed by one by an angle β about the y axis, and finally one by an angle γ about the z axis.[2] Thus

$$\begin{aligned} &O(\alpha,\beta,\gamma) \\ &= \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} . \end{aligned} \tag{14.3}$$

The domain of the angles α, β and γ is given by

$$0 \le \alpha < 2\pi, \qquad 0 \le \beta < \pi, \qquad 0 \le \gamma < 2\pi. \tag{14.4}$$

Clearly $O(\alpha,\beta,\gamma)$ belongs to SO(3), since each factor is an element of SO(3). Conversely it is shown in Exercise (14.1) that every real orthogonal 3×3 matrix with determinant 1 can be factorized as in (14.3).

The Lie algebra of SO(3) was introduced in Sect. 13.7. It is made up of all imaginary antisymmetric 3×3 matrices. Such a matrix can always be written as a linear combination of matrices I_x, I_y and I_z given in (4.6) and (4.7). Hence an element of SO(3) close to the identity can be written as

$$O = I - \mathrm{i}(\delta\theta_x I_x + \delta\theta_y I_y + \delta\theta_z I_z) = I - \mathrm{i}(\hat{n}\delta\theta)\cdot \boldsymbol{I}, \tag{14.5}$$

where I_x, I_y and I_z were combined into a "vector" $\boldsymbol{I}$, and where a unit vector $\hat{n}$ and an infinitesimal parameter $\delta\theta$ were introduced so that $\hat{n}\delta\theta = (\delta\theta_x, \delta\theta_y, \delta\theta_z)$.

---

[2] It is important to note that when defining a product of rotations, we use fixed axes $(\hat{x}, \hat{y}, \hat{z})$. Euler angles are often defined by a different convention, in which axes rotate with each rotation. Define $O'(\alpha,\beta,\gamma)$ as the product of the following three rotations: one by an angle α about the z axis, followed by one by an angle β about the new y axis, and finally one by an angle γ about the new z axis. Then one can show ([73], p. 8) that $O'(\alpha,\beta,\gamma) = O(\gamma,\beta,\alpha)$.

The commutation relations of matrices I_x, I_y and I_z are given in (4.10)–(4.12). These matrices can be written in terms of the Levi-Civita symbol defined in (7.54). If $(I_i)_{jk}$ denotes the (jk) element of matrix I_i, one can see that

$$(I_i)_{jk} = -\mathrm{i}\varepsilon_{ijk}\,. \tag{14.6}$$

The (jk) element of (14.5) can therefore be written as

$$O_{jk} = \delta_{jk} - \delta\theta \sum_i n_i \varepsilon_{ijk}\,.$$

Following (14.1) this rotation transforms vector $\boldsymbol{r}$ into a vector $\boldsymbol{r}'$ given by

$$\boldsymbol{r}' = \boldsymbol{r} + \delta\theta(\hat{n} \times \boldsymbol{r})\,. \tag{14.7}$$

This shows that the infinitesimal parameters $(\delta\theta_x, \delta\theta_y, \delta\theta_z) = \hat{n}\delta\theta$ correspond to a rotation by an angle $\delta\theta$ about the $\hat{n}$ axis.

Without using the present terminology, we have obtained in Chap. 4 all irreducible representations of the Lie algebra of SO(3). They are given in (4.26)–(4.28) and (4.13). Each representation is associated with a nonnegative number j, either integral or half-integral. Representation j is made up of three Hermitian matrices I_x, I_y and I_z, with dimension $(2j+1) \times (2j+1)$, that obey the commutation relations (4.10)–(4.12).

Irreducible representations of SO(3) can be obtained from irreducible representations of its Lie algebra by exponentiation. Following Chap. 13, unitary matrices of the representation j are given by

$$D^{(j)}(\hat{n};\theta) = \exp\{-\mathrm{i}(\hat{n}\theta) \cdot \boldsymbol{I}\}\,, \tag{14.8}$$

where the components of $\boldsymbol{I}$ are generators of the representation j. Strictly speaking, however, only integral values of j correspond to true representations of SO(3). To justify this statement let us consider the matrix $D^{(j)}(\hat{z};2\pi)$. Making use of (4.26) one easily sees that

$$D^{(j)}(\hat{z};2\pi) = \exp(-2\pi\mathrm{i}I_z) = (-1)^{2j}I\,, \tag{14.9}$$

where I is the identity matrix in the space of dimension $2j+1$. If $j=1$ then $D^{(1)}(\hat{z};2\pi) = I$, which corresponds to the fact that a rotation by 2π about the z axis coincides with the null rotation. For any integral j, $D^{(j)}(\hat{z};2\pi) = I$. But for half-integral j, $D^{(j)}(\hat{z};2\pi) = -I$. This means that rigorously, matrices $D^{(j)}$ for half-integral j are not genuine representations of SO(3). For every j, however,

$$D^{(j)}(\hat{z};4\pi) = I\,. \tag{14.10}$$

To properly understand what happens with half-integral values of j, it is useful to go back to the spin 1/2 representation explicitly obtained in Chap. 4. In (4.66) and (4.67) we found that

$$D^{(1/2)}(\hat{n};\theta) = \exp\left\{-\mathrm{i}(\hat{n}\theta)\cdot\frac{\sigma}{2}\right\}$$
$$= \begin{pmatrix} \cos\frac{\theta}{2} - \mathrm{i}n_z\sin\frac{\theta}{2} & -(\mathrm{i}n_x + n_y)\sin\frac{\theta}{2} \\ (-\mathrm{i}n_x + n_y)\sin\frac{\theta}{2} & \cos\frac{\theta}{2} + \mathrm{i}n_z\sin\frac{\theta}{2} \end{pmatrix}.$$

One can see that to parametrize an arbitrary spatial rotation $D^{(1)}(\hat{n};\theta)$, it is sufficient to pick θ in the interval $[0,2\pi)$ and to pick the $\hat{n}$ axis in the upper hemisphere (i.e. such that $n_z \geq 0$). Indeed a rotation by θ about $\hat{n}$ is equivalent to a rotation by $2\pi - \theta$ about $-\hat{n}$, and a rotation by $2\pi + \theta$ is equivalent to a rotation by θ. Such values of $\hat{n}$ and θ are not enough, however, to parametrize $D^{(1/2)}$. Indeed

$$D^{(1/2)}(\hat{n};2\pi+\theta) = -D^{(1/2)}(\hat{n};\theta). \tag{14.11}$$

The set of $D^{(1/2)}$ matrices is parametrized by picking $\hat{n}$ anywhere in the upper hemisphere and θ in the interval $[0,4\pi)$.

Observing matrices $D^{(1/2)}(\hat{n};\theta)$ carefully, one sees they are all unitary and have determinant 1. This means they all belong to the group SU(2). Conversely it is shown in Exercise (14.2) that every unitary 2×2 matrix with determinant 1 can be written in the form $D^{(1/2)}(\hat{n};\theta)$, with $n_z \geq 0$ and $0 \leq \theta < 4\pi$. Matrices $D^{(1/2)}(\hat{n};\theta)$ therefore make up a faithful representation of SU(2).

From a group-theoretical point of view, these observations follow from the fact that SO(3) is homomorphic to SU(2). The product laws of these two groups are the same. But to every element of SO(3) there correspond two elements of SU(2). Matrices $D^{(j)}(\hat{n};\theta)$ make up, for integral j, a faithful representation of SO(3) and, for half-integral j, a faithful representation of SU(2).[3]

The use of representations of SU(2) is no doubt legitimate in group theory, but is it in quantum mechanics? In other words can it be that under rotations, a physical system transforms through operators which, strictly speaking, do not make up a representation of the rotation group? To answer this question, we should recall that the correspondence between quantum states and state vectors is not one-to-one. Thus for any real α, the normalized vectors $|\psi\rangle$ and $e^{\mathrm{i}\alpha}|\psi\rangle$ correspond to the same quantum state. Suppose we associate a unitary operator $U(O)$ with each rotation O. For $U(O)$ to represent O it is enough that one has

$$U(O_1)U(O_2) = \exp\{\mathrm{i}\alpha(O_1,O_2)\}\,U(O_1O_2), \tag{14.12}$$

---

[3] Groups SU(2) and SO(3) are *connected*, which means that each of their elements can be related to the identity by a continuous curve in parameter space. One shows that SU(2) is *simply connected*, which means that any closed curve in parameter space can be brought to a point continuously. SO(3), on the other hand, is not simply connected. Note that O(3) is not connected, since its elements with determinant -1 cannot be brought to the identity in a continuous way.

instead of the strict relation (13.20). A *projective representation* of a group is a correspondence that satisfies (14.12).[4] One can show ([238], Chaps. 20 and 21) that all projective representations of SO(3) are true representations of SU(2). This justifies the introduction of half-integral-spin representations and shows that, together with integral-spin representations, they account for all state spaces associated with spatial rotations.

Note that for any j, matrices $D^{(j)}$ can be parametrized by Euler angles. In general one defines

$$D^{(j)}(\alpha,\beta,\gamma) = \exp(-\mathrm{i}\gamma I_z)\exp(-\mathrm{i}\beta I_y)\exp(-\mathrm{i}\alpha I_z). \tag{14.13}$$

For integral j the domain of parameters α, β and γ was given in (14.4). For half-integral j the domain of α and γ stays the same, but β must be picked in the interval $[0, 2\pi)$. Matrices $D^{(1/2)}(\alpha,\beta,\gamma)$ were obtained in Chap. 4. Matrices $D^{(j)}(\alpha,\beta,\gamma)$ for arbitrary j will be obtained in Sect. 14.2.

In Sect. 13.2 a representation of the rotation group in the space of kets $|\boldsymbol{r}\rangle$ was defined as

$$|\boldsymbol{r}'\rangle = |O(\hat{n};\theta)\boldsymbol{r}\rangle = \mathcal{R}(\hat{n};\theta)|\boldsymbol{r}\rangle\,.$$

We close this section by obtaining the generators of this representation.

We saw in Sect. 13.8 that generators of the translation group are components of the momentum operator multiplied by $\hbar^{-1}$. Hence in general

$$|\boldsymbol{r}+\delta\boldsymbol{r}\rangle = |\boldsymbol{r}\rangle - \frac{\mathrm{i}}{\hbar}\delta\boldsymbol{r}\cdot\boldsymbol{P}|\boldsymbol{r}\rangle\,.$$

Using (14.7) one can write

$$\begin{aligned}\mathcal{R}(\hat{n};\delta\theta)|\boldsymbol{r}\rangle &= |\boldsymbol{r}+\delta\theta(\hat{n}\times\boldsymbol{r})\rangle \\ &= |\boldsymbol{r}\rangle - \frac{\mathrm{i}}{\hbar}(\delta\theta\,\hat{n}\times\boldsymbol{r})\cdot\boldsymbol{P}|\boldsymbol{r}\rangle \\ &= |\boldsymbol{r}\rangle - \frac{\mathrm{i}}{\hbar}\delta\theta\,\hat{n}\cdot(\boldsymbol{r}\times\boldsymbol{P})\,|\boldsymbol{r}\rangle \\ &= |\boldsymbol{r}\rangle - \frac{\mathrm{i}}{\hbar}(\hat{n}\delta\theta)\cdot(\boldsymbol{R}\times\boldsymbol{P})\,|\boldsymbol{r}\rangle\,,\end{aligned}$$

where $\boldsymbol{R}$ is the position operator. The last equality follows from the fact that different components of $\boldsymbol{R}$ and $\boldsymbol{P}$ commute. But $\boldsymbol{R}\times\boldsymbol{P} = \boldsymbol{L}$, the angular momentum operator. Thus

$$\mathcal{R}(\hat{n};\delta\theta)|\boldsymbol{r}\rangle = \left\{I - \mathrm{i}(\hat{n}\delta\theta)\cdot\left(\hbar^{-1}\boldsymbol{L}\right)\right\}|\boldsymbol{r}\rangle\,. \tag{14.14}$$

This equation implies that generators of the rotation group, in the representation associated with kets $|\boldsymbol{r}\rangle$, are components of the angular momentum operator multiplied by $\hbar^{-1}$.

---

[4] Projective representations are examined, for instance, in [105], Chap. 12, [218], Sect. 3.10 and [232], Sect. 2.7.

For a finite rotation

$$\mathcal{R}(\hat{n};\theta) = \exp\left\{-\mathrm{i}\frac{\hat{n}\theta}{\hbar}\cdot \boldsymbol{L}\right\}. \tag{14.15}$$

Obviously this representation is reducible. It is interesting to point out that in group-theoretical language, Sect. 7.3 can be interpreted in the following way. The representation $\mathcal{R}(\hat{n};\theta)$ of the rotation group in the space of kets $|\boldsymbol{r}\rangle$ decomposes into irreducible representations. These act on spaces of vectors $\{|\psi_{lm}\rangle\}$, where l is an integer and $-l \leq m \leq l$. The projection of $|\psi_{lm}\rangle$ on $|\boldsymbol{r}\rangle$ is given by

$$\langle \boldsymbol{r}|\psi_{lm}\rangle = f_l(r)Y_{lm}(\theta,\phi), \tag{14.16}$$

where $f_l(r)$ is a function of r (the same for all values of m associated with an irreducible representation) and Y_{lm} is a spherical harmonic. Note that representations with integral j only can be realized in terms of kets $|\boldsymbol{r}\rangle$.

14.2 Finite Rotations

Rotation matrices $D^{(j)}$ were defined in Sect. 14.1 for both parametrizations (α,β,γ) and $(\hat{n};\theta)$ of rotations:

$$D^{(j)}(\alpha,\beta,\gamma) = \exp\{-\mathrm{i}\gamma[I_z]\}\exp\{-\mathrm{i}\beta[I_y]\}\exp\{-\mathrm{i}\alpha[I_z]\}, \tag{14.17}$$

$$D^{(j)}(\hat{n};\theta) = \exp\{-\mathrm{i}(\hat{n}\theta)\cdot[\boldsymbol{I}]\}. \tag{14.18}$$

It is useful here to recall the bracket notation, to emphasize that $[I_x]$, $[I_y]$ and $[I_z]$ are matrices. They are given by (4.26)–(4.28) and (4.17). In particular $[I_z]$ is diagonal.

At this stage we should make the physical interpretation of matrices $D^{(j)}$ more precise. Let $\{|j,m\rangle\}$ be an orthonormal basis of the state space of a spin j particle, so that $|j,m\rangle$ represents the state where the spin projection along the $\hat{z}$ axis is equal to $m\hbar$. Let I_x, I_y and I_z be operators whose matrices in that basis are given by $[I_x]$, $[I_y]$ and $[I_z]$. Define an operator $\mathcal{D}^{(j)}(\hat{n};\theta)$ as

$$\mathcal{D}^{(j)}(\hat{n};\theta) = \exp\{-\mathrm{i}(\hat{n}\theta)\cdot \boldsymbol{I}\}. \tag{14.19}$$

Clearly

$$D^{(j)}_{mm'}(\hat{n};\theta) = \langle j,m|\mathcal{D}^{(j)}(\hat{n};\theta)|j,m'\rangle. \tag{14.20}$$

Let $\hat{u}$ be a unit vector and let $\hat{u}' = O(\hat{n};\theta)\hat{u}$. Let $|j,m;\hat{u}\rangle$ be a vector representing a state where the spin component along $\hat{u}$ is equal to $m\hbar$, i.e. an eigenvector of operator $\hat{u}\cdot\boldsymbol{I}$ with eigenvalue m. Let $|j,m;\hat{u}'\rangle$ be an eigenvector of $\hat{u}'\cdot\boldsymbol{I}$ with eigenvalue m. We will show that[5]

---

[5] Note that $|j,m;\hat{u}'\rangle$ is defined up to a phase only, which corresponds to the fact that the rotation $O(\hat{n};\theta)$ relating $\hat{u}'$ to $\hat{u}$ is not unique.

$$|j, m; \hat{u}'\rangle = \mathcal{D}^{(j)}(\hat{n}, \theta)|j, m; \hat{u}\rangle . \tag{14.21}$$

In particular,

$$\begin{aligned} |j, m; \hat{z}'\rangle &= \mathcal{D}^{(j)}(\hat{n}; \theta)|j, m; \hat{z}\rangle = \mathcal{D}^{(j)}(\hat{n}; \theta)|j, m\rangle \\ &= \sum_{m'} |j, m'\rangle\langle j, m'|\mathcal{D}^{(j)}(\hat{n}; \theta)|j, m\rangle \\ &= \sum_{m'} D^{(j)}_{m'm}(\hat{n}; \theta)|j, m'\rangle . \end{aligned} \tag{14.22}$$

Thus matrices $D^{(j)}$ are coefficients of the expansion of vectors $|j, m; \hat{z}'\rangle$ as linear combinations of vectors $|j, m\rangle$.[6]

To prove property (14.21) we establish the following lemma, which we will use repeatedly.

Lemma Let $\hat{u}' = O(\hat{n}; \theta)\hat{u}$. Then

$$\mathcal{D}^{(j)}(\hat{n}; \theta)\,(\hat{u} \cdot \boldsymbol{I})\left\{\mathcal{D}^{(j)}(\hat{n}; \theta)\right\}^{-1} = \hat{u}' \cdot \boldsymbol{I} . \tag{14.23}$$

Proof Because a finite rotation is the product of a large number of infinitesimal rotations, it is enough to prove (14.23) for a rotation $\delta\theta$. Using the result of Exercise (4.6) one finds

$$\begin{aligned} &\mathcal{D}^{(j)}(\hat{n}; \delta\theta)\,(\hat{u} \cdot \boldsymbol{I})\left\{\mathcal{D}^{(j)}(\hat{n}; \delta\theta)\right\}^{-1} \\ &\quad = \exp\{-\mathrm{i}(\hat{n}\delta\theta) \cdot \boldsymbol{I}\}\,(\hat{u} \cdot \boldsymbol{I}) \exp\{\mathrm{i}(\hat{n}\delta\theta) \cdot \boldsymbol{I}\} \\ &\quad = \hat{u} \cdot \boldsymbol{I} - \mathrm{i}\delta\theta\,[\hat{n} \cdot \boldsymbol{I}, \hat{u} \cdot \boldsymbol{I}] \\ &\quad = \{\hat{u} + \delta\theta(\hat{n} \times \hat{u})\} \cdot \boldsymbol{I} \\ &\quad = \hat{u}' \cdot \boldsymbol{I} . \end{aligned}$$

♣

From (14.23) one has

$$\begin{aligned} (\hat{u}' \cdot \boldsymbol{I})\left\{\mathcal{D}^{(j)}(\hat{n}; \theta)|j, m; \hat{u}\rangle\right\} &= \mathcal{D}^{(j)}(\hat{n}; \theta)\,(\hat{u} \cdot \boldsymbol{I})\,|j, m; \hat{u}\rangle \\ &= m\left\{\mathcal{D}^{(j)}(\hat{n}; \theta)|j, m; \hat{u}\rangle\right\} , \end{aligned}$$

from which (14.21) follows.

Let $\boldsymbol{r}' = O(\hat{n}; \theta)\boldsymbol{r}$ and $\hat{u}' = O(\hat{n}; \theta)\hat{u}$. If the vectors $|j, m\rangle$ are defined in the space of kets $|\boldsymbol{r}\rangle$, then clearly

$$\langle \boldsymbol{r}'|j, m; \hat{u}'\rangle = \langle \boldsymbol{r}|j, m; \hat{u}\rangle . \tag{14.24}$$

---

[6] Matrix $D^{(j)}$ coincides with matrix U in (4.3). It allows to compute the probability that a spin j atom goes through arbitrarily oriented Stern–Gerlach devices.

This is possible only if $\mathcal{D}^{(j)}(\hat{n};\theta)$ coincides with the appropriate restriction of the operator $\mathcal{R}(\hat{n};\theta)$ defined in Sect. 13.2.

Let us now go back to the computation of matrices $D^{(j)}$ defined in (14.13). Since $[I_z]$ is diagonal one can write, owing to (4.26),

$$[\exp(-\mathrm{i}\alpha I_z)]_{mm'} = \mathrm{e}^{-\mathrm{i}m\alpha}\delta_{mm'}. \tag{14.25}$$

Thus knowledge of $D^{(j)}(\alpha,\beta,\gamma)$ boils down to evaluating the exponential of $(i\beta I_y)$.[7] Define a matrix $d^{(j)}$ as

$$d^{(j)}_{mm'}(\beta) = [\exp(-\mathrm{i}\beta I_y)]_{mm'} = D^{(j)}_{mm'}(0,\beta,0). \tag{14.26}$$

We then have

$$D^{(j)}_{mm'}(\alpha,\beta,\gamma) = \mathrm{e}^{-\mathrm{i}m\gamma}d^{(j)}_{mm'}(\beta)\mathrm{e}^{-\mathrm{i}m'\alpha}. \tag{14.27}$$

Let us turn to matrices $D^{(j)}(\hat{n};\theta)$. Let $(\bar{\theta},\bar{\phi})$ be the polar and azimuthal angles of the unit vector $\hat{n}$ with respect to $\hat{z}$. By applying (14.23) twice one can see that[8]

$$\hat{n}\cdot\boldsymbol{I} = \exp(-\mathrm{i}\bar{\phi}I_z)\exp(-\mathrm{i}\bar{\theta}I_y)I_z\exp(\mathrm{i}\bar{\theta}I_y)\exp(\mathrm{i}\bar{\phi}I_z). \tag{14.28}$$

Making use of the result of Exercise (2.16) one gets

$$\begin{aligned}&\exp\{-\mathrm{i}(\hat{n}\theta)\cdot\boldsymbol{I}\}\\ &\quad = \exp(-\mathrm{i}\bar{\phi}I_z)\exp(-\mathrm{i}\bar{\theta}I_y)\exp(-\mathrm{i}\theta I_z)\exp(\mathrm{i}\bar{\theta}I_y)\exp(\mathrm{i}\bar{\phi}I_z)\,.\end{aligned} \tag{14.29}$$

Matrices $D^{(j)}(\hat{n};\theta)$, therefore, also follow from the evaluation of $d^{(j)}$.

We now compute matrices $d^{(j)}(\beta)$ explicitly. For this purpose it is convenient to introduce a representation of the Lie algebra of the rotation group in terms of differential operators. Let ξ_+ and ξ_- be two complex variables and let ∂_+ and ∂_- be defined as

$$\partial_+ \equiv \frac{\partial}{\partial\xi_+}, \qquad \partial_- \equiv \frac{\partial}{\partial\xi_-}. \tag{14.30}$$

Let us make the following correspondence:

---

[7] For simplicity, square brackets are again suppressed on matrices $[I_i]$. The context will tell if I_i is the generator of the abstract operator $\mathcal{D}^{(j)}$ or of the matrix $D^{(j)}$.

[8] Equation (14.28) can also be obtained in another way. Note that in the case where $j = 1/2$, i.e. $\boldsymbol{I} = \boldsymbol{\sigma}/2$, that equation was proved in Sect. 4.6. From (2.58) one can see that the right-hand side can be written entirely in terms of commutators of components of $\boldsymbol{I}$. Since commutators are the same for all j, (14.28) is true for every j. A similar argument can be used to prove (14.23) without employing infinitesimal transformations.

$$\begin{aligned} I_+ &\to \xi_+\partial_- \,, \\ I_- &\to \xi_-\partial_+ \,, \\ I_z &\to \frac{1}{2}(\xi_+\partial_+ - \xi_-\partial_-) \,. \end{aligned} \tag{14.31}$$

It is easy to check that this correspondence preserves the commutation relations (4.19) of operators I_+, I_- and I_z. Thus we have here a representation of the Lie algebra of the rotation group in terms of differential operators. One checks that

$$\begin{aligned} I_x &= \frac{1}{2}(I_+ + I_-) \quad \to \frac{1}{2}(\xi_-\partial_+ + \xi_+\partial_-) \,, \\ I_y &= -\frac{\mathrm{i}}{2}(I_+ - I_-) \to \frac{\mathrm{i}}{2}(\xi_-\partial_+ - \xi_+\partial_-) \,, \\ \boldsymbol{I}\cdot\boldsymbol{I} &= I_x^2 + I_y^2 + I_z^2 \quad \to \Xi(\Xi+1) \,, \end{aligned} \tag{14.32}$$

where a differential operator Ξ was defined as

$$\Xi = \frac{1}{2}(\xi_+\partial_+ + \xi_-\partial_-) \,. \tag{14.33}$$

Monomials ξ_+ and ξ_- can be associated with vectors $|\frac{1}{2},\frac{1}{2}\rangle$ and $|\frac{1}{2},-\frac{1}{2}\rangle$ of the spin 1/2 representation. Indeed

$$\begin{aligned} I_+\xi_+ &= 0 \,, & I_+\xi_- &= \xi_+ \,, \\ I_-\xi_+ &= \xi_- \,, & I_-\xi_- &= 0 \,, \\ I_z\xi_+ &= \tfrac{1}{2}\xi_+ \,, & I_z\xi_- &= -\tfrac{1}{2}\xi_- \,, \\ \boldsymbol{I}\cdot\boldsymbol{I}\,\xi_+ &= \tfrac{3}{4}\xi_+ \,, & \boldsymbol{I}\cdot\boldsymbol{I}\,\xi_- &= \tfrac{3}{4}\xi_- \,. \end{aligned}$$

Denote by ξ'_+ and ξ'_- the variables obtained from ξ_+ and ξ_- through a rotation by β about the y axis. Since $D^{(j)}(0,\beta,0) = d^{(j)}(\beta)$ one must have

$$\begin{aligned} \xi'_+ &= d^{(1/2)}_{++}(\beta)\,\xi_+ + d^{(1/2)}_{-+}(\beta)\,\xi_- \,, \\ \xi'_- &= d^{(1/2)}_{+-}(\beta)\,\xi_+ + d^{(1/2)}_{--}(\beta)\,\xi_- \,. \end{aligned}$$

But matrices $d^{(1/2)}$ can also be extracted from (4.65). Thus

$$\xi'_+ = \cos\frac{\beta}{2}\,\xi_+ + \sin\frac{\beta}{2}\,\xi_- \,, \tag{14.34}$$

$$\xi'_- = -\sin\frac{\beta}{2}\,\xi_+ + \cos\frac{\beta}{2}\,\xi_- \,. \tag{14.35}$$

With the help of ξ_+ and ξ_- one gets a new realization of the spin 1/2 representation. But with them one can also build the representation for arbitrary spin j. The latter is defined on the set of monomials $(\xi_+)^p(\xi_-)^q$, where $p+q=2j$. Indeed consider the correspondence

$$|j,m\rangle \;\rightarrow\; \{(j+m)!(j-m)!\}^{-1/2}\xi_+^{j+m}\xi_-^{j-m}. \tag{14.36}$$

One checks (Exercise 14.3) that the action of differential operators I_+, I_- and I_z on monomials indeed corresponds to matrix elements given in (4.26)–(4.28).

From the previous relation and the definition of ξ'_+ and ξ'_- one must have

$$\frac{(\xi'_+)^{j+m'}(\xi'_-)^{j-m'}}{\{(j+m')!(j-m')!\}^{1/2}} = \sum_{m=-j}^{j} d^{(j)}_{mm'}(\beta)\frac{(\xi_+)^{j+m}(\xi_-)^{j-m}}{\{(j+m)!(j-m)!\}^{1/2}}. \tag{14.37}$$

On the other hand, from the binomial expansion one shows that (Exercise 14.4)

$$\begin{aligned}
&\{(j+m')!(j-m')!\}^{-1/2}(\xi'_+)^{j+m'}(\xi'_-)^{j-m'}\\
&\quad= \{(j+m')!(j-m')!\}^{-1/2}\left\{\cos\frac{\beta}{2}\,\xi_+ + \sin\frac{\beta}{2}\,\xi_-\right\}^{j+m'}\\
&\qquad\times\left\{-\sin\frac{\beta}{2}\,\xi_+ + \cos\frac{\beta}{2}\,\xi_-\right\}^{j-m'}\\
&\quad= \sum_{m=-j}^{j}\left\{\left[\frac{(j+m)!(j-m)!}{(j+m')!(j-m')!}\right]^{1/2}\sum_{\lambda}\binom{j+m'}{j-m-\lambda}\binom{j-m'}{\lambda}\right.\\
&\qquad\left.\times(-1)^{j-m'-\lambda}\left(\cos\frac{\beta}{2}\right)^{m'+m+2\lambda}\left(\sin\frac{\beta}{2}\right)^{2j-m'-m-2\lambda}\right\}\\
&\qquad\times\frac{(\xi_+)^{j+m}(\xi_-)^{j-m}}{\{(j+m)!(j-m)!\}^{1/2}}.
\end{aligned} \tag{14.38}$$

Use was made of the notation

$$\binom{p}{q} = \begin{cases}\frac{p!}{q!(p-q)!} & \text{if} \quad 0\le q\le p\,,\\ 0 & \text{otherwise}\,.\end{cases} \tag{14.39}$$

Comparing (14.37) and (14.38) one obtains for matrices $d^{(j)}_{mm'}(\beta)$

$$\begin{aligned}
d^{(j)}_{mm'}(\beta) &= \left[\frac{(j+m)!(j-m)!}{(j+m')!(j-m')!}\right]^{1/2}\sum_{\lambda}\binom{j+m'}{j-m-\lambda}\\
&\quad\times\binom{j-m'}{\lambda}(-1)^{j-m'-\lambda}\left(\cos\frac{\beta}{2}\right)^{m'+m+2\lambda}\left(\sin\frac{\beta}{2}\right)^{2j-m'-m-2\lambda}
\end{aligned} \tag{14.40}$$

The variable λ assumes all values for which the factorials' arguments are non-negative integers. Comparison of (14.40) and (7.176) yields matrices $d^{(j)}_{mm'}(\beta)$ in terms of Jacobi polynomials as

$$d^{(j)}_{mm'}(\beta) = \left[\frac{(j+m)!(j-m)!}{(j+m')!(j-m')!}\right]^{1/2} \left(\cos\frac{\beta}{2}\right)^{m+m'} \left(-\sin\frac{\beta}{2}\right)^{m-m'} \times P^{(m-m', m+m')}_{j-m}(\cos\beta)\,. \tag{14.41}$$

14.3 *D* Matrices as Wave Functions

In classical mechanics the configuration of a rigid body is completely determined by its center-of-mass position and its orientation. In the absence of external forces the center-of-mass coordinates can be eliminated by an appropriate choice of reference frame. Rigid-body configurations are then in one-to-one correspondence with spatial rotations, i.e. elements of SO(3).

The wave function of a quantum system is defined in the configuration space of the corresponding classical system. In this section we will parametrize rotations by means of Euler angles. The wave function of the quantum rigid body can thus be written as $\psi(\alpha, \beta, \gamma)$.[9] It is very useful to normalize it in the following way:

$$\frac{1}{8\pi^2}\int_0^{2\pi} \mathrm{d}\alpha \int_0^{\pi} \sin\beta\, \mathrm{d}\beta \int_0^{2\pi} \mathrm{d}\gamma\, |\psi(\alpha,\beta,\gamma)|^2 = 1. \tag{14.42}$$

One can show ([223], Sect. 9.2) that with this choice of volume element, the integral is invariant under a redefinition of origin in configuration space.

We now construct a representation of the rotation group in the space of functions $\psi(\alpha, \beta, \gamma)$. Let $O(\bar\alpha, \bar\beta, \bar\gamma)$ be an element of SO(3) and let $\mathcal{R}(\bar\alpha, \bar\beta, \bar\gamma)$ be the operator defined so that[10]

$$\mathcal{R}(\bar\alpha,\bar\beta,\bar\gamma)\psi(\alpha,\beta,\gamma) = \psi'(\alpha,\beta,\gamma) = \psi\left[O^{-1}(\bar\alpha,\bar\beta,\bar\gamma)(\alpha,\beta,\gamma)\right]. \tag{14.43}$$

Here $O^{-1}(\bar\alpha,\bar\beta,\bar\gamma)(\alpha,\beta,\gamma)$ stands for the point in configuration space that parametrizes the rotation $O^{-1}(\bar\alpha,\bar\beta,\bar\gamma)O(\alpha,\beta,\gamma)$. In other words operator $\mathcal{R}$ transforms a function ψ into a function ψ' which coincides with ψ evaluated at a new value of its argument.

One easily checks that (14.43) does define a representation, in the sense that the product of two operators $\mathcal{R}(O_2)\mathcal{R}(O_1)$ coincides with $\mathcal{R}(O_2 O_1)$.

Let us compute generators $\mathcal{I}_x$, $\mathcal{I}_y$ and $\mathcal{I}_z$ in representation $\mathcal{R}$. For this purpose we first note that, owing to (14.3),

---

[9] One could introduce kets $|\alpha,\beta,\gamma\rangle$ so that $\psi(\alpha,\beta,\gamma) = \langle\alpha,\beta,\gamma|\psi\rangle$ but we will work with wave functions only.

[10] To prevent any ambiguity, we should point out that operators $\mathcal{R}(\bar\alpha,\bar\beta,\bar\gamma)$ do not act in the space of kets $|\boldsymbol{r}\rangle$ and therefore do not coincide with operators $\mathcal{R}$ in Sect. 13.2. We use $\mathcal{R}$ to denote reducible representations of the rotation group, while operators $\mathcal{D}^{(j)}$ and matrices $D^{(j)}$ denote irreducible representations.

$$\begin{aligned}\left\{1+\delta\gamma\frac{\partial}{\partial\gamma}\right\}\psi(\alpha,\beta,\gamma) &= \psi(\alpha,\beta,\gamma+\delta\gamma)\\ &= \psi\left\{\left[O(\alpha,\beta,\gamma+\delta\gamma)O^{-1}(\alpha,\beta,\gamma)\right](\alpha,\beta,\gamma)\right\}\\ &= \mathcal{R}\left\{O(\alpha,\beta,\gamma)O^{-1}(\alpha,\beta,\gamma+\delta\gamma)\right\}\psi(\alpha,\beta,\gamma)\\ &= \mathcal{R}\left\{O(\hat{z};-\delta\gamma)\right\}\psi(\alpha,\beta,\gamma)\\ &= (I+\mathrm{i}\delta\gamma\mathcal{I}_z)\psi(\alpha,\beta,\gamma)\,.\end{aligned}$$

Hence one gets

$$\frac{\partial}{\partial\gamma} = \mathrm{i}\mathcal{I}_z\,. \tag{14.44}$$

Similarly

$$\begin{aligned}\left\{1+\delta\beta\frac{\partial}{\partial\beta}\right\}\psi(\alpha,\beta,\gamma) &= \mathcal{R}\left\{O(\alpha,\beta,\gamma)O^{-1}(\alpha,\beta+\delta\beta,\gamma)\right\}\psi(\alpha,\beta,\gamma)\\ &= \mathcal{R}\left\{O(\hat{z};\gamma)O(\hat{y};-\delta\beta)O^{-1}(\hat{z};\gamma)\right\}\psi(\alpha,\beta,\gamma)\,.\end{aligned}$$

Noting that $O(\hat{y};-\delta\beta) = I+\mathrm{i}\delta\beta I_y$ and using (14.23), true for any representation, one sees that

$$O(\hat{z};\gamma)O(\hat{y};-\delta\beta)O^{-1}(\hat{z};\gamma) = I+\mathrm{i}\delta\beta(\cos\gamma\, I_y-\sin\gamma\, I_x).$$

Hence[11]

$$\left\{1+\delta\beta\frac{\partial}{\partial\beta}\right\}\psi(\alpha,\beta,\gamma) = \left\{I+\mathrm{i}\delta\beta(\cos\gamma\,\mathcal{I}_y-\sin\gamma\,\mathcal{I}_x)\right\}\psi(\alpha,\beta,\gamma),$$

from which we get

$$\frac{\partial}{\partial\beta} = \mathrm{i}(\cos\gamma\,\mathcal{I}_y-\sin\gamma\,\mathcal{I}_x)\,. \tag{14.45}$$

By a similar argument one finds that (Exercise 14.5)

$$\frac{\partial}{\partial\alpha} = \mathrm{i}(\cos\beta\,\mathcal{I}_z+\sin\gamma\,\sin\beta\,\mathcal{I}_y+\cos\gamma\,\sin\beta\,\mathcal{I}_x). \tag{14.46}$$

Generators $\mathcal{I}_x$, $\mathcal{I}_y$ and $\mathcal{I}_z$ are now obtained by inverting (14.44), (14.45) and (14.46):

$$\mathcal{I}_x = -\mathrm{i}\left\{\frac{\cos\gamma}{\sin\beta}\frac{\partial}{\partial\alpha}-\sin\gamma\frac{\partial}{\partial\beta}-\cos\gamma\,\cot\beta\frac{\partial}{\partial\gamma}\right\}, \tag{14.47}$$

$$\mathcal{I}_y = -\mathrm{i}\left\{\frac{\sin\gamma}{\sin\beta}\frac{\partial}{\partial\alpha}+\cos\gamma\frac{\partial}{\partial\beta}-\sin\gamma\,\cot\beta\frac{\partial}{\partial\gamma}\right\}, \tag{14.48}$$

$$\mathcal{I}_z = -\mathrm{i}\frac{\partial}{\partial\gamma}\,. \tag{14.49}$$

---

[11] As elsewhere, we use the same symbol I to denote the identity operator in different spaces.

It is easy to check that these operators obey the commutation relations (4.10)–(4.12). Evaluating the sum of squares of generators one finds that

$$\begin{aligned}\mathcal{I}\cdot\mathcal{I} &= \mathcal{I}_x^2+\mathcal{I}_y^2+\mathcal{I}_z^2\\ &= -\frac{\partial^2}{\partial\beta^2}-\cot\beta\frac{\partial}{\partial\beta}-\frac{1}{\sin^2\beta}\left\{\frac{\partial^2}{\partial\alpha^2}+\frac{\partial^2}{\partial\gamma^2}-2\cos\beta\frac{\partial^2}{\partial\alpha\partial\gamma}\right\}.\end{aligned} \tag{14.50}$$

One can see from (14.27) and (14.49) that the $D^{(j)}_{mm'}(\alpha,\beta,\gamma)$ are eigenfunctions of $\mathcal{I}_z$ with eigenvalue $-m$. Let us show that the $D^{(j)}_{mm'}$ are also eigenfunctions of $\mathcal{I}\cdot\mathcal{I}$. For this purpose we establish the following lemma.

Lemma Let I_x, I_y and I_z be generators of the spin j representation of the rotation group, and let

$$\mathcal{D}^{(j)}(\alpha,\beta,\gamma)=\exp(-\mathrm{i}\gamma I_z)\exp(-\mathrm{i}\beta I_y)\exp(-\mathrm{i}\alpha I_z).$$

Then

$$\mathcal{I}\cdot\mathcal{I}\mathcal{D}^{(j)}(\alpha,\beta,\gamma)=\mathcal{D}^{(j)}(\alpha,\beta,\gamma)\boldsymbol{I}\cdot\boldsymbol{I}. \tag{14.51}$$

Proof Let us apply operator (14.50) on $\mathcal{D}^{(j)}(\alpha,\beta,\gamma)$. One finds that

$$\begin{aligned}&\mathcal{I}\cdot\mathcal{I}\,\mathcal{D}^{(j)}(\alpha,\beta,\gamma)\\ &= -\left\{\frac{\partial^2}{\partial\beta^2}+\cot\beta\frac{\partial}{\partial\beta}+\frac{1}{\sin^2\beta}\left[\frac{\partial^2}{\partial\alpha^2}+\frac{\partial^2}{\partial\gamma^2}-2\cos\beta\frac{\partial^2}{\partial\alpha\partial\gamma}\right]\right\}\\ &\quad\times \mathrm{e}^{-\mathrm{i}\gamma I_z}\mathrm{e}^{-\mathrm{i}\beta I_y}\mathrm{e}^{-\mathrm{i}\alpha I_z}\\ &= -\mathrm{e}^{-\mathrm{i}\gamma I_z}\left\{\mathrm{e}^{-\mathrm{i}\beta I_y}(-I_y^2)+\cot\beta\,\mathrm{e}^{-\mathrm{i}\beta I_y}(-\mathrm{i}I_y)\right.\\ &\quad+\frac{1}{\sin^2\beta}\left[\mathrm{e}^{-\mathrm{i}\beta I_y}(-I_z^2)-I_z^2\mathrm{e}^{-\mathrm{i}\beta I_y}\right]\\ &\quad\left.-\frac{2\cos\beta}{\sin^2\beta}(-\mathrm{i}I_z)\mathrm{e}^{-\mathrm{i}\beta I_y}(-\mathrm{i}I_z)\right\}\mathrm{e}^{-\mathrm{i}\alpha I_z}\\ &= \mathcal{D}^{(j)}(\alpha,\beta,\gamma)\left\{\mathrm{e}^{\mathrm{i}\alpha I_z}I_y^2\mathrm{e}^{-\mathrm{i}\alpha I_z}+\mathrm{i}\cot\beta\,\mathrm{e}^{\mathrm{i}\alpha I_z}I_y\mathrm{e}^{-\mathrm{i}\alpha I_z}+\frac{1}{\sin^2\beta}I_z^2\right.\\ &\quad+\frac{1}{\sin^2\beta}\mathrm{e}^{\mathrm{i}\alpha I_z}\mathrm{e}^{\mathrm{i}\beta I_y}I_z^2\mathrm{e}^{-\mathrm{i}\beta I_y}\mathrm{e}^{-\mathrm{i}\alpha I_z}\\ &\quad\left.-\frac{2\cos\beta}{\sin^2\beta}\mathrm{e}^{\mathrm{i}\alpha I_z}\mathrm{e}^{\mathrm{i}\beta I_y}I_z\mathrm{e}^{-\mathrm{i}\beta I_y}\mathrm{e}^{-\mathrm{i}\alpha I_z}I_z\right\}.\end{aligned} \tag{14.52}$$

From (14.23) one can see that

$$\begin{aligned}\mathrm{e}^{\mathrm{i}\alpha I_z}I_y\mathrm{e}^{-\mathrm{i}\alpha I_z}&=\sin\alpha\,I_x+\cos\alpha\,I_y\,,\\ \mathrm{e}^{\mathrm{i}\alpha I_z}\mathrm{e}^{\mathrm{i}\beta I_y}I_z\mathrm{e}^{-\mathrm{i}\beta I_y}\mathrm{e}^{-\mathrm{i}\alpha I_z}&=-\sin\beta\,\cos\alpha\,I_x+\sin\beta\,\sin\alpha\,I_y+\cos\beta\,I_z\,.\end{aligned}$$

Furthermore

$$\mathrm{e}^{\mathrm{i}\alpha I_z} I_y^2 \mathrm{e}^{-\mathrm{i}\alpha I_z} = \left(\mathrm{e}^{\mathrm{i}\alpha I_z} I_y \mathrm{e}^{-\mathrm{i}\alpha I_z}\right)^2 ,$$

and similarly with the term involving I_z^2. Substituting these relations in (14.52) and manipulating, one gets the result. ♣

Let us now apply $\boldsymbol{\mathcal{I}} \cdot \boldsymbol{\mathcal{I}}$ on $D^{(j)}_{mm'}(\alpha, \beta, \gamma)$. From (14.20) one has

$$\begin{aligned} \boldsymbol{\mathcal{I}} \cdot \boldsymbol{\mathcal{I}}\, D^{(j)}_{mm'}(\alpha,\beta,\gamma) &= \boldsymbol{\mathcal{I}} \cdot \boldsymbol{\mathcal{I}}\, \langle j, m | \mathcal{D}^{(j)}(\alpha,\beta,\gamma) | j, m' \rangle \\ &= \langle j, m | \left\{ \boldsymbol{\mathcal{I}} \cdot \boldsymbol{\mathcal{I}}\, \mathcal{D}^{(j)}(\alpha,\beta,\gamma) \right\} | j, m' \rangle \\ &= \langle j, m | \mathcal{D}^{(j)}(\alpha,\beta,\gamma) \boldsymbol{I} \cdot \boldsymbol{I} | j, m' \rangle \\ &= j(j+1) \langle j, m | \mathcal{D}^{(j)}(\alpha,\beta,\gamma) | j, m' \rangle \\ &= j(j+1) D^{(j)}_{mm'}(\alpha,\beta,\gamma) \,. \end{aligned} \tag{14.53}$$

Thus the functions $D^{(j)}_{mm'}(\alpha, \beta, \gamma)$ are wave functions of the quantum rigid body, representing states wherein the total angular momentum and its component along z are well-defined.

In Chap. 20 we shall come back to rigid-body states, in the context of molecules. We close this section by giving a few additional properties of matrices $D^{(j)}$.[12]

Let l be an integer. From (14.27) one sees at once that $D^{(l)}_{m0}(\alpha, \beta, \gamma)$ does not depend on α. One can thus write

$$-\left\{ \frac{\partial^2}{\partial \beta^2} + \cot\beta \frac{\partial}{\partial \beta} + \frac{1}{\sin^2\beta} \frac{\partial^2}{\partial \gamma^2} \right\} D^{(l)}_{m0}(\alpha,\beta,\gamma) = l(l+1) D^{(l)}_{m0}(\alpha,\beta,\gamma). \tag{14.54}$$

Comparing this equation with (7.81) and (7.83) and remembering that $D^{(l)}_{m0}(\alpha, \beta, \gamma)$ depends on γ only through the factor $\mathrm{e}^{-\mathrm{i}m\gamma}$, one concludes that $D^{(l)}_{m0}(\alpha, \beta, \gamma)$ is proportional to $Y^*_{lm}(\beta, \gamma)$. In fact one can show that

$$D^{(l)}_{m0}(\alpha,\beta,\gamma) = \left(\frac{4\pi}{2l+1} \right)^{1/2} Y^*_{lm}(\beta,\gamma). \tag{14.55}$$

Furthermore

$$D^{(l)}_{0m}(\alpha,\beta,\gamma) = (-1)^m \left(\frac{4\pi}{2l+1} \right)^{1/2} Y^*_{lm}(\beta,\alpha). \tag{14.56}$$

---

[12] Proofs can be found in [73] and [223]. We point out that our matrices $D^{(j)}_{mm'}(\alpha, \beta, \gamma)$ and $d^{(j)}_{mm'}(\beta)$ are denoted by $D^{(j)}_{mm'}(-\alpha, -\beta, -\gamma)$ and $d^{(j)}_{mm'}(-\beta)$ in [73], and by $\mathrm{i}^{m'-m} D^{(j)}_{mm'}(\gamma, \beta, \alpha)$ and $d^{(j)}_{mm'}(\beta)$ in [223].

To the spherical harmonics' orthogonality relation there correspond the following properties of the functions $D^{(j)}$, for integral or half-integral j:

$$\begin{aligned}\int_0^{2\pi} \mathrm{d}\alpha \int_0^{\pi} \sin\beta \, \mathrm{d}\beta \int_0^{2\pi} \mathrm{d}\gamma \left[D^{(j_1)}_{m_1 m_1'}(\alpha,\beta,\gamma)\right]^* D^{(j_2)}_{m_2 m_2'}(\alpha,\beta,\gamma) \\ = \frac{8\pi^2}{2j_1+1} \delta_{j_1 j_2} \delta_{m_1 m_2} \delta_{m_1' m_2'} \, . \qquad (14.57)\end{aligned}$$

14.4 Interaction of Two Angular Momenta

Let two quantum systems carry angular momenta $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$. Let $|j_1, m_1\rangle$ stand for basis vectors of the first system's (angular-momentum) state space $\mathcal{V}^1$ and let $|j_2, m_2\rangle$ stand for basis vectors of $\mathcal{V}^2$.

The appropriate tool to deal with the total system is the tensor product formalism. The space $\mathcal{V} = \mathcal{V}^1 \otimes \mathcal{V}^2$ has a basis of vectors $|j_1, m_1\rangle \otimes |j_2, m_2\rangle$, that will more often be denoted by $|j_1, j_2; m_1, m_2\rangle$. These are eigenvectors of the operators

$$\begin{aligned} J_{1z} \otimes I(2) &: \; m_1 \hbar \, , \\ \boldsymbol{J}_1^2 \otimes I(2) &: \; j_1(j_1+1)\hbar^2 \, , \\ I(1) \otimes J_{2z} &: \; m_2 \hbar \, , \\ I(1) \otimes \boldsymbol{J}_2^2 &: \; j_2(j_2+1)\hbar^2 \, , \end{aligned}$$

where corresponding eigenvalues are also indicated.

Section 13.7 shows that one can define in $\mathcal{V}$ a representation of the rotation group whose generators are given by

$$\boldsymbol{J} = \boldsymbol{J}_1 \otimes I(2) + I(1) \otimes \boldsymbol{J}_2 \, . \qquad (14.58)$$

Operator $\boldsymbol{J}$ is naturally interpreted as the total system's total angular momentum. Obviously operators J_x, J_y and J_z obey the commutation relations of the Lie algebra of the rotation group, and $\boldsymbol{J}^2$ commutes with each component. Note that

$$\boldsymbol{J}^2 = \boldsymbol{J}_1^2 \otimes I(2) + I(1) \otimes \boldsymbol{J}_2^2 + 2\boldsymbol{J}_1 \otimes \boldsymbol{J}_2 \, , \qquad (14.59)$$

where

$$\boldsymbol{J}_1 \otimes \boldsymbol{J}_2 = J_{1x} \otimes J_{2x} + J_{1y} \otimes J_{2y} + J_{1z} \otimes J_{2z} \, . \qquad (14.60)$$

Basis vectors $|j_1, j_2; m_1, m_2\rangle$ are eigenvectors of J_z, but not of $\boldsymbol{J}^2$. On the other hand, since $\mathcal{V}$ is a representation space, one can certainly decompose it into a direct sum of subspaces, in each of which the rotation group is represented irreducibly. This actually means finding a new basis of $\mathcal{V}$ whose vectors, denoted by $|j, m\rangle$, are eigenvectors of $\boldsymbol{J}^2$ and of J_z:

$$J_z|j,m\rangle = m\hbar|j,m\rangle\,,$$
$$\boldsymbol{J}^2|j,m\rangle = j(j+1)\hbar^2|j,m\rangle\,.$$

Here $-j \leq m \leq j$, and j assumes in general different values. Vectors $|j,m\rangle$ represent states wherein $\boldsymbol{J}^2, J_z, \boldsymbol{J}_1^2$ and $\boldsymbol{J}_2^2$ are well-defined, while vectors $|j_1,j_2;m_1,m_2\rangle$ represent states wherein $J_{1z}, J_{2z}, \boldsymbol{J}_1^2$ and $\boldsymbol{J}_2^2$ are well-defined.

To illustrate the relevance of the foregoing discussion we consider two noninteracting particles with masses μ_1 and μ_2, in spherically symmetric potentials. The total system's Hamiltonian can be written as

$$\begin{aligned} H(1,2) &= H(1)\otimes I(2) + I(1)\otimes H(2) \\ &= \left(\frac{1}{2\mu_1}\boldsymbol{P}_1^2 + V_1(R_1)\right) + \left(\frac{1}{2\mu_2}\boldsymbol{P}_2^2 + V_2(R_2)\right), \end{aligned} \tag{14.61}$$

where on the last line identity operators were suppressed. If $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ denote the two particles' orbital angular momentum operators, then clearly

$$[\boldsymbol{L}_1, H(1,2)] = 0\,, \tag{14.62}$$

$$[\boldsymbol{L}_2, H(1,2)] = 0\,, \tag{14.63}$$

$$[\boldsymbol{L}_1 + \boldsymbol{L}_2, H(1,2)] = [\boldsymbol{L}, H(1,2)] = 0\,. \tag{14.64}$$

Thus the total system's Hamiltonian can be diagonalized simultaneously with operators $L_{1z}, L_{2z}, \boldsymbol{L}_1^2$ and $\boldsymbol{L}_2^2$ or with operators $L_z, \boldsymbol{L}^2, \boldsymbol{L}_1^2$ and $\boldsymbol{L}_2^2$. Eigenvectors of H can be picked indifferently in the form $|l_1,l_2;m_{l_1},m_{l_2}\rangle$ or in the form $|l,m_l\rangle$.

Suppose now that particles interact by means of a term like $C\boldsymbol{L}_1\cdot\boldsymbol{L}_2$, where C is a constant. Such an interaction term may not be very realistic if one deals with two isolated particles but for electrons in the same orbital, the approximation may be adequate. One shows that (Exercise 14.7)

$$[\boldsymbol{L}_1, C\boldsymbol{L}_1\cdot\boldsymbol{L}_2] \neq 0 \neq [\boldsymbol{L}_2, C\boldsymbol{L}_1\cdot\boldsymbol{L}_2]\,, \tag{14.65}$$

$$\left[\boldsymbol{L}_1^2, C\boldsymbol{L}_1\cdot\boldsymbol{L}_2\right] = 0 = \left[\boldsymbol{L}_2^2, C\boldsymbol{L}_1\cdot\boldsymbol{L}_2\right], \tag{14.66}$$

$$[\boldsymbol{L}_1 + \boldsymbol{L}_2, C\boldsymbol{L}_1\cdot\boldsymbol{L}_2] = 0\,. \tag{14.67}$$

Thus angular momentum components of the two particles no longer separately commute with the Hamiltonian. Components of total angular momentum, however, still commute. This means that eigenvectors of H must be picked in the form $|l,m_l\rangle$ rather than in the form $|l_1,l_2;m_{l_1},m_{l_2}\rangle$.

The case of a spin s particle in a spherically symmetric potential leads to the same conclusion. Although there is only one physical system, the particle has a spin and an orbital angular momentum, so that the tensor product formalism applies. If the Hamiltonian is given by

$$H_0 = \frac{1}{2\mu}\boldsymbol{P}^2 + V(R)\,, \tag{14.68}$$

then components of $\boldsymbol{S}$ and $\boldsymbol{L}$ all commute with H. Defining $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$, one can write eigenvectors of H in the form $|j, m\rangle$ or in the form $|l, s; m_l, m_s\rangle$. Suppose now that the Hamiltonian has a spin–orbit interaction term equal to $\xi(R)\boldsymbol{L}\cdot\boldsymbol{S}$. Then $\boldsymbol{L}$ and $\boldsymbol{S}$ no longer separately commute with the Hamiltonian, whereas $\boldsymbol{L} + \boldsymbol{S}$ still commutes. Only vectors like $|j, m\rangle$ are then appropriate to write down the Hamiltonian's eigenvectors.

14.5 Total Angular Momentum

Let us now turn to the determination of vectors $|j, m\rangle$ in terms of vectors $|j_1, j_2; m_1, m_2\rangle$. Recall that the former are eigenvectors of $\boldsymbol{J}_1^2$, $\boldsymbol{J}_2^2$, $\boldsymbol{J}^2$ and J_z, while the latter are eigenvectors of $\boldsymbol{J}_1^2$, $\boldsymbol{J}_2^2$, J_{1z} and J_{2z}.

We first note that the $|j_1, j_2; m_1, m_2\rangle$ themselves are eigenvectors of J_z, corresponding to the eigenvalue $(m_1 + m_2)\hbar$. Indeed

$$\begin{aligned} J_z|j_1, j_2; m_1, m_2\rangle &= (J_{1z} + J_{2z})|j_1, m_1\rangle \otimes |j_2, m_2\rangle \\ &= \{J_{1z}|j_1, m_1\rangle\} \otimes |j_2, m_2\rangle + |j_1, m_1\rangle \otimes \{J_{2z}|j_2, m_2\rangle\} \\ &= (m_1 + m_2)\hbar|j_1, m_1\rangle \otimes |j_2, m_2\rangle \,. \end{aligned} \tag{14.69}$$

Moreover, it is clear that only linear combinations of vectors $|j_1, j_2; m_1, m_2\rangle$ for which $m_1 + m_2 = m$, with m fixed, are eigenvectors of J_z. Thus one can write

$$\begin{aligned} |j, m\rangle &= \sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} |j_1, j_2; m_1, m_2\rangle\langle j_1, j_2; m_1, m_2|j, m\rangle \\ &= \sum_{m_1+m_2=m} \langle j_1, j_2; m_1, m_2|j, m\rangle|j_1, j_2; m_1, m_2\rangle \,, \end{aligned} \tag{14.70}$$

where the double summation on m_1 and m_2 is restricted to values of m_1 and m_2 such that $m_1 + m_2 = m$. The numbers $\langle j_1, j_2; m_1, m_2|j, m\rangle$ are called *Clebsch–Gordan coefficients.* They are elements of a $(2j_1 + 1)(2j_2 + 1) \times (2j_1 + 1)(2j_2 + 1)$ matrix with rows and columns labeled by pairs (m_1, m_2) and (j, m), respectively. Since these coefficients transform an orthonormal basis into another, the matrix is unitary.

There is only one vector $|j_1, j_2; m_1, m_2\rangle$ for which $m_1 + m_2 = j_1 + j_2$. It is $|j_1, j_2; j_1, j_2\rangle$, which corresponds to the maximum value of m. This vector must therefore coincide with a $|j, m\rangle$ for which $m = j = j_1 + j_2$, that is,

$$|j, j\rangle = |j_1, j_2; j_1, j_2\rangle \,. \tag{14.71}$$

To illustrate the determination of other $|j, m\rangle$ vectors we will work with an example. Let $j_1 = 1$ and $j_2 = 1/2$. Then

$$|\tfrac{3}{2}, \tfrac{3}{2}\rangle = |1, \tfrac{1}{2}; 1, \tfrac{1}{2}\rangle \,.$$

There are four vectors $|j,m\rangle$ for which $j=3/2$. One can find them all by applying operator $J_- = J_{1-} + J_{2-}$ to both sides of the previous equation. For $|\frac{3}{2},\frac{1}{2}\rangle$ one has

$$\begin{aligned}
|\tfrac{3}{2},\tfrac{1}{2}\rangle &= \left(\hbar\sqrt{3}\right)^{-1} J_-|\tfrac{3}{2},\tfrac{3}{2}\rangle \\
&= \left(\hbar\sqrt{3}\right)^{-1} \left\{(J_{1-}+J_{2-})|1,\tfrac{1}{2};1,\tfrac{1}{2}\rangle\right\} \\
&= \left(\hbar\sqrt{3}\right)^{-1} \left\{(J_{1-}|1,1\rangle)\otimes|\tfrac{1}{2},\tfrac{1}{2}\rangle + |1,1\rangle\otimes(J_{2-}|\tfrac{1}{2},\tfrac{1}{2}\rangle)\right\} \\
&= \left(\hbar\sqrt{3}\right)^{-1} \left\{\hbar\sqrt{2}|1,0\rangle\otimes|\tfrac{1}{2},\tfrac{1}{2}\rangle + |1,1\rangle\otimes\hbar|\tfrac{1}{2},-\tfrac{1}{2}\rangle\right\} \\
&= \sqrt{\tfrac{2}{3}}|1,\tfrac{1}{2};0,\tfrac{1}{2}\rangle + \sqrt{\tfrac{1}{3}}|1,\tfrac{1}{2};1,-\tfrac{1}{2}\rangle\,.
\end{aligned}$$

Similarly

$$\begin{aligned}
|\tfrac{3}{2},-\tfrac{1}{2}\rangle &= \sqrt{\tfrac{1}{3}}|1,\tfrac{1}{2};-1,\tfrac{1}{2}\rangle + \sqrt{\tfrac{2}{3}}|1,\tfrac{1}{2};0,-\tfrac{1}{2}\rangle\,, \\
|\tfrac{3}{2},-\tfrac{3}{2}\rangle &= |1,\tfrac{1}{2};-1,-\tfrac{1}{2}\rangle\,.
\end{aligned}$$

From the six basis vectors $|1,\frac{1}{2};m_1,m_2\rangle$, we extracted four orthonormal linear combinations, all of them eigenvectors of operators $\boldsymbol{J}^2$ and J_z. The linear combination of $|1,\frac{1}{2};0,\frac{1}{2}\rangle$ and $|1,\frac{1}{2};1,-\frac{1}{2}\rangle$ which is orthogonal to $|\frac{3}{2},\frac{1}{2}\rangle$ must be a vector of the form $|\frac{1}{2},\frac{1}{2}\rangle$. Thus we can write

$$|\tfrac{1}{2},\tfrac{1}{2}\rangle = -\sqrt{\tfrac{1}{3}}|1,\tfrac{1}{2};0,\tfrac{1}{2}\rangle + \sqrt{\tfrac{2}{3}}|1,\tfrac{1}{2};1,-\tfrac{1}{2}\rangle.$$

Should $|\frac{1}{2},\frac{1}{2}\rangle$ be multiplied by a factor $e^{i\alpha}$, the orthogonality condition on $|\frac{1}{2},\frac{1}{2}\rangle$ and $|\frac{3}{2},\frac{1}{2}\rangle$ would still be met. That factor is determined by specifying that $\langle 1,\frac{1}{2};1,-\frac{1}{2}|\frac{1}{2},\frac{1}{2}\rangle$ is real and positive. Finally, $|\frac{1}{2},-\frac{1}{2}\rangle$ is obtained through application of operator $J_- = J_{1-} + J_{2-}$ on the last equation. One gets

$$|\tfrac{1}{2},-\tfrac{1}{2}\rangle = -\sqrt{\tfrac{2}{3}}|1,\tfrac{1}{2};-1,\tfrac{1}{2}\rangle + \sqrt{\tfrac{1}{3}}|1,\tfrac{1}{2};0,-\tfrac{1}{2}\rangle.$$

Thus addition of two angular momenta $j_1 = 1$ and $j_2 = 1/2$ yields a total angular momentum $j = 3/2$ or $j = 1/2$. In group-theoretical language, the tensor product of representations $j_1 = 1$ and $j_2 = 1/2$ decomposes into two irreducible representations for which $j = 3/2$ and $j = 1/2$.

Let us now go back to the general case. The tensor product of representations j_1 and j_2 decomposes into $2\times\min(j_1,j_2)+1$ irreducible representations, where $\min(j_1,j_2)$ is the smaller of j_1 and j_2. Values of j go from $|j_1-j_2|$ to j_1+j_2 (by unit steps), and there is exactly one irreducible representation for each value of j. As it should be, the number of $|j,m\rangle$ vectors is equal to

$$\sum_{j=|j_1-j_2|}^{j_1+j_2} (2j+1) = (2j_1+1)(2j_2+1). \tag{14.72}$$

Clebsch–Gordan coefficients, in the general case, are obtained by the method just illustrated. Repeated application of operator $J_- = J_{1-} + J_{2-}$ to the vector $|j_1+j_2, j_1+j_2\rangle = |j_1, j_2; j_1, j_2\rangle$ yields all vectors $|j_1+j_2, m\rangle$. Next one finds the vector $|j_1+j_2-1, j_1+j_2-1\rangle$ by applying the orthogonality condition to $|j_1+j_2, j_1+j_2-1\rangle$, and the convention that $\langle j_1, j_2; j_1, j-j_1|j, j\rangle$ is real and positive. The same procedure is followed till all $|j, m\rangle$ vectors have been found.

14.6 Clebsch–Gordan Coefficients

The sets $\{|j, m\rangle\}$ and $\{|j_1, j_2; m_1, m_2\rangle\}$ make up two orthonormal bases. One can thus write

$$\sum_{j=|j_1-j_2|}^{j_1+j_2} \sum_{m=-j}^{j} \langle j_1, j_2; m_1, m_2|j, m\rangle\langle j, m|j_1, j_2; m_1', m_2'\rangle = \delta_{m_1 m_1'}\delta_{m_2 m_2'} , \tag{14.73}$$

$$\sum_{m_1=-j_1}^{j_1} \sum_{m_2=-j_2}^{j_2} \langle j, m|j_1, j_2; m_1, m_2\rangle\langle j_1, j_2; m_1, m_2|j', m'\rangle = \delta_{mm'}\delta_{jj'} . \tag{14.74}$$

These relations express the unitarity of the Clebsch–Gordan coefficient matrix, a property that was mentioned in Sect. 14.5.

We understand that in (14.74) values of j are constrained by the inequality $|j_1 - j_2| \leq j \leq j_1 + j_2$, and by the fact that j differs from $j_1 + j_2$ by an integer. Actually, Clebsch–Gordan coefficients vanish unless these conditions are satisfied. They also vanish if $m \neq m_1+m_2$. This means that the coefficient matrix is block-diagonal, each block corresponding to matrix elements labeled by a given value of m.

With the phase conventions we use, Clebsch–Gordan coefficients are all real. To see this recall that the action of J_+ and J_- on a vector $|j, m\rangle$ yields a real multiple of $|j, m\pm 1\rangle$, and that $\langle j_1, j_2; j_1, j-j_1|j, j\rangle$ is real and positive.[13]

One can show (Exercise 14.14) that Clebsch–Gordan coefficients satisfy the following recurrence relations:

$$\begin{aligned}\langle j_1, j_2; m_1, m_2|j, m+1\rangle &= \{(j-m)(j+m+1)\}^{-1/2} \\ &\times \Big[\{(j_1+m_1)(j_1-m_1+1)\}^{1/2}\langle j_1, j_2; m_1-1, m_2|j, m\rangle \\ &\quad + \{(j_2+m_2)(j_2-m_2+1)\}^{1/2}\langle j_1, j_2; m_1, m_2-1|j, m\rangle\Big] ,\end{aligned} \tag{14.75}$$

---

[13] Hence the symbol $\langle jm|j_1, j_2; m_1, m_2\rangle$ in (14.73) and (14.74), which stands for the complex conjugate of the coefficient $\langle j_1, j_2; m_1, m_2|j, m\rangle$, is in fact equal to it.

$$\begin{aligned}\langle j_1, j_2; m_1, m_2 | j, m-1\rangle &= \{(j+m)(j-m+1)\}^{-1/2} \\ &\times \Big[\{(j_1-m_1)(j_1+m_1+1)\}^{1/2}\langle j_1, j_2; m_1+1, m_2 | j, m\rangle \\ &\quad + \{(j_2-m_2)(j_2+m_2+1)\}^{1/2}\langle j_1, j_2; m_1, m_2+1 | j, m\rangle\Big] .\end{aligned} \tag{14.76}$$

The first relation is obtained by applying operator $J_+ = J_{1+} + J_{2+}$ to both sides of (14.70) and rearranging. The second one is obtained by applying $J_- = J_{1-} + J_{2-}$.

The recurrence relations, orthogonality conditions and phase conventions lead, through a rather involved calculation, to an explicit formula for the Clebsch–Gordan coefficients ([59], Sect. 9.11; [73], Sect. 3.6):

$$\begin{aligned}&\langle j_1, j_2; m_1, m_2 | j, m\rangle \\ &= \delta_{m_1+m_2,m}\left[\frac{(2j+1)(j_1+j_2-j)!(j_1-j_2+j)!(-j_1+j_2+j)!}{(j_1+j_2+j+1)!}\right]^{1/2} \\ &\times\ [(j_1+m_1)!(j_1-m_1)!(j_2+m_2)!(j_2-m_2)!(j+m)!(j-m)!]^{1/2} \\ &\times \sum_k \{(-1)^k k!(j_1+j_2-j-k)!(j_1-m_1-k)!(j_2+m_2-k)! \\ &\qquad \times(j-j_2+m_1+k)!(j-j_1-m_2+k)!\}^{-1} .\end{aligned} \tag{14.77}$$

The variable k is restricted to values for which the factorials' arguments are nonnegative integers.

The Clebsch–Gordan coefficient matrix is given in Fig. 14.1, for small values of j_1 and j_2. More complete tables can be found in [16], p. 172.

Several properties of the Clebsch–Gordan coefficients are more simply expressed through the Wigner $3j$ *symbol* defined as

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = (-1)^{j_1-j_2-m_3}(2j_3+1)^{-1/2}\langle j_1, j_2; m_1, m_2 | j_3, -m_3\rangle. \tag{14.78}$$

One can show ([73], Sect. 3.7) that $3j$ symbols are invariant under even permutations of their columns, and that an odd permutation multiplies then by a factor $(-1)^{j_1+j_2+j_3}$. This implies that

$$\langle j_1, j_2; m_1, m_2 | j, m\rangle = (-1)^{j_1+j_2-j}\langle j_2, j_1; m_2, m_1 | j, m\rangle. \tag{14.79}$$

Hence vectors that result from the coupling of j_1 and j_2 do not always have the same phase as the ones that result from the coupling of j_2 and j_1. Furthermore one has

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = (-1)^{j_1+j_2+j_3}\begin{pmatrix} j_1 & j_2 & j_3 \\ -m_1 & -m_2 & -m_3 \end{pmatrix} . \tag{14.80}$$

$j_1 \times j_2$		J J ⋯ M M ⋯
m_1	m_2	Coefficients
m_1	m_2	
⋮	⋮	

2 × 1/2

m_1	m_2	5/2 5/2
2	1/2	1

m_1	m_2	5/2 3/2	3/2 3/2
2	-1/2	1/5	4/5
1	1/2	4/5	-1/5

m_1	m_2	5/2 1/2	3/2 1/2
1	-1/2	2/5	3/5
0	1/2	3/5	-2/5

m_1	m_2	5/2 -1/2	3/2 -1/2
0	-1/2	3/5	2/5
-1	1/2	2/5	-3/5

m_1	m_2	5/2 -3/2	3/2 -3/2
-1	-1/2	4/5	1/5
-2	1/2	1/5	-4/5

m_1	m_2	5/2 -5/2
-2	-1/2	1

3/2 × 1/2

m_1	m_2	2 2
3/2	1/2	1

m_1	m_2	2 1	1 1
3/2	-1/2	1/4	3/4
1/2	1/2	3/4	-1/4

m_1	m_2	2 0	1 0
1/2	-1/2	1/2	1/2
-1/2	1/2	1/2	-1/2

m_1	m_2	2 -1	1 -1
-1/2	-1/2	3/4	1/4
-3/2	1/2	1/4	-3/4

m_1	m_2	2 -2
-3/2	-1/2	1

1/2 × 1/2

m_1	m_2	1 1
1/2	1/2	1

m_1	m_2	1 0	0 0
1/2	-1/2	1/2	1/2
-1/2	1/2	1/2	-1/2

m_1	m_2	1 -1
-1/2	-1/2	1

1 × 1/2

m_1	m_2	3/2 3/2
1	1/2	1

m_1	m_2	3/2 1/2	1/2 1/2
1	-1/2	1/3	2/3
0	1/2	2/3	-1/3

m_1	m_2	3/2 -1/2	1/2 -1/2
0	-1/2	2/3	1/3
-1	1/2	1/3	-2/3

m_1	m_2	3/2 -3/2
-1	-1/2	1

3/2 × 1

m_1	m_2	5/2 5/2
3/2	1	1

m_1	m_2	5/2 3/2	3/2 3/2
3/2	0	2/5	3/5
1/2	1	3/5	-2/5

m_1	m_2	5/2 1/2	3/2 1/2	1/2 1/2
3/2	-1	1/10	2/5	1/2
1/2	0	3/5	1/15	-1/3
-1/2	1	3/10	-8/15	1/6

m_1	m_2	5/2 -1/2	3/2 -1/2	1/2 -1/2
1/2	-1	3/10	8/15	1/6
-1/2	0	3/5	-1/15	-1/3
-3/2	1	1/10	-2/5	1/2

m_1	m_2	5/2 -3/2	3/2 -3/2
-1/2	-1	3/5	2/5
-3/2	0	2/5	-3/5

m_1	m_2	5/2 -5/2
-3/2	-1	1

2 × 1

m_1	m_2	3 3
2	1	1

m_1	m_2	3 2	2 2
2	0	1/3	2/3
1	1	2/3	-1/3

m_1	m_2	3 1	2 1	1 1
2	-1	1/15	1/3	3/5
1	0	8/15	1/6	-3/10
0	1	6/15	-1/2	1/10

m_1	m_2	3 0	2 0	1 0
1	-1	1/5	1/2	3/10
0	0	3/5	0	-2/5
-1	1	1/5	-1/2	3/10

m_1	m_2	3 -1	2 -1	1 -1
0	-1	6/15	1/2	1/10
-1	0	8/15	-1/6	-3/10
-2	1	1/15	-1/3	3/5

m_1	m_2	3 -2	2 -2
-1	-1	2/3	1/3
-2	0	1/3	-2/3

m_1	m_2	3 -3
-2	-1	1

1 × 1

m_1	m_2	2 2
1	1	1

m_1	m_2	2 1	1 1
1	0	1/2	1/2
0	1	1/2	-1/2

m_1	m_2	2 0	1 0	0 0
1	-1	1/6	1/2	1/3
0	0	2/3	0	-1/3
-1	1	1/6	-1/2	1/3

m_1	m_2	2 -1	1 -1
0	-1	1/2	1/2
-1	0	1/2	-1/2

m_1	m_2	2 -2
-1	-1	1

Fig. 14.1. Clebsch–Gordan coefficients. A square root is understood on each coefficient, that is, $-1/3$ means $-\sqrt{1/3}$

The Clebsch–Gordan coefficients can be used to express the product of rotation matrix elements and, consequently, of spherical harmonics. From (14.20) one has

$$D^{(j_1)}_{m_1 m_1'}(\alpha,\beta,\gamma) D^{(j_2)}_{m_2 m_2'}(\alpha,\beta,\gamma)$$
$$= \langle j_1, m_1 | \mathcal{D}^{(j_1)}(\alpha,\beta,\gamma) | j_1, m_1' \rangle \langle j_2, m_2 | \mathcal{D}^{(j_2)}(\alpha,\beta,\gamma) | j_2, m_2' \rangle$$

$$= [\langle j_1, m_1| \otimes \langle j_2, m_2|]\, \mathcal{D}^{(j_1)}(\alpha,\beta,\gamma) \otimes \mathcal{D}^{(j_2)}(\alpha,\beta,\gamma)$$
$$\times\, [|j_1, m_1'\rangle \otimes |j_2, m_2'\rangle]\,.$$

Using the inverse of relation (14.70) and realizing that operators $\mathcal{D}^{(j_1)} \otimes \mathcal{D}^{(j_2)}$ are block-diagonal (one block for each value of j) in the basis of vectors $|j,m\rangle$, one finds that

$$\begin{aligned}
&D^{(j_1)}_{m_1 m_1'}(\alpha,\beta,\gamma) D^{(j_2)}_{m_2 m_2'}(\alpha,\beta,\gamma)\\
&\quad= \left[\sum_{j,m} \langle j_1, j_2; m_1, m_2 | j, m\rangle \langle j, m|\right] \mathcal{D}^{(j_1)}(\alpha,\beta,\gamma) \otimes \mathcal{D}^{(j_2)}(\alpha,\beta,\gamma)\\
&\qquad\times \left[\sum_{j',m'} |j', m'\rangle \langle j', m' | j_1, j_2; m_1', m_2'\rangle\right]\\
&\quad= \sum_{j,m,m'} \langle j_1, j_2; m_1, m_2 | j, m\rangle \langle j, m' | j_1, j_2; m_1', m_2'\rangle D^{(j)}_{mm'}(\alpha,\beta,\gamma)\,.
\end{aligned} \tag{14.81}$$

Letting $m_1' = 0 = m_2'$ and using (14.55) together with the reality of Clebsch–Gordan coefficients, one gets for spherical harmonics

$$\begin{aligned}
Y_{l_1 m_1}(\beta,\gamma) Y_{l_2 m_2}(\beta,\gamma) &= \sum_{l,m} \left[\frac{(2l_1+1)(2l_2+1)}{4\pi(2l+1)}\right]^{1/2}\\
&\quad\times \langle l_1, l_2; m_1, m_2 | l, m\rangle \langle l_1, l_2; 0, 0 | l, 0\rangle Y_{lm}(\beta,\gamma)\,.
\end{aligned} \tag{14.82}$$

We close this section by mentioning briefly the coupling of three angular momenta. The total system's state space has a basis of vectors $|j_1, m_1\rangle \otimes |j_2, m_2\rangle \otimes |j_3, m_3\rangle$. We look for eigenvectors $|j,m\rangle$ of $\boldsymbol{J}\cdot\boldsymbol{J}$ and J_z, where $\boldsymbol{J} = \boldsymbol{J_1} + \boldsymbol{J_2} + \boldsymbol{J_3}$. To obtain them one can proceed in (at least) two different ways:

i) One first couples j_1 and j_2 to obtain vectors $|j', m'\rangle$. Next one couples j' and j_3. Let us denote vectors so obtained by $|[(j_1, j_2)j', j_3]j, m\rangle$.
ii) One first couples j_2 and j_3 to obtain vectors $|j'', m''\rangle$. Next one couples j_1 and j''. Let us denote vectors so obtained by $|[j_1, (j_2, j_3)j'']j, m\rangle$.

These two avenues lead to different results, essentially due to the fact that in general there are more than one $|j,m\rangle$ vector for given values of j and m. The scalar product $\langle [(j_1, j_2)j', j_3]j, m | [j_1, (j_2, j_3)j'']j, m\rangle$, however, does not depend on m. This follows from (14.94), which will be proved in Sect. 14.7. The $6j$ *symbol* is defined so that

$$\begin{aligned}
&\langle [(j_1, j_2)j', j_3]j, m | [j_1, (j_2, j_3)j'']j, m\rangle\\
&\quad= (-1)^{j_1+j_2+j_3+j}\, [(2j'+1)(2j''+1)]^{1/2} \begin{Bmatrix} j_1 & j_2 & j' \\ j_3 & j & j'' \end{Bmatrix}.
\end{aligned} \tag{14.83}$$

Properties of $6j$ symbols, widely used in complex atomic calculations, are developed in [73], Chap. 6 and in [149].

14.7 The Wigner–Eckart Theorem

Let $\boldsymbol{J}$ be the total angular momentum operator of a quantum system. In the state space a rotation $O(\alpha, \beta, \gamma)$ is represented by the operator

$$\mathcal{R}(\alpha,\beta,\gamma) = \exp\left\{-\mathrm{i}\gamma\frac{J_z}{\hbar}\right\}\exp\left\{-\mathrm{i}\beta\frac{J_y}{\hbar}\right\}\exp\left\{-\mathrm{i}\alpha\frac{J_z}{\hbar}\right\}. \qquad (14.84)$$

Thus an active rotation transforms vector $|\psi\rangle$ into $\mathcal{R}|\psi\rangle$ and operator Q into $\mathcal{R}Q\mathcal{R}^{-1}$.

Let $\{Q(p,q)\}$ be a set of $2p+1$ operators labeled by a variable q going from $-p$ to p by unit steps. The set $\{Q(p,q)\}$ makes up an *irreducible tensor operator* of rank p if the following relation is satisfied:

$$\mathcal{R}(\alpha,\beta,\gamma)Q(p,q)\left[\mathcal{R}(\alpha,\beta,\gamma)\right]^{-1} = \sum_{q'=-p}^{p} D^{(p)}_{q'q}(\alpha,\beta,\gamma)Q(p,q'). \qquad (14.85)$$

By restricting (14.85) to infinitesimal rotations one obtains the following commutation relations, which can be used as well to define irreducible tensor operators:

$$[J_z, Q(p,q)] = q\hbar Q(p,q)\,, \qquad (14.86)$$

$$[J_+, Q(p,q)] = \{(p-q)(p+q+1)\}^{1/2}\,\hbar Q(p,q+1)\,, \qquad (14.87)$$

$$[J_-, Q(p,q)] = \{(p+q)(p-q+1)\}^{1/2}\,\hbar Q(p,q-1)\,. \qquad (14.88)$$

There is an obvious analogy between these relations and the result of applying operators J_z, J_+ and J_- on kets $|j,m\rangle$ with $j=p$ and $m=q$. Actually one easily shows (Exercise 14.16) that if $|0,0\rangle$ stands for a state of zero total angular momentum, then $\{|p,q\rangle\langle 0,0|\}$ is an irreducible tensor operator of rank p.

Every operator that commutes with $\boldsymbol{J}$ is an irreducible tensor operator of rank zero. Among irreducible tensor operators of rank 1 one has

$$\left\{\begin{matrix}\frac{-1}{\sqrt{2}}(X+\mathrm{i}Y)\\ Z\\ \frac{1}{\sqrt{2}}(X-\mathrm{i}Y)\end{matrix}\right\},\quad \left\{\begin{matrix}\frac{-1}{\sqrt{2}}(P_x+\mathrm{i}P_y)\\ P_z\\ \frac{1}{\sqrt{2}}(P_x-\mathrm{i}P_y)\end{matrix}\right\},\quad \left\{\begin{matrix}\frac{-1}{\sqrt{2}}J_+\\ J_z\\ \frac{1}{\sqrt{2}}J_-\end{matrix}\right\}, \qquad (14.89)$$

which are the position, momentum and total angular momentum operators. An irreducible tensor operator of rank 2 is given by

$$\left\{\begin{array}{c} \frac{1}{2}(X^2 - Y^2) + \mathrm{i}XY \\ -(X + \mathrm{i}Y)Z \\ \frac{1}{\sqrt{6}}(-X^2 - Y^2 + 2Z^2) \\ (X - \mathrm{i}Y)Z \\ \frac{1}{2}(X^2 - Y^2) - \mathrm{i}XY \end{array}\right\}, \tag{14.90}$$

which is essentially the quadrupole moment operator.

Let $\{Q_1(p_1, q_1)\}$ and $\{Q_2(p_2, q_2)\}$ be two irreducible tensor operators of ranks p_1 and p_2. Starting with Q_1 and Q_2, one can construct irreducible tensor operator $\{Q(p, q)\}$ in the following way:

$$Q(p,q) = \sum_{q_1=-p_1}^{p_1} \sum_{q_2=-p_2}^{p_2} Q_1(p_1, q_1) Q_2(p_2, q_2) \langle p_1, p_2; q_1, q_2 | p, q\rangle . \tag{14.91}$$

This relation means that irreducible tensor operators combine as basis vectors of irreducible representations of the rotation group. The proof that the commutation relations of $Q(p,q)$ with J_z, J_+ and J_- coincide with (14.86)–(14.88) follows from the recurrence relations (14.75) and (14.76) of the Clebsch–Gordan coefficients.

We now derive a very important result about matrix elements of irreducible tensor operators. Let us first write (14.91) for the case where $Q_2(p_2, q_2) = |p_2, q_2\rangle\langle 0, 0|$. One easily shows that for given p,

$$|p,q\rangle = \sum_{q_1} \sum_{q_2} Q_1(p_1, q_1) |p_2, q_2\rangle \left(\langle p_1, p_2; q_1, q_2 | p, q\rangle\right) \tag{14.92}$$

is a basis for the $(2p+1)$ representation of the rotation group. Note, however, that vectors $|p,q\rangle$ are not necessarily normalized, although for a given p they all have the same norm.

The Clebsch–Gordan coefficient matrix is unitary and real. Hence (14.92) can be inverted as

$$Q_1(p_1, q_1)|p_2, q_2\rangle = \sum_{p=|p_1-p_2|}^{p_1+p_2} \sum_{q=-p}^{p} |p,q\rangle \left(\langle p_1, p_2; q_1, q_2 | p, q\rangle\right) . \tag{14.93}$$

Projecting (14.93) on $\langle p', q'|$ one gets

$$\langle p', q' | Q_1(p_1, q_1) | p_2, q_2\rangle = N(p', p_1, p_2, Q_1) \langle p_1, p_2; q_1, q_2 | p', q'\rangle .$$

Here the orthogonality of vectors $|p,q\rangle$ was used and a normalization constant N was introduced as

$$N(p', p_1, p_2, Q_1) = \langle p', q' | p', q'\rangle .$$

The fact that N only depends on the arguments shown follows from (14.92).

The matrix elements of $Q_1(p_1, q_1)$ are more often written as

$$\begin{aligned}
&\langle p,q|Q_1(p_1,q_1)|p_2,q_2\rangle \\
&\quad = (-1)^{p+p_1-p_2}(2p+1)^{-1/2}\langle p||Q_1(p_1)||p_2\rangle\,(\langle p_1,p_2;q_1,q_2|p,q\rangle) \\
&\quad = (-1)^{p-q}\langle p||Q_1(p_1)||p_2\rangle \begin{pmatrix} p_1 & p_2 & p \\ q_1 & q_2 & -q \end{pmatrix}. \qquad (14.94)
\end{aligned}$$

The quantity $\langle p||Q_1(p_1)||p_2\rangle$ is called a *reduced matrix element* and it is directly related to the normalization constant N.

Equation (14.94) embodies the *Wigner–Eckart theorem.* Its meaning is the following. Whenever $|p,q\rangle$ and $|p_2,q_2\rangle$ are associated with irreducible representations of the rotation group, the matrix element $\langle p,q|Q_1(p_1,q_1)|p_2,q_2\rangle$ of an irreducible tensor operator depends on indices q, q_1 and q_2 only through a Clebsch–Gordan coefficient. Knowledge of one nonzero matrix element is enough to get all the others (for p, p_1 and p_2 fixed).

The Wigner–Eckart theorem will be used in Chaps. 18 and 19. Among its numerous applications it often helps to realize that various matrix elements vanish. This happens if the Clebsch–Gordan coefficient does vanish. For example one sees at once that matrix elements of operators $\boldsymbol{J}$, $\boldsymbol{R}$ and $\boldsymbol{P}$ vanish for states of zero total angular momentum.

We should point out that the Wigner–Eckart theorem also applies when kets $|p,q\rangle$ and $|p_2,q_2\rangle$ are eigenvectors of a partial angular momentum operator (orbital angular momentum or spin, for instance), provided that $Q_1(p_1,q_1)$ and that operator have appropriate commutation relations.

14.8 Enlarged Symmetry

In Sect. 13.4 we pointed out that irreducible representations of a Hamiltonian's symmetry group are related to vector spaces associated with degenerate eigenvalues. We now further examine the nature of this relation.

The Hamiltonian of a particle in a spherically symmetric potential can be written as

$$H = \frac{1}{2\mu}\boldsymbol{P}\cdot\boldsymbol{P} + V(R).$$

Eigenfunctions of H are given as $u_{nl}(r)Y_{lm}(\theta,\phi)$. For almost all potentials the quantum number m accounts for the full degeneracy of eigenvalues. And correspondingly, eigenfunctions with the same n and l but different m form a basis for the l-th irreducible representation of SO(3), the symmetry group of H.

For potentials associated with an isotropic harmonic oscillator (R^2) and a hydrogen atom (R^{-1}), however, the magnetic quantum number does not fully explain the degeneracy. We will see that in both cases, the symmetry group of the Hamiltonian is larger than the group of rotations.

The Hamiltonian of an isotropic harmonic oscillator can be written as

$$H = \frac{1}{2\mu}\left\{P_x^2 + P_y^2 + P_z^2\right\} + \frac{1}{2}\mu\omega^2\left\{X^2 + Y^2 + Z^2\right\}. \tag{14.95}$$

Formally, this is really three independent oscillators. Eigenvalues of H are therefore given by (Exercise 7.12)

$$E = \hbar\omega\left(n_x + \frac{1}{2}\right) + \hbar\omega\left(n_y + \frac{1}{2}\right) + \hbar\omega\left(n_z + \frac{1}{2}\right) = \hbar\omega\left(n + \frac{3}{2}\right), \tag{14.96}$$

where $n = n_x + n_y + n_z$ is a nonnegative integer. Thus there is degeneracy with respect to two quantum numbers, and one can see that the degree of degeneracy is equal to $\frac{1}{2}(n+1)(n+2)$. But we already showed in Exercise (13.15) that the Hamiltonian (14.95) is invariant under the group SU(3). In spaces associated with degenerate eigenvalues, SO(3) is represented reducibly and SU(3) irreducibly ([183], Chap. 14; [200], Sect. 30).

Let us now turn to the R^{-1} potential. In terms of relative coordinates and reduced mass the Hamiltonian of a hydrogen atom (or, more generally, of a hydrogen-like ion) can be written as

$$H = \frac{1}{2\mu}\boldsymbol{P}\cdot\boldsymbol{P} - \frac{Ze^2}{R}.$$

Define three operators A_x, A_y and A_z as[14]

$$\boldsymbol{A} = \frac{1}{2\mu}(\boldsymbol{P}\times\boldsymbol{L} - \boldsymbol{L}\times\boldsymbol{P}) - Ze^2\frac{\boldsymbol{R}}{R} \tag{14.97}$$

or, equivalently,

$$A_i = \frac{1}{\mu}\left(\sum_{j,k}\varepsilon_{ijk}P_jL_k - \mathrm{i}\hbar P_i\right) - Ze^2\frac{X_i}{R}. \tag{14.98}$$

One can show that these operators satisfy the following commutation relations:[15]

$$[A_i, H] = 0, \quad [L_i, A_j] = \mathrm{i}\hbar\sum_k \varepsilon_{ijk}A_k, \quad [A_i, A_j] = -\frac{2\mathrm{i}\hbar}{\mu}H\sum_k \varepsilon_{ijk}L_k, \tag{14.99}$$

---

[14] This definition is motivated by the analogy with the Runge–Lenz vector which, in the classical Kepler problem, is conserved ([136], p. 39).

[15] The proof of these relations, although rather cumbersome, follows entirely from the commutation relations of $\boldsymbol{R}$, $\boldsymbol{P}$ and $\boldsymbol{L}$ and presents no conceptual problem. The computation can be based on the properties of the Levi-Civita symbol ([100], Chap. 14) or brought to expressions involving $\boldsymbol{R}$ and $\boldsymbol{P}$ only ([79], App. A). The latter reference gives a detailed discussion of the enlarged symmetry of the R^{-1} potential. See also [241], Chap. 21.

$$\sum_i L_i A_i = \sum_i A_i L_i = 0, \quad \sum_i A_i A_i = (Ze^2)^2 + \frac{2}{\mu} H \left\{ \boldsymbol{L} \cdot \boldsymbol{L} + \hbar^2 \right\}. \tag{14.100}$$

The six operators L_i and A_i commute with the Hamiltonian. Due to the last of (14.99), however, they do not make up a Lie algebra. Nevertheless, one can show that the set of all operators $H^n L_i$ and $H^n A_i$, for all nonnegative integers n, make up a Lie algebra of operators that commute with H. This algebra has infinite dimension.

To avoid dealing with an infinite algebra, let us not focus on the whole state space but rather on a subspace corresponding to a specific eigenvalue of H. Let E be in the discrete spectrum of H and let $\mathcal{V}$ be the space of eigenvectors of H with eigenvalue E. In $\mathcal{V}$ define an operator $\tilde{A}_i$ as

$$\tilde{A}_i = \sqrt{\frac{-\mu}{2E}} A_i . \tag{14.101}$$

Then in the subspace $\mathcal{V}$,

$$[L_i, H] = 0, \qquad [\tilde{A}_i, H] = 0 \tag{14.102}$$

and

$$\begin{aligned} [L_i, L_j] &= \mathrm{i}\hbar \sum_k \varepsilon_{ijk} L_k , \\ [L_i, \tilde{A}_j] &= \mathrm{i}\hbar \sum_k \varepsilon_{ijk} \tilde{A}_k , \\ [\tilde{A}_i, \tilde{A}_j] &= \mathrm{i}\hbar \sum_k \varepsilon_{ijk} L_k . \end{aligned} \tag{14.103}$$

Hence the six operators L_i, $\tilde{A}_i$ ($i = 1, 2, 3$) make up, when restricted to $\mathcal{V}$, a Lie algebra of operators that commute with H. Equations (14.100) become

$$\sum_i L_i \tilde{A}_i = \sum_i \tilde{A}_i L_i = 0 , \tag{14.104}$$

$$\sum_i \tilde{A}_i \tilde{A}_i + \sum_i L_i L_i + \hbar^2 + \frac{\mu}{2E} (Ze^2)^2 = 0 . \tag{14.105}$$

To fully understand the nature of the Lie algebra (14.103), it is useful to define six more operators F_i and G_i as

$$F_i = \frac{1}{2}(L_i - \tilde{A}_i), \qquad G_i = \frac{1}{2}(L_i + \tilde{A}_i). \tag{14.106}$$

Obviously F_i and G_i commute with H. Their commutation relations are given by

$$[F_i, F_j] = \mathrm{i}\hbar \sum_k \varepsilon_{ijk} F_k, \quad [G_i, G_j] = \mathrm{i}\hbar \sum_k \varepsilon_{ijk} G_k, \quad [F_i, G_j] = 0. \tag{14.107}$$

Thus F_i commutes with G_j but the F_i themselves, like the G_i, have the same commutation relations as angular momentum. The Lie algebra therefore coincides with that of the group SO(3) $\otimes$ SO(3).

Irreducible representations of SO(3) $\otimes$ SO(3) are labeled by two numbers f and g which, in general, can independently assume any nonnegative integral or half-integral value. Eigenvalues of $\sum_i F_i F_i$ and $\sum_i G_i G_i$ are equal to $f(f+1)\hbar^2$ and $g(g+1)\hbar^2$ respectively. In the present case, however, f and g are not independent. Indeed

$$\begin{aligned}
&\sum_i F_i F_i - \sum_i G_i G_i \\
&\quad = \frac{1}{4}\sum_i (L_i - \tilde{A}_i)(L_i - \tilde{A}_i) - \frac{1}{4}\sum_i (L_i + \tilde{A}_i)(L_i + \tilde{A}_i) \\
&\quad = -\frac{1}{2}\sum_i (L_i \tilde{A}_i + \tilde{A}_i L_i) = 0\,,
\end{aligned} \tag{14.108}$$

hence $f = g$. On the other hand, (14.105) entails that

$$\begin{aligned}
0 &= \sum_i \tilde{A}_i \tilde{A}_i + \sum_i L_i L_i + \hbar^2 + \frac{\mu}{2E}(Ze^2)^2 \\
&= \sum_i (G_i - F_i)(G_i - F_i) + \sum_i (G_i + F_i)(G_i + F_i) + \hbar^2 + \frac{\mu}{2E}(Ze^2)^2 \\
&= 2\sum_i F_i F_i + 2\sum_i G_i G_i + \hbar^2 + \frac{\mu}{2E}(Ze^2)^2\,.
\end{aligned}$$

This implies that eigenvalues satisfy

$$0 = 4f(f+1)\hbar^2 + \hbar^2 + \frac{\mu}{2E}(Ze^2)^2 = (2f+1)^2\hbar^2 + \frac{\mu}{2E}(Ze^2)^2\,.$$

Hence we get

$$E = -\frac{\mu(Ze^2)^2}{2\hbar^2(2f+1)^2} = -\frac{\mu Z^2 e^4}{2\hbar^2 n^2}\,, \tag{14.109}$$

where n is a positive integer. This coincides with hydrogen-like energy levels (9.20) and correctly reproduces the degree of degeneracy $n^2 = (2f+1)^2$ found in (9.23). Each energy level thus corresponds to only one irreducible representation of SO(3) $\otimes$ SO(3).

Note that even though F_i and G_i can correspond to representations with half-integral spin, the orbital angular momentum $L_i = F_i + G_i$, owing to the rules of addition of angular momenta, corresponds to integral-spin representations only.

Exercises

14.1. Show that any real orthogonal 3×3 matrix with determinant 1 can be parametrized by means of Euler angles.

14.2. Show that any unitary 2×2 matrix with determinant 1 can be written in the form $D^{(1/2)}(\hat{n}, \theta)$, with $n_z \geq 0$ and $0 \leq \theta < 4\pi$.

14.3. Check that the action of differential operators I_+, I_- and I_z on the right-hand side of (14.36) corresponds to matrix elements given by (4.26)–(4.28).

14.4. Use the binomial expansion to obtain the elements of matrix $d^{(j)}(\beta)$.

14.5. Derive (14.46).

14.6.
a) From (14.3) show that if $O(\alpha, \beta, \gamma) = O(\phi_1, \theta_1, 0)O(0, \theta_2, \pi - \phi_2)$ then

$$\cos\beta = \cos\theta_1 \, \cos\theta_2 + \sin\theta_1 \, \sin\theta_2 \, \cos(\phi_1 - \phi_2).$$

b) Use (14.55) and (14.56) and the fact that matrices $D^{(l)}_{mm'}$ represent the rotation group to derive formula (7.174) for the addition of spherical harmonics.

14.7. Derive (14.65)–(14.67).

14.8. Let $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ be spin operators for two spin 1/2 particles. Show that the projectors on states with total spin 1 and 0 are respectively given by

$$P_1 = \frac{3}{4} + \hbar^{-2}\boldsymbol{S}_1 \cdot \boldsymbol{S}_2 \quad \text{and} \quad P_0 = \frac{1}{4} - \hbar^{-2}\boldsymbol{S}_1 \cdot \boldsymbol{S}_2 .$$

14.9. Compute Clebsch–Gordan coefficients for $j_1 = 3/2$ and $j_2 = 1/2$.

14.10. Compute Clebsch–Gordan coefficients for $j_1 = 1$ and $j_2 = 1$.

14.11. Compute Clebsch–Gordan coefficients for $j_1 = 3/2$ and $j_2 = 1$.

14.12. In Exercise (14.11) one finds that $|\frac{5}{2}, \frac{5}{2}\rangle = |\frac{3}{2}, 1; \frac{3}{2}, 1\rangle$. Check that ket $|\frac{5}{2}, \frac{5}{2}\rangle$ is indeed an eigenvector of $\boldsymbol{J}^2$, by letting the operator $(\boldsymbol{J}_1 + \boldsymbol{J}_2)^2$ act on the vector $|\frac{3}{2}, 1; \frac{3}{2}, 1\rangle$.

14.13. Let an angular momentum l be coupled to a spin $s = 1/2$. Show by induction that

$$\begin{aligned} |l + \tfrac{1}{2}, m\rangle = {} & \left\{\frac{l+\frac{1}{2}+m}{2l+1}\right\}^{1/2} |l, \tfrac{1}{2}; m - \tfrac{1}{2}, \tfrac{1}{2}\rangle \\ & + \left\{\frac{l+\frac{1}{2}-m}{2l+1}\right\}^{1/2} |l, \tfrac{1}{2}; m + \tfrac{1}{2}, -\tfrac{1}{2}\rangle , \end{aligned}$$

$$|l-\tfrac{1}{2},m\rangle = -\left\{\tfrac{l+\frac{1}{2}-m}{2l+1}\right\}^{1/2}|l,\tfrac{1}{2};m-\tfrac{1}{2},\tfrac{1}{2}\rangle + \left\{\tfrac{l+\frac{1}{2}+m}{2l+1}\right\}^{1/2}|l,\tfrac{1}{2};m+\tfrac{1}{2},-\tfrac{1}{2}\rangle\,.$$

14.14. Derive the recurrence formulas (14.75) and (14.76).

14.15. Use properties of the Clebsch–Gordan coefficients and $3j$ symbols to show that

$$\langle j,j;m,-m|0,0\rangle = (2j+1)^{-1/2}(-1)^{j-m}\,,$$

$$\sum_{m=-j}^{j}(-1)^{j-m}\langle j,j;m,-m|j',0\rangle = (2j+1)^{1/2}\delta_{0,j'}\,.$$

14.16. Show that $\{|p,q\rangle\langle 0,0|\}$, for p fixed and $-p \le q \le p$, makes up an irreducible tensor operator of rank p.

14.17. Let $\boldsymbol{P}$ and $\boldsymbol{L}$ be the momentum and angular momentum operators of a spinless quantum particle. Use the commutation relations of L_i and P_j to show that $Q(1,q)$ is an irreducible tensor operator of rank 1, where

$$Q(1,1) = \frac{-1}{\sqrt{2}}(P_x + \mathrm{i}P_y)\,,$$
$$Q(1,0) = P_z\,,$$
$$Q(1,-1) = \frac{1}{\sqrt{2}}(P_x - \mathrm{i}P_y)\,.$$

14.18. Let X_q and P_q be the $+1$, 0 and -1 components of the position and momentum operators given in (14.89). Show that

$$\boldsymbol{X}\cdot\boldsymbol{P} = \sum_q (-1)^q X_q P_{-q} = \sum_q X_q P_q^\dagger\,.$$

These relations generalize to any irreducible tensor operator of rank 1 whose Cartesian components are Hermitian.

14.19. Use formula (14.91) to find, from two different copies of the operator $Q(1,q)$ of Exercise (14.17), (i) an irreducible tensor operator of rank 2, (ii) an irreducible tensor operator of rank 1 and (iii) an irreducible tensor operator of rank 0.

14.20. Show that the commutator of operator $J_- = L_-$ with the five components of the quadrupole moment operator (14.90) indeed corresponds to relation (14.88) applied to an irreducible tensor operator of rank 2.

14.21.

a) From results obtained in Sect. 14.8 show that the wave function of the ground state of hydrogen-like ions satisfies the following equations:

$$L_i \psi(\boldsymbol{r}) = 0\,, \qquad A_i \psi(\boldsymbol{r}) = 0\,.$$

b) From these equations derive the form of the wave function.

15 Dirac's Relativistic Equation

The behavior of electrons, in atoms and molecules in particular, is normally described by the Schrödinger equation. For several reasons, however, this is not entirely satisfactory. The Schrödinger equation is not invariant under the coordinate transformations of the special theory of relativity. This means that it cannot correctly account for relativistic effects which, for inner electrons, are often significant. Furthermore, the Schrödinger equation is not really suitable for the introduction of electronic spin. Both shortcomings are fixed by Dirac's relativistic equation. That equation and the associated formalism have many applications in elementary particle theory. It is on the basis of his equation that Dirac predicted the existence of the positron.[1]

15.1 The Lorentz Group

According to the theory of special relativity, there exists a set of equivalent reference frames that move with respect to each other at constant velocities. The Lorentz group is, loosely speaking, the set of coordinate transformations that relate an arbitrary pair of these reference frames.

More precisely, let us introduce a quadruple x^μ ($\mu = 0, 1, 2, 3$), called the *contravariant position four-vector* and defined so that

$$x^0 = ct, \qquad x^1 = x, \qquad x^2 = y, \qquad x^3 = z. \tag{15.1}$$

Let us also introduce a diagonal matrix $g_{\mu\nu}$, called the *covariant metric tensor* and defined so that

$$g_{00} = 1, \qquad g_{11} = g_{22} = g_{33} = -1, \qquad g_{\mu\nu} = 0 \quad \text{if} \quad \mu \neq \nu. \tag{15.2}$$

Let x be the column matrix with elements x^μ and g the square matrix with elements $g_{\mu\nu}$. Consider the bilinear form $x^T g x$, given explicitly by

$$x^T g x = c^2 t^2 - x^2 - y^2 - z^2. \tag{15.3}$$

---

[1] The Dirac equation is discussed in detail in [29], Chaps. 1–5, [69], Chap. 11 and [139], Chaps. 3 and 4. For the historical side see [106].

L. Marchildon, *Quantum Mechanics*

Then the *Lorentz group* is defined as the set of all linear transformations of x into $x' = \Lambda x$ that leave the bilinear form $x^T g x$ invariant, i.e. such that

$$x'^T g x' = x^T \Lambda^T g \Lambda x = x^T g x\,. \tag{15.4}$$

A necessary and sufficient condition for this equation to be satisfied for every four-vector x is that

$$\Lambda^T g \Lambda = g\,. \tag{15.5}$$

In this chapter we will use the common conventions of special relativity. We define a *covariant four-vector* x_μ as[2]

$$x_\mu = \sum_{\nu=0}^{3} g_{\mu\nu} x^\nu \equiv g_{\mu\nu} x^\nu\,. \tag{15.6}$$

A summation from 0 to 3 is understood on any index appearing twice in the same term (once covariantly, one contravariantly). The tensor $g_{\mu\nu}$ and its matrix inverse $g^{\mu\nu}$ (the *contravariant metric tensor*) are used to lower and raise indices, respectively. The tensor $g_\mu{}^\nu$ coincides with the Kronecker delta. Since it is symmetric, it can be written irrespective of the order of indices as g_μ^ν. Components of the transformation equation $x' = \Lambda x$ can be written as

$$x'^\mu = \Lambda^\mu{}_\nu x^\nu\,. \tag{15.7}$$

To understand the nature of the Lorentz group, let us recall Sect. 13.7. Matrix g obviously coincides with what was then denoted by $I(3,1)$. Hence the Lorentz group is isomorphic to the Lie group $\mathrm{O}(3,1)$. In the present context we will rather denote it by L.

The group L can be divided in two disjoint subsets L_+ and L_- according as $\det(\Lambda) = \pm 1$. The subset L_+ is a subgroup of L which coincides with the group $\mathrm{SO}(3,1)$. It is called the *proper Lorentz group.*

There is another useful way to divide the Lorentz group in two disjoint subsets. First note that the (00) component of (15.5) can be written as

$$\left(\Lambda^0{}_0\right)^2 - \sum_{i=1}^{3} \left(\Lambda^i{}_0\right)^2 = 1\,. \tag{15.8}$$

One sees at once that $|\Lambda^0{}_0| \geq 1$. We denote by $L^\uparrow$ the subset of L made of all elements for which $\Lambda^0{}_0 \geq 1$, and by $L^\downarrow$ the subset made of all elements for which $\Lambda^0{}_0 \leq -1$. The set $L^\uparrow$ is a subgroup of L. It is called the *orthochronous Lorentz group.*

---

[2] Thus $x_1 = -x$, $x_2 = -y$ and $x_3 = -z$, in contrast with notations used everywhere else in this book. So is it with all other three-dimensional vectors. Throughout this chapter we work in the coordinate representation.

Let us now apply both divisions together. One sees that the Lorentz group divides in four disjoint subsets that can be denoted by $L_+^\uparrow$, $L_+^\downarrow$, $L_-^\uparrow$ and $L_-^\downarrow$. The subgroup $L_+^\uparrow$ is called the *proper orthochronous Lorentz group.*[3]

Two elements of the Lorentz group are important enough to deserve particular notation:

$$P=\begin{pmatrix}1&0&0&0\\0&-1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix},\qquad T=\begin{pmatrix}-1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}. \tag{15.9}$$

Clearly the action of P on the four-vector x coincides with space inversion, while the action of T on x coincides with time reversal. One easily checks that any element of $L_+^\downarrow$ can be obtained as the product of PT and an element of $L_+^\uparrow$, that any element of $L_-^\uparrow$ can be obtained as the product of P and an element of $L_+^\uparrow$ and that any element of $L_-^\downarrow$ can be obtained as the product of T and an element of $L_+^\uparrow$.

The Lorentz group generators are found by the methods developed in Chap. 13. They are imaginary matrices Σ that satisfy the equation

$$\Sigma^T g+g\Sigma=0\,. \tag{15.10}$$

One can find six linearly independent matrices that satisfy (15.10). It is convenient to label these matrices with pairs $(\kappa\lambda)$ such that

$$\Sigma_{\kappa\lambda}=-\Sigma_{\lambda\kappa}\,,\qquad \kappa,\lambda=0,1,2,3. \tag{15.11}$$

One should note here that $\Sigma_{\kappa\lambda}$ is a matrix, and not the $(\kappa\lambda)$ element of a matrix Σ. In principle the $\Sigma_{\kappa\lambda}$ can be picked in different ways, but the following choice is particularly appropriate:

$$\Sigma_{12}=\begin{pmatrix}0&0&0&0\\0&0&-\mathrm{i}&0\\0&\mathrm{i}&0&0\\0&0&0&0\end{pmatrix},\qquad \Sigma_{23}=\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&-\mathrm{i}\\0&0&\mathrm{i}&0\end{pmatrix}, \tag{15.12}$$

$$\Sigma_{13}=\begin{pmatrix}0&0&0&0\\0&0&0&-\mathrm{i}\\0&0&0&0\\0&\mathrm{i}&0&0\end{pmatrix},\qquad \Sigma_{01}=\begin{pmatrix}0&-\mathrm{i}&0&0\\-\mathrm{i}&0&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix}, \tag{15.13}$$

$$\Sigma_{02}=\begin{pmatrix}0&0&-\mathrm{i}&0\\0&0&0&0\\-\mathrm{i}&0&0&0\\0&0&0&0\end{pmatrix},\qquad \Sigma_{03}=\begin{pmatrix}0&0&0&-\mathrm{i}\\0&0&0&0\\0&0&0&0\\-\mathrm{i}&0&0&0\end{pmatrix}. \tag{15.14}$$

---

[3] The terminology may vary. $L_+^\uparrow$ is sometimes called the *proper Lorentz group.*

Let us denote by $(\Sigma_{\kappa\lambda})^{\mu}{}_{\nu}$ the $(\mu\nu)$ element of the matrix $\Sigma_{\kappa\lambda}$. One easily sees that

$$(\Sigma_{\kappa\lambda})^{\mu}{}_{\nu} = \mathrm{i}(g^{\mu}_{\kappa} g_{\lambda\nu} - g_{\kappa\nu} g^{\mu}_{\lambda})\,. \tag{15.15}$$

From (15.15) or from the explicit form of matrices $\Sigma_{\kappa\lambda}$ one checks that (Exercise 15.1)

$$[\Sigma_{\kappa\lambda}, \Sigma_{\mu\nu}] = -\mathrm{i}(g_{\kappa\mu}\Sigma_{\lambda\nu} + g_{\lambda\nu}\Sigma_{\kappa\mu} - g_{\kappa\nu}\Sigma_{\lambda\mu} - g_{\lambda\mu}\Sigma_{\kappa\nu})\,. \tag{15.16}$$

These commutation relations make up the Lie algebra of the Lorentz group.[4]

The parametrization of a Lorentz transformation close to the identity involves six infinitesimal constants. Again it is convenient to label them with a pair $(\kappa\lambda)$ and to choose them so that $a^{\kappa\lambda} = -a^{\lambda\kappa}$. One then writes

$$\Lambda = I - \frac{\mathrm{i}}{2} a^{\kappa\lambda} \Sigma_{\kappa\lambda}\,, \tag{15.17}$$

the factor 1/2 coming from the double sum which results in each term being counted twice. The effect of an infinitesimal transformation on components x^{μ} can be written as

$$\begin{aligned} x'^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu} &= \left\{ I - \frac{\mathrm{i}}{2} a^{\kappa\lambda} \Sigma_{\kappa\lambda} \right\}^{\mu}{}_{\nu} x^{\nu} \\ &= \left\{ g^{\mu}_{\nu} - \frac{\mathrm{i}}{2} a^{\kappa\lambda} \mathrm{i} \left[g^{\mu}_{\kappa} g_{\lambda\nu} - g_{\kappa\nu} g^{\mu}_{\lambda} \right] \right\} x^{\nu} \\ &= x^{\mu} + a^{\mu}{}_{\nu} x^{\nu}\,. \end{aligned} \tag{15.18}$$

Likewise

$$x'_{\mu} = x_{\mu} + a_{\mu}{}^{\nu} x_{\nu}\,. \tag{15.19}$$

---

[4] From matrices $\Sigma_{\kappa\lambda}$ one can define six matrices F_i and G_i $(i = 1, 2, 3)$ as

$$\begin{aligned} F_1 &= \frac{1}{2}(\Sigma_{23} + \mathrm{i}\Sigma_{01})\,, & F_2 &= \frac{1}{2}(\Sigma_{31} + \mathrm{i}\Sigma_{02})\,, & F_3 &= \frac{1}{2}(\Sigma_{12} + \mathrm{i}\Sigma_{03})\,, \\ G_1 &= \frac{1}{2}(\Sigma_{23} - \mathrm{i}\Sigma_{01})\,, & G_2 &= \frac{1}{2}(\Sigma_{31} - \mathrm{i}\Sigma_{02})\,, & G_3 &= \frac{1}{2}(\Sigma_{12} - \mathrm{i}\Sigma_{03})\,. \end{aligned}$$

It is easy to check that up to a factor the F_i and G_i satisfy the commutation relations (14.107), i.e. they make up two independent copies of the Lie algebra of the rotation group. This does not entail that the Lie algebra of the Lorentz group is isomorphic to that of SO(3) × SO(3), since F_i and G_i (due to imaginary factors) are not generators of the Lorentz group. Nevertheless, the finite-dimensional irreducible representations of the Lorentz group can be built from pairs of irreducible representations of SO(3) ([183], Chap. 22).

Looking back at Sect. 14.8 we realize that in vector spaces associated with positive values of the energy of hydrogen, the symmetry algebra is isomorphic to that of the Lorentz group.

Clearly generators Σ_{12}, Σ_{23} and Σ_{13} correspond to spatial rotations. These relate reference frames at rest with respect to each other. Σ_{01}, Σ_{02} and Σ_{03} generate *special Lorentz transformations*, which relate reference frames moving with respect to each other. A finite Lorentz transformation can be written as

$$\Lambda = \exp\left\{-\frac{\mathrm{i}}{2}a^{\kappa\lambda}\Sigma_{\kappa\lambda}\right\}. \tag{15.20}$$

Here constants $a^{\mu\nu}$ are finite. The exponential is explicitly calculated in Exercise (15.2), for special Lorentz transformations.

Any matrix written in the form (15.20) represents a proper orthochronous Lorentz transformation. Indeed Λ can be transformed into the identity in a continuous way. Conversely, one can show that any element of $L_+^\uparrow$ can be written in the form (15.20).

Any quadruple A^μ that transforms like x^μ is called a *contravariant four-vector*, whereas any quadruple B_μ that transforms like x_μ is called a *covariant four-vector.* Matrices Λ, which transform A^μ into A'^μ, make up a faithful irreducible representation of the Lorentz group called the *fundamental representation.* Obviously, matrices $g\Lambda g$ transform B_μ into B'_μ and make up an equivalent representation. Any expression of the kind $A^\mu B_\mu$ is a scalar, i.e. it is invariant under Lorentz transformations.

A *contravariant tensor of rank n* is an object like $T^{\kappa\lambda\mu\nu\cdots}$ (n exponents) which, under Lorentz transformations, transforms like $A^\kappa A^\lambda A^\mu A^\nu \ldots$ One similarly defines *covariant* and mixed tensor. All these objects make up vector spaces on which the Lorentz group is represented, in general reducibly.

It is not difficult to show that derivatives $\partial/\partial x_\mu$ transform as a contravariant four-vector, while derivatives $\partial/\partial x^\mu$ transform as a covariant four-vector. Let ϕ be a scalar function of x^μ. One easily shows that the equation

$$\frac{\partial^2\phi}{\partial x_\mu \partial x^\mu}+\left(\frac{mc}{\hbar}\right)^2\phi \equiv \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2}-\frac{\partial^2\phi}{\partial x^2}-\frac{\partial^2\phi}{\partial y^2}-\frac{\partial^2\phi}{\partial z^2}+\left(\frac{mc}{\hbar}\right)^2\phi = 0, \tag{15.21}$$

where m has dimensions of a mass, is invariant under Lorentz transformations. This is the *Klein–Gordon* equation. On the other hand, one is easily convinced that the Schrödinger equation is not invariant under Lorentz transformations.

The Klein–Gordon equation, even though invariant under Lorentz transformations, is not suitable for the description of electrons. Indeed it doesn't involve spin variables in any way. In Sect. 15.2 we will develop a representation of the Lorentz group that includes the spin 1/2 representation of the rotation group. With its help we will construct a Lorentz-invariant equation capable of describing the electronic spin.

15.2 The Dirac Equation

Different approaches can be used to build the required representation of the Lorentz group. The origin and motivation of the one chosen here may not be entirely obvious. It does, however, have the advantage of being generalizable to all groups SO(p,q).

The starting point consists in assuming that four matrices γ_μ ($\mu = 0$, 1, 2, 3) exist which satisfy the following relations:

$$\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2g_{\mu\nu}I\,. \tag{15.22}$$

Here I is the identity matrix with the same dimension as matrices γ_μ. Note that γ_μ and γ_ν anticommute, i.e. $\gamma_\nu\gamma_\mu = -\gamma_\mu\gamma_\nu$ if $\mu \neq \nu$. Eventually, we will provide an explicit realization of the γ_μ in terms of 4 × 4 matrices. For the time being, however, we will not specify them further.

From the γ_μ let us define six matrices $\sigma_{\mu\nu}$ as

$$\sigma_{\mu\nu} = \frac{\mathrm{i}}{2}[\gamma_\mu, \gamma_\nu]\,. \tag{15.23}$$

Obviously $\sigma_{\mu\nu} = -\sigma_{\nu\mu}$, and $\sigma_{\mu\nu}$ vanishes if $\mu = \nu$. Note that one can write

$$\sigma_{\mu\nu} = \mathrm{i}(\gamma_\mu\gamma_\nu - g_{\mu\nu}I)\,. \tag{15.24}$$

Let us now evaluate the commutator of matrices $\sigma_{\kappa\lambda}$ and γ_μ. We have

$$\begin{aligned}
[\sigma_{\kappa\lambda}, \gamma_\mu] &= \mathrm{i}\,[\gamma_\kappa\gamma_\lambda - g_{\kappa\lambda}I, \gamma_\mu] \\
&= \mathrm{i}(\gamma_\kappa\gamma_\lambda\gamma_\mu - \gamma_\mu\gamma_\kappa\gamma_\lambda) \\
&= \mathrm{i}\,\{\gamma_\kappa(-\gamma_\mu\gamma_\lambda + 2g_{\lambda\mu}) - (-\gamma_\kappa\gamma_\mu + 2g_{\kappa\mu})\gamma_\lambda\} \\
&= 2\mathrm{i}(g_{\lambda\mu}\gamma_\kappa - g_{\kappa\mu}\gamma_\lambda)\,.
\end{aligned} \tag{15.25}$$

From this result the commutator of $\sigma_{\kappa\lambda}$ and $\sigma_{\mu\nu}$ follows:

$$\begin{aligned}
[\sigma_{\kappa\lambda}, \sigma_{\mu\nu}] &= \mathrm{i}\,[\sigma_{\kappa\lambda}, \gamma_\mu\gamma_\nu - g_{\mu\nu}I] \\
&= \mathrm{i}\,\{\gamma_\mu\,[\sigma_{\kappa\lambda}, \gamma_\nu] + [\sigma_{\kappa\lambda}, \gamma_\mu]\,\gamma_\nu\} \\
&= \mathrm{i}(2\mathrm{i})\,\{\gamma_\mu\,(g_{\lambda\nu}\gamma_\kappa - g_{\kappa\nu}\gamma_\lambda) + (g_{\lambda\mu}\gamma_\kappa - g_{\kappa\mu}\gamma_\lambda)\,\gamma_\nu\} \\
&= (-2\mathrm{i})(g_{\lambda\nu}\sigma_{\kappa\mu} + g_{\kappa\mu}\sigma_{\lambda\nu} - g_{\kappa\nu}\sigma_{\lambda\mu} - g_{\lambda\mu}\sigma_{\kappa\nu})\,.
\end{aligned} \tag{15.26}$$

This is what we were looking for. We have shown that matrices $\sigma_{\mu\nu}/2$ satisfy the commutation relations (15.16) of the Lorentz-group generators. Hence they generate a matrix representation of the Lorentz group. To parameters $a^{\mu\nu}$ there corresponds a matrix M given by

$$M = \exp\left\{-\frac{\mathrm{i}}{4}a^{\mu\nu}\sigma_{\mu\nu}\right\}. \tag{15.27}$$

This is called the *spinor representation*. Vectors on which it acts are called *spinors*.[5]

With the help of spinors we now construct an equation that is invariant under proper orthochronous Lorentz transformations. For this purpose, let us first prove the following lemma.

Lemma Let $\{a^{\kappa\lambda}\}$ be parameters of a Lorentz transformation, A^μ a contravariant vector and M the matrix of the spinor representation defined above. Then

$$M^{-1}\gamma_\mu A'^\mu M = \gamma_\mu A^\mu . \tag{15.28}$$

Proof It is enough to prove the result for an infinitesimal transformation, as a finite transformation is the limit of a product of infinitesimal transformations. If $a^{\kappa\lambda}$ is infinitesimal

$$\begin{aligned} M^{-1}\gamma_\mu A'^\mu M &= \left\{I + \frac{\mathrm{i}}{4}a^{\kappa\lambda}\sigma_{\kappa\lambda}\right\}\gamma_\mu\left\{I - \frac{\mathrm{i}}{4}a^{\kappa\lambda}\sigma_{\kappa\lambda}\right\}(A^\mu + a^\mu{}_\nu A^\nu) \\ &= \left\{\gamma_\mu + \frac{\mathrm{i}}{4}a^{\kappa\lambda}\left[\sigma_{\kappa\lambda}, \gamma_\mu\right]\right\}(A^\mu + a^\mu{}_\nu A^\nu) \\ &= (\gamma_\mu + a_\mu{}^\kappa\gamma_\kappa)(A^\mu + a^\mu{}_\nu A^\nu) \\ &= \gamma_\mu A^\mu . \end{aligned}$$

♣

Let ψ be a spinor which, under Lorentz transformations, transforms as $\psi \to \psi' = M\psi$. Clearly the equation

$$-\mathrm{i}\gamma_\mu\frac{\partial}{\partial x_\mu}\psi + \frac{mc}{\hbar}\psi = 0 \tag{15.29}$$

is invariant under proper orthochronous Lorentz transformations. Indeed multiplying both sides of (15.29) by M one has

$$-\mathrm{i}M\gamma_\mu\frac{\partial}{\partial x_\mu}\left\{M^{-1}M\psi\right\} + \frac{mc}{\hbar}M\psi = 0 .$$

Using the fact that $\partial/\partial x_\mu$ is a contravariant vector we see that

$$-\mathrm{i}\gamma_\mu\frac{\partial}{\partial x'_\mu}\psi' + \frac{mc}{\hbar}\psi' = 0 .$$

Equation (15.29) is called the *Dirac equation*. To see that it is suitable to describe the behavior of electrons, there remains to check that it involves spin 1/2 matrices.

---

[5] Strictly speaking the spinor representation associates two matrices M with each matrix Λ, as was the case with half-integral spin representations of the rotation group. Matrices M written as exponentials represent the proper orthochronous Lorentz group. The transformation of a spinor under space inversion and time reversal will be investigated in Sect. 15.4.

For this purpose we introduce an explicit realization of matrices γ_μ. Let σ^0 be the identity matrix in two dimensions and let σ^i ($i = 1, 2, 3$) be the Pauli spin matrices. It is easy to check that the following choice fulfills all defining conditions (15.22) of the γ_μ:

$$\gamma_0 = \begin{pmatrix} \sigma^0 & 0 \\ 0 & -\sigma^0 \end{pmatrix}, \qquad \gamma_i = \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix}. \tag{15.30}$$

Matrices $\sigma_{\mu\nu}$ are then given by

$$\sigma_{0i} = \begin{pmatrix} 0 & -\mathrm{i}\sigma^i \\ -\mathrm{i}\sigma^i & 0 \end{pmatrix}, \qquad \sigma_{ij} = \begin{pmatrix} \sigma^k & 0 \\ 0 & \sigma^k \end{pmatrix}, \tag{15.31}$$

where, in the equation for σ_{ij}, indices i, j and k are a cyclic permutation of 1, 2 and 3. Thus the rotation group, a subgroup of $L_+^\uparrow$, is generated in the spinor representation by

$$\frac{1}{2}\sigma_{ij} = \begin{pmatrix} \frac{1}{2}\sigma^k & 0 \\ 0 & \frac{1}{2}\sigma^k \end{pmatrix}. \tag{15.32}$$

This means that the spinor representation, when restricted to rotations, involves two spin 1/2 representations.

15.3 Plane-Wave Solution

The spinor ψ that appears in Dirac's equation is a four-component column which, in general, depends on space-time coordinates x_μ. We now look for a plane-wave solution of the equation, that is,

$$\psi(x_\mu) = \exp\left\{-\frac{\mathrm{i}}{\hbar}p^\nu x_\nu\right\} u. \tag{15.33}$$

Here u is a constant spinor and p^ν is a constant four-vector with the dimensions of momentum.

Substituting expression (15.33) in (15.29) one finds that u must satisfy the equation

$$\gamma_\mu p^\mu u - mcu = (\gamma_\mu p^\mu - mcI)u = 0. \tag{15.34}$$

Let the operator $(\gamma_\nu p^\nu + mcI)$ act on both sides of (15.34). One can see that a nontrivial solution for u will be obtained only if

$$p_\mu p^\mu - m^2c^2 = 0. \tag{15.35}$$

This is the relativistic equation for the *momentum four-vector*. Writing

$$p^{\mu} = \left(\frac{E}{c}, p_x, p_y, p_z \right), \tag{15.36}$$

one sees that

$$E = \pm \left\{ c^2 \boldsymbol{p}^2 + m^2 c^4 \right\}^{1/2}. \tag{15.37}$$

Equation (15.34) can be written explicitly as

$$\begin{pmatrix} \frac{E}{c} - mc & 0 & -p_z & -(p_x - \mathrm{i}p_y) \\ 0 & \frac{E}{c} - mc & -(p_x + \mathrm{i}p_y) & p_z \\ p_z & p_x - \mathrm{i}p_y & -\frac{E}{c} - mc & 0 \\ p_x + \mathrm{i}p_y & -p_z & 0 & -\frac{E}{c} - mc \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = 0. \tag{15.38}$$

This equation has a nontrivial solution for u if and only if the determinant of the matrix vanishes. This condition on the determinant coincides with (15.37). For any three-dimensional vector $\boldsymbol{p}$, (15.38) has four linearly independent solutions. The first two correspond to a positive value of E, the last two to a negative value. One can pick the solutions as

$$E > 0: \quad u^{(1)}(\boldsymbol{p}) = \begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E+mc^2} \\ \frac{c(p_x+\mathrm{i}p_y)}{E+mc^2} \end{pmatrix}, \quad u^{(2)}(\boldsymbol{p}) = \begin{pmatrix} 0 \\ 1 \\ \frac{c(p_x-\mathrm{i}p_y)}{E+mc^2} \\ \frac{-cp_z}{E+mc^2} \end{pmatrix}; \tag{15.39}$$

$$E < 0: \quad u^{(3)}(\boldsymbol{p}) = \begin{pmatrix} \frac{-cp_z}{|E|+mc^2} \\ \frac{-c(p_x+\mathrm{i}p_y)}{|E|+mc^2} \\ 1 \\ 0 \end{pmatrix}, \quad u^{(4)}(\boldsymbol{p}) = \begin{pmatrix} \frac{-c(p_x-\mathrm{i}p_y)}{|E|+mc^2} \\ \frac{cp_z}{|E|+mc^2} \\ 0 \\ 1 \end{pmatrix}. \tag{15.40}$$

In the case where $c|\boldsymbol{p}| \ll |E|$, i.e. in the nonrelativistic case, the last two components are small (with respect to one) for positive-energy solutions, and the first two components are small for negative-energy solutions. If $\boldsymbol{p} = 0$ the four solutions simply become

$$u^{(1)}(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad u^{(2)}(0) = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad u^{(3)}(0) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad u^{(4)}(0) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{15.41}$$

Obviously these solutions are eigenvectors of σ_{12}, hence they represent states wherein the z component of spin is well-defined.

One can build solutions of the Dirac equation for which any given spin component is well-defined, by letting act on spinors $u^{(i)}(0)$ the rotation operator

$$\exp\left(-\frac{\mathrm{i}}{4}\sum_{k,l=1}^{3} a^{kl}\sigma_{kl}\right). \tag{15.42}$$

One also checks (Exercise 15.3) that spinors $u^{(i)}(\boldsymbol{p})$ are obtained (up to a constant multiplicative factor) from spinors $u^{(i)}(0)$ through a Lorentz transformation.

We have just found that Dirac's equation has negative-energy solutions. In general the energy spectrum of a quantum system is required to have a lower bound, so that there exists an absolutely stable state. But the Dirac equation has solutions for all negative E such that $E < -mc^2$, hence the spectrum does not have a lower bound. Dirac postulated that in nature, all negative-energy states are occupied. Consequently, a positive-energy electron cannot normally go to a negative-energy level, due to the Pauli exclusion principle. What happens, however, if a negative-energy electron is excited to a positive-energy level? A hole is then created in negative-energy states, which for all practical purposes looks like a positive-energy particle with a charge opposite to that of the electron. Dirac and others suggested that these holes would behave exactly as particles, which they called *positrons*. The positron was discovered experimentally by Anderson in 1932.

In Dirac's theory electron excitation and hole formation are interpreted as the creation of an electron-positron pair, while the falling back of an electron to a negative-energy state is interpreted as pair annihilation. Nowadays quantum field theory no longer views positrons as holes, but as particles very closely related to electrons. Dirac's theory has nevertheless remarkably predicted the existence of a particle with the same mass as that of the electron, and opposite charge.

15.4 Properties of the Dirac Equation

Dirac's equation, developed in Sect. 15.2 for a free particle, can easily be generalized to include the effect of an electromagnetic field. We know that an electromagnetic field is suitably described by a *potential four-vector* $A^\mu = (c^{-1}\phi, A_x, A_y, A_z)$. In general A^μ depends on space-time coordinates x_ν. As with Schrödinger's equation, the field can be incorporated in Dirac's equation through the substitution

$$-\mathrm{i}\hbar\frac{\partial}{\partial x_\mu} \to -\mathrm{i}\hbar\frac{\partial}{\partial x_\mu} + q_e A^\mu, \tag{15.43}$$

where q_e denotes the electron's charge. Dirac's equation in the presence of an electromagnetic field is therefore given by

$$-\mathrm{i}\gamma_\mu\frac{\partial}{\partial x_\mu}\psi + \frac{q_e}{\hbar}\gamma_\mu A^\mu\psi + \frac{mc}{\hbar}\psi = 0. \tag{15.44}$$

Clearly the generalized equation is still invariant under proper orthochronous Lorentz transformations.

Dirac's field ψ has a probabilistic interpretation similar to that of the Schrödinger wave function. To develop this interpretation, let us first note that in the representation we use, matrices γ_μ satisfy

$$\gamma_0 \gamma_\mu^\dagger \gamma_0 = \gamma_\mu \, . \tag{15.45}$$

Taking the matrix adjoint of (15.44) and letting $\bar{\psi} = \psi^\dagger \gamma_0$, one easily sees that

$$\mathrm{i}\left(\frac{\partial}{\partial x_\mu}\bar{\psi}\right)\gamma_\mu + \frac{q_e}{\hbar}A^\mu \bar{\psi}\gamma_\mu + \frac{mc}{\hbar}\bar{\psi} = 0 \, .$$

Let us project this equation on ψ and multiply from the left by $\bar{\psi}$ both sides of (15.44). With some manipulations one gets

$$\frac{\partial}{\partial x_\mu}\left(\bar{\psi}\gamma_\mu\psi\right) = 0 \, . \tag{15.46}$$

This is the continuity equation, which can be written more explicitly as

$$\frac{\partial}{\partial t}\left(\psi^\dagger\psi\right) + \boldsymbol{\nabla}\cdot\boldsymbol{S} = 0 \, . \tag{15.47}$$

Here the three-vector $\boldsymbol{S}$ has components $S^i = -c\bar{\psi}\gamma_i\psi$. Equation (15.47) entails that $\psi^\dagger\psi$ can be associated with a probability density and $\boldsymbol{S}$ with a probability current.

In Sect. 15.1 space inversion P and time reversal T were introduced as Lorentz group elements. Dirac's equation is invariant under space inversion and time reversal, provided the fields ψ and A^μ transform appropriately. For space inversion

$$[x'] = P[x], \qquad [A'] = P[A], \qquad \psi' = \gamma_0\psi, \tag{15.48}$$

while for time reversal

$$[x'] = T[x], \qquad [A'] = -T[A], \qquad \psi' = \gamma_1\gamma_3\psi^* \, . \tag{15.49}$$

Square brackets around x and A emphasize matrix notation. That Dirac's equation is invariant under both these transformations is checked in Exercises (15.5) and (15.6).

The physical significance of space inversion and time reversal is clear. But the Dirac equation is also invariant under another discrete transformation, called *charge conjugation* and given by (Exercise 15.7)

$$[x'] = [x], \qquad [A'] = -[A], \qquad \psi' = -\mathrm{i}\gamma_2\psi^* \, . \tag{15.50}$$

From our explicit matrix representation of γ_2 one can see that charge conjugation transforms a spinor associated with a positive energy into a spinor associated with a negative energy, and conversely. Thus a positron in a field $-A^\mu$ behaves as an electron in a field A^μ.[6]

In the presence of an electromagnetic field the Dirac equation, like Schrödinger's, is invariant under a gauge transformation given by (Exercise 15.8)

$$\psi' = \exp\left\{\mathrm{i}\frac{q_e\chi}{\hbar}\right\}\psi, \qquad A'_\mu = A_\mu - \frac{\partial\chi}{\partial x^\mu}. \tag{15.51}$$

Here χ is an arbitrary function of x, y, z and t.

It is useful to define a matrix γ_5 as $\gamma_5 = -\mathrm{i}\gamma_0\gamma_1\gamma_2\gamma_3$. For any μ and ν

$$\gamma_5\gamma_\mu + \gamma_\mu\gamma_5 = 0, \qquad [\gamma_5, \sigma_{\mu\nu}] = 0. \tag{15.52}$$

In our representation

$$\gamma_5 = \begin{pmatrix} 0 & \sigma^0 \\ \sigma^0 & 0 \end{pmatrix}. \tag{15.53}$$

The usefulness of γ_5 comes from the fact that $\frac{1}{2}(I+\gamma_5)$ and $\frac{1}{2}(I-\gamma_5)$ are projectors on two-dimensional spaces. Let us define

$$\psi_R = \frac{1}{2}(I+\gamma_5)\psi, \qquad \psi_L = \frac{1}{2}(I-\gamma_5)\psi. \tag{15.54}$$

Letting each of the two projectors act on Dirac's equation (15.44) one finds that

$$-\mathrm{i}\gamma_\mu\frac{\partial}{\partial x_\mu}\psi_R + \frac{q_e}{\hbar}\gamma_\mu A^\mu\psi_R + \frac{mc}{\hbar}\psi_L = 0\,, \tag{15.55}$$

$$-\mathrm{i}\gamma_\mu\frac{\partial}{\partial x_\mu}\psi_L + \frac{q_e}{\hbar}\gamma_\mu A^\mu\psi_L + \frac{mc}{\hbar}\psi_R = 0\,. \tag{15.56}$$

In the case where $m = 0$, the Dirac equation separates in two independent equations for ψ_R and ψ_L. Because γ_5 commutes with $\sigma_{\mu\nu}$, each of these is invariant under proper orthochronous Lorentz transformations. Under space inversion ψ_R transforms into ψ'_L and ψ_L into ψ'_R. Spinors ψ_R and ψ_L are called *chiral spinors.*

---

[6] Specifically, charge conjugation transforms $u^{(1)}(\boldsymbol{p})$ into $u^{(4)}(-\boldsymbol{p})$ and $u^{(2)}(\boldsymbol{p})$ into $-u^{(3)}(-\boldsymbol{p})$. In Dirac's interpretation a positron is associated with the absence of a negative-energy electron. A positron's momentum and its spin projection are therefore opposite to those of the electronic state. Thus charge conjugation transforms an electronic state into a positron state with the same momentum and same spin projection. Exercise (15.3) shows that a Lorentz transformation changes the electron's momentum and the positron's momentum in the same way.

15.5 Electron in an Electromagnetic Field

Let us now go back to Dirac's equation for an electron in an electromagnetic field. One can write it as

$$\left\{\gamma_\mu\left[-\mathrm{i}\frac{\partial}{\partial x_\mu}+\frac{q_eA^\mu}{\hbar}\right]+\frac{mc}{\hbar}I\right\}\psi=0. \tag{15.57}$$

For the argument that follows it is useful to obtain a second-order equation for ψ. Let the following operator act on the previous equation:

$$\left\{\gamma_\nu\left[-\mathrm{i}\frac{\partial}{\partial x_\nu}+\frac{q_eA^\nu}{\hbar}\right]-\frac{mc}{\hbar}I\right\}.$$

Then one easily finds that

$$\gamma_\nu\gamma_\mu\left[\mathrm{i}\frac{\partial}{\partial x_\nu}-\frac{q_eA^\nu}{\hbar}\right]\left[\mathrm{i}\frac{\partial}{\partial x_\mu}-\frac{q_eA^\mu}{\hbar}\right]\psi-\left(\frac{mc}{\hbar}\right)^2\psi=0. \tag{15.58}$$

Let C^μ be an operator-valued four-vector that does not act in the space of γ_μ matrices. Then

$$\begin{aligned}\gamma_\nu\gamma_\mu C^\nu C^\mu &= \frac{1}{2}(\gamma_\nu\gamma_\mu C^\nu C^\mu+\gamma_\mu\gamma_\nu C^\mu C^\nu+\gamma_\nu\gamma_\mu C^\mu C^\nu-\gamma_\nu\gamma_\mu C^\mu C^\nu)\\ &= \frac{1}{2}\gamma_\nu\gamma_\mu\left[C^\nu,C^\mu\right]+g_{\mu\nu}IC^\mu C^\nu\\ &= \frac{i}{2}\sigma_{\mu\nu}\left[C^\nu,C^\mu\right]+C^\mu C_\mu I\,.\end{aligned} \tag{15.59}$$

In particular

$$\begin{aligned}&\gamma_\nu\gamma_\mu\left[\mathrm{i}\frac{\partial}{\partial x_\nu}-\frac{q_eA^\nu}{\hbar}\right]\left[\mathrm{i}\frac{\partial}{\partial x_\mu}-\frac{q_eA^\mu}{\hbar}\right]\\ &\quad=\frac{\mathrm{i}}{2}\sigma_{\mu\nu}\left(-\frac{\mathrm{i}q_e}{\hbar}\right)\left\{\left[\frac{\partial}{\partial x_\nu},A^\mu\right]+\left[A^\nu,\frac{\partial}{\partial x_\mu}\right]\right\}\\ &\qquad+\left[\mathrm{i}\frac{\partial}{\partial x_\mu}-\frac{q_eA^\mu}{\hbar}\right]\left[\mathrm{i}\frac{\partial}{\partial x^\mu}-\frac{q_eA_\mu}{\hbar}\right]\\ &\quad=\frac{q_e}{2\hbar}\sigma_{\mu\nu}F^{\mu\nu}+\left[\mathrm{i}\frac{\partial}{\partial x_\mu}-\frac{q_eA^\mu}{\hbar}\right]\left[\mathrm{i}\frac{\partial}{\partial x^\mu}-\frac{q_eA_\mu}{\hbar}\right],\end{aligned}$$

where identity matrices are understood and where

$$F^{\mu\nu}=\frac{\partial A^\mu}{\partial x_\nu}-\frac{\partial A^\nu}{\partial x_\mu} \tag{15.60}$$

is the *electromagnetic field tensor*. Equation (15.58) can therefore be written as

$$\left[\mathrm{i}\frac{\partial}{\partial x_\mu}-\frac{q_e A^\mu}{\hbar}\right]\left[\mathrm{i}\frac{\partial}{\partial x^\mu}-\frac{q_e A_\mu}{\hbar}\right]\psi+\frac{q_e}{2\hbar}\sigma_{\mu\nu}F^{\mu\nu}\psi-\left(\frac{mc}{\hbar}\right)^2\psi=0. \tag{15.61}$$

Equation (15.61) is still completely general.

In the case where the electromagnetic field does not depend on time, (15.61) has solutions for which

$$\psi(x^\mu)=\psi(\boldsymbol{r},t)=\psi(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}Et/\hbar}. \tag{15.62}$$

Function $\psi(\boldsymbol{r})$ then satisfies the equation

$$\left[\left(\frac{E-q_e\phi}{c\hbar}\right)^2-\left(\frac{mc}{\hbar}\right)^2\right]\psi(\boldsymbol{r}) \\ -\left[\mathrm{i}\boldsymbol{\nabla}+\frac{q_e\boldsymbol{A}}{\hbar}\right]\cdot\left[\mathrm{i}\boldsymbol{\nabla}+\frac{q_e\boldsymbol{A}}{\hbar}\right]\psi(\boldsymbol{r})+\frac{q_e}{2\hbar}\sigma_{\mu\nu}F^{\mu\nu}\psi(\boldsymbol{r})=0\,. \tag{15.63}$$

We now investigate two specific cases more closely: (i) time-independent magnetic field and vanishing electric field; and (ii) vanishing magnetic field and time-independent radial electric field. The first case will yield the value of the electron's gyromagnetic ratio, whereas the second one coincides with the problem of an electron in the central-field model.

i) Time-Independent Magnetic Field and Vanishing Electric Field

Here the scalar potential can be set equal to zero and the vector potential is time independent. Let $\boldsymbol{\sigma}$ denote a "vector" with components $\sigma^1=\sigma_x$, $\sigma^2=\sigma_y$ and $\sigma^3=\sigma_z$. One can write (15.63) as

$$\left[\left(\frac{E}{c\hbar}\right)^2-\left(\frac{mc}{\hbar}\right)^2\right]\psi(\boldsymbol{r})-\left[\mathrm{i}\boldsymbol{\nabla}+\frac{q_e\boldsymbol{A}}{\hbar}\right]\cdot\left[\mathrm{i}\boldsymbol{\nabla}+\frac{q_e\boldsymbol{A}}{\hbar}\right]\psi(\boldsymbol{r}) \\ +\frac{q_e}{\hbar}\boldsymbol{B}\cdot\begin{pmatrix}\boldsymbol{\sigma} & 0\\ 0 & \boldsymbol{\sigma}\end{pmatrix}\psi(\boldsymbol{r})=0\,. \tag{15.64}$$

It is convenient to represent $\psi(\boldsymbol{r})$ as

$$\psi(\boldsymbol{r})=\begin{pmatrix}\chi_1(\boldsymbol{r})\\ \chi_2(\boldsymbol{r})\end{pmatrix}, \tag{15.65}$$

where $\chi_1(\boldsymbol{r})$ and $\chi_2(\boldsymbol{r})$ are two-component columns. Let us define

$$E'=\frac{E^2-m^2c^4}{2mc^2}\,. \tag{15.66}$$

Clearly E' coincides with energy in the nonrelativistic limit. With these notations one gets for χ_1 (and also for χ_2) the following equation:

$$E'\chi_1 = \frac{1}{2m}\,[\mathrm{i}\hbar\boldsymbol{\nabla} + q_e\boldsymbol{A}]\cdot[\mathrm{i}\hbar\boldsymbol{\nabla} + q_e\boldsymbol{A}]\,\chi_1 - \frac{q_e\hbar}{2m}\boldsymbol{\sigma}\cdot\boldsymbol{B}\chi_1 . \tag{15.67}$$

This is the eigenvalue equation of the Schrödinger Hamiltonian, for a particle of charge q_e in a magnetic field. The last term entails that the electron has a magnetic moment given by $q_e\hbar\boldsymbol{\sigma}/2m$. That moment is proportional to the electron's spin $\hbar\boldsymbol{\sigma}/2$, the proportionality factor (called the *gyromagnetic ratio*) being equal to q_e/m. In classical physics the gyromagnetic ratio is equal to $q_e/2m$. This ratio also applies, in quantum mechanics, to orbital magnetic moments and angular momenta.

Thus Dirac's equation predicts that the gyromagnetic ratio associated with the electron's spin is twice larger than the ratio associated with orbital angular momentum. The relative difference between the experimental value of the electron's gyromagnetic ratio and the value predicted by the Dirac equation is of the order of one part in 10^3.[7] The correct value of the gyromagnetic ratio could not have been predicted from Schrödinger's equation or classical considerations.

ii) Vanishing Magnetic Field and Time-Independent Radial Electric Field

Here the vector potential vanishes and the scalar potential is time independent and spherically symmetric. This case applies to an electron in the central-field model.

At this stage it is convenient to go back to the original form of Dirac's equation. Letting $\boldsymbol{A} = 0$ and substituting the expression

$$\psi(\boldsymbol{r},t) = \psi(\boldsymbol{r})\mathrm{e}^{-\mathrm{i}Et/\hbar} \tag{15.68}$$

in (15.44), one gets for $\psi(\boldsymbol{r})$

$$\gamma_0\left\{-\frac{E}{\hbar c} + \frac{q_e\phi}{\hbar c}\right\}\psi(\boldsymbol{r}) - \mathrm{i}\boldsymbol{\gamma}\cdot\boldsymbol{\nabla}\psi(\boldsymbol{r}) + \frac{mc}{\hbar}\psi(\boldsymbol{r}) = 0. \tag{15.69}$$

The symbol $\boldsymbol{\gamma}$ denotes a "vector" with components $\gamma^i = -\gamma_i$ ($i = 1, 2, 3$). The last equation is equivalent to an equation of the kind

$$H\psi(\boldsymbol{r}) = E\psi(\boldsymbol{r}), \tag{15.70}$$

where

$$H = q_e\phi I - \mathrm{i}c\hbar\gamma_0\boldsymbol{\gamma}\cdot\boldsymbol{\nabla} + mc^2\gamma_0 . \tag{15.71}$$

Let us define an operator $\boldsymbol{J}$ as

---

[7] A more accurate value of the gyromagnetic ratio can be calculated by quantum field theory. The relative difference between experimental and theoretical values is lower than one part in 10^{10} [162].

$$\boldsymbol{J} = -\mathrm{i}\hbar I\, \boldsymbol{r} \times \boldsymbol{\nabla} + \frac{\hbar}{2}\begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}. \tag{15.72}$$

Clearly $\boldsymbol{J}$ represents the sum of the electron's orbital and spin angular momenta. One can check that (Exercise 15.10)

$$[H, \boldsymbol{J}] = 0, \tag{15.73}$$

i.e. the total angular momentum commutes with Dirac's Hamiltonian.

We now show that Dirac's equation makes a specific prediction for the coupling between the electron's spin and orbital magnetic moments. Once more we write[8]

$$\psi(\boldsymbol{r}) = \begin{pmatrix} \chi_1(\boldsymbol{r}) \\ \chi_2(\boldsymbol{r}) \end{pmatrix}. \tag{15.74}$$

The Hamiltonian's eigenvalue equation then becomes

$$\left(E - mc^2 - q_e\phi\right)\chi_1 + \mathrm{i}c\hbar\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\chi_2 = 0, \tag{15.75}$$

$$\left(E + mc^2 - q_e\phi\right)\chi_2 + \mathrm{i}c\hbar\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\chi_1 = 0. \tag{15.76}$$

From (15.76) one gets

$$\chi_2 = \left\{E + mc^2 - q_e\phi\right\}^{-1}(-\mathrm{i}c\hbar)\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\chi_1. \tag{15.77}$$

Substituting this expression in (15.75) one obtains an equation for χ_1 only:

$$\left(E' - q_e\phi\right)\chi_1 + (\hbar c)^2\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\left\{E' + 2mc^2 - q_e\phi\right\}^{-1}\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\chi_1 = 0. \tag{15.78}$$

Here $E' = E - mc^2$. Equation (15.78) is still exact.

Now consider a situation where relativistic effects are weak, i.e. where $E' \ll mc^2$. One can write

$$\left\{E' + 2mc^2 - q_e\phi\right\}^{-1} \approx \frac{1}{2mc^2}\left[1 - \frac{E' - q_e\phi}{2mc^2}\right]. \tag{15.79}$$

Let $\boldsymbol{a}$ and $\boldsymbol{b}$ be operators that do not act in spin space. One easily checks that

$$(\boldsymbol{\sigma}\cdot\boldsymbol{a})(\boldsymbol{\sigma}\cdot\boldsymbol{b}) = \boldsymbol{a}\cdot\boldsymbol{b} + \mathrm{i}\boldsymbol{\sigma}\cdot(\boldsymbol{a}\times\boldsymbol{b}). \tag{15.80}$$

From this identity one finds that

$$\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\left[1 - \frac{E' - q_e\phi}{2mc^2}\right]\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}$$

---

[8] The separation of Dirac's equation in two two-component equations can also be effected for an arbitrary electromagnetic field. See [29], Chap 4 and [58].

$$= \left[1 - \frac{E' - q_e\phi}{2mc^2}\right] \boldsymbol{\sigma}\cdot\boldsymbol{\nabla}\boldsymbol{\sigma}\cdot\boldsymbol{\nabla} + \frac{q_e}{2mc^2}\left[\boldsymbol{\sigma}\cdot\boldsymbol{\nabla},\phi\right]\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}$$

$$= \left[1 - \frac{E' - q_e\phi}{2mc^2}\right]\nabla^2 + \frac{q_e}{2mc^2}\frac{\mathrm{d}\phi}{\mathrm{d}r}\left[\frac{\partial}{\partial r} + \mathrm{i}\boldsymbol{\sigma}\cdot\left(\frac{\boldsymbol{r}}{r}\times\boldsymbol{\nabla}\right)\right],$$

where the fact that ϕ is spherically symmetric was used. Substituting this result in (15.78) one gets

$$(E' - q_e\phi)\chi_1 + \frac{\hbar^2}{2m}\left[\left(1 - \frac{E' - q_e\phi}{2mc^2}\right)\nabla^2 + \frac{q_e}{2mc^2}\frac{\mathrm{d}\phi}{\mathrm{d}r}\left(\frac{\partial}{\partial r} - \frac{1}{\hbar r}\boldsymbol{\sigma}\cdot\boldsymbol{L}\right)\right]\chi_1 = 0\,, \qquad (15.81)$$

where $\boldsymbol{L} = -\mathrm{i}\hbar\boldsymbol{r}\times\boldsymbol{\nabla}$.

In the term $\frac{E'-q_e\phi}{2mc^2}$ one can replace $E' - q_e\phi$ by $(-\hbar^2/2m)\nabla^2$. One obtains

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + q_e\phi\right]\chi_1 - \frac{1}{2mc^2}\left[\left(\frac{\hbar^2}{2m}\nabla^2\right)^2 + q_e\frac{\hbar^2}{2m}\frac{\mathrm{d}\phi}{\mathrm{d}r}\left(\frac{\partial}{\partial r} - \frac{1}{\hbar r}\boldsymbol{\sigma}\cdot\boldsymbol{L}\right)\right]\chi_1 = E'\chi_1\,. \qquad (15.82)$$

This is the eigenvalue equation associated with Schrödinger's equation, to which a few more terms are added. The operators

$$-\frac{1}{2mc^2}\left[\left(\frac{\hbar^2}{2m}\nabla^2\right)^2 + q_e\frac{\hbar^2}{2m}\frac{\mathrm{d}\phi}{\mathrm{d}r}\frac{\partial}{\partial r}\right] \qquad (15.83)$$

represent different relativistic effects, while

$$\frac{1}{2m^2c^2}\frac{q_e}{r}\frac{\mathrm{d}\phi}{\mathrm{d}r}\left(\frac{\hbar\boldsymbol{\sigma}}{2}\right)\cdot\boldsymbol{L} \qquad (15.84)$$

is a *spin–orbit interaction* term. Dirac's equation thus makes a specific prediction on the interaction between the electron's spin and orbital magnetic moments.

15.6 r ^{1} Potential

Insofar as relativistic corrections are weak, (15.82) describes rather well the behavior of an electron in a spherically symmetric potential. Nevertheless it is useful to go back to the exact equations (15.75) and (15.76). We will first show that for any spherically symmetric potential $\phi(r)$ these equations separate in spherical coordinates. Next we will see that for the r^{-1} potential they have exact solutions.

For a spherically symmetric potential the relativistic Hamiltonian commutes with the total angular momentum operator (15.72). Hence eigenfunctions of H can be picked as eigenfunctions of $\boldsymbol{J}\cdot\boldsymbol{J}$ and J_z. It is convenient to define a matrix operator $\boldsymbol{\tau}$ as

$$\boldsymbol{\tau} = -\mathrm{i}\hbar\boldsymbol{r}\times\boldsymbol{\nabla}+\frac{\hbar}{2}\boldsymbol{\sigma} = \boldsymbol{L}+\frac{\hbar}{2}\boldsymbol{\sigma}, \tag{15.85}$$

so that

$$\boldsymbol{J} = \begin{pmatrix} \boldsymbol{\tau} & 0 \\ 0 & \boldsymbol{\tau} \end{pmatrix}. \tag{15.86}$$

The operator $\boldsymbol{\tau}$ represents the total angular momentum of an electron in the nonrelativistic approximation.

Denote by $\Omega_{jlm}(\boldsymbol{r})$ an eigenfunction of operators $\boldsymbol{\tau}\cdot\boldsymbol{\tau}$, $\boldsymbol{L}\cdot\boldsymbol{L}$, $(\hbar/2)^2\boldsymbol{\sigma}\cdot\boldsymbol{\sigma}$ and τ_z, corresponding to eigenvalues $j(j+1)\hbar^2$, $l(l+1)\hbar^2$, $(3/4)\hbar^2$ and $m\hbar$ respectively. These functions can be built from spherical harmonics of index l and two-component spinors. They were obtained in Exercise (14.13). In the present notation

$$\Omega_{l\pm1/2,l,m} = \begin{pmatrix} \pm\sqrt{\frac{l+\frac{1}{2}\pm m}{2l+1}}Y_{l,m-1/2} \\ \sqrt{\frac{l+\frac{1}{2}\mp m}{2l+1}}Y_{l,m+1/2} \end{pmatrix}. \tag{15.87}$$

Let $\bar{l} = 2j - l$. That is, $\bar{l}$ is equal to $j + 1/2$ if l is equal to $j - 1/2$ and conversely. One shows in Exercise (15.12) that

$$(\boldsymbol{\sigma}\cdot\hat{r})\,\Omega_{jlm} = -\Omega_{j\bar{l}m}. \tag{15.88}$$

But then, for any function $f(r)$,

$$\begin{aligned}
(\boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\,f(r)\Omega_{jlm} &= (\boldsymbol{\sigma}\cdot\boldsymbol{\nabla})\,(\boldsymbol{\sigma}\cdot\hat{r})\,(\boldsymbol{\sigma}\cdot\hat{r})\,f(r)\Omega_{jlm} \\
&= \{(\boldsymbol{\sigma}\cdot\hat{r})\,(\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}) + [(\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}),(\boldsymbol{\sigma}\cdot\hat{r})]\}\,f(r)(-1)\Omega_{j\bar{l}m} \\
&= -\left\{\hat{r}\cdot\boldsymbol{\nabla}+\mathrm{i}\boldsymbol{\sigma}\cdot(\hat{r}\times\boldsymbol{\nabla})+\frac{2}{r}-2\mathrm{i}\boldsymbol{\sigma}\cdot(\hat{r}\times\boldsymbol{\nabla})\right\}f(r)\Omega_{j\bar{l}m} \\
&= -\left\{\frac{\partial}{\partial r}+\frac{2}{r}+\frac{1}{\hbar r}\boldsymbol{\sigma}\cdot\boldsymbol{L}\right\}f(r)\Omega_{j\bar{l}m} \\
&= -\left\{\frac{\partial}{\partial r}+\frac{2}{r}+\frac{1}{r}\left[\frac{\boldsymbol{\tau}\cdot\boldsymbol{\tau}}{\hbar^2}-\frac{\boldsymbol{L}\cdot\boldsymbol{L}}{\hbar^2}-\frac{\boldsymbol{\sigma}}{2}\cdot\frac{\boldsymbol{\sigma}}{2}\right]\right\}f(r)\Omega_{j\bar{l}m} \\
&= -\left\{\frac{\mathrm{d}f}{\mathrm{d}r}+\frac{1}{r}\left[1+k(j,\bar{l})\right]f\right\}\Omega_{j\bar{l}m}
\end{aligned} \tag{15.89}$$

where (15.80) was used and where

$$k(j,\bar{l}) = \begin{cases} \bar{l}+1 & \text{if } j = \bar{l}+1/2\,, \\ -\bar{l} & \text{if } j = \bar{l}-1/2\,. \end{cases} \tag{15.90}$$

Note that $k(j,\bar{l})$ is a nonvanishing integer and that $k(j,\bar{l}) = -k(j,l)$.

The result just obtained suggests the following Ansatz:

$$\chi_1 = f(r)\Omega_{jlm}, \qquad \chi_2 = \mathrm{i}g(r)\Omega_{j\bar{l}m}. \tag{15.91}$$

Substituting these relations in (15.75) and (15.76) and manipulating, one finds that

$$\left\{\frac{E + mc^2 - q_e\phi}{\hbar c}\right\} g(r) - \frac{\mathrm{d}f}{\mathrm{d}r} - \frac{1}{r}\left[1 - k(j,l)\right] f(r) = 0\,, \tag{15.92}$$

$$\left\{\frac{E - mc^2 - q_e\phi}{\hbar c}\right\} f(r) + \frac{\mathrm{d}g}{\mathrm{d}r} + \frac{1}{r}\left[1 + k(j,l)\right] g(r) = 0\,. \tag{15.93}$$

The angular and the spin dependence have completely disappeared. There remain two equations for two unknown functions $f(r)$ and $g(r)$. In general these equations must be solved numerically. For the r^{-1} potential, however, they can be solved exactly.

We let

$$\phi(r) = -\frac{Zq_e}{4\pi\varepsilon_0 r}\,. \tag{15.94}$$

Equations (15.92) and (15.93) become

$$\left\{\frac{E + mc^2 + Ze^2/r}{\hbar c}\right\} g - \frac{\mathrm{d}f}{\mathrm{d}r} - \frac{1-k}{r} f = 0\,, \tag{15.95}$$

$$\left\{\frac{E - mc^2 + Ze^2/r}{\hbar c}\right\} f + \frac{\mathrm{d}g}{\mathrm{d}r} + \frac{1+k}{r} g = 0\,, \tag{15.96}$$

where use was made of (9.4). Let us define constants

$$\beta_1 = \frac{mc^2 + E}{\hbar c}\,, \qquad \beta_2 = \frac{mc^2 - E}{\hbar c}\,, \tag{15.97}$$

$$\beta = \sqrt{\beta_1\beta_2} = \frac{1}{\hbar c}\sqrt{m^2c^4 - E^2}\,, \qquad \gamma = \frac{Ze^2}{\hbar c}\,, \tag{15.98}$$

and let us make the following change of variables:

$$\zeta = \beta r, \qquad f = \frac{1}{\zeta}\mathrm{e}^{-\zeta}F, \qquad g = \frac{1}{\zeta}\mathrm{e}^{-\zeta}G. \tag{15.99}$$

One finds that

$$\frac{\mathrm{d}F}{\mathrm{d}\zeta} - F - \frac{k}{\zeta}F - \left\{\frac{\beta_1}{\beta} + \frac{\gamma}{\zeta}\right\} G = 0\,, \tag{15.100}$$

$$\frac{\mathrm{d}G}{\mathrm{d}\zeta} - G + \frac{k}{\zeta}G - \left\{\frac{\beta_2}{\beta} - \frac{\gamma}{\zeta}\right\} F = 0\,. \tag{15.101}$$

We are interested in bound-state energies. The nonrelativistic energy is then negative, which means that β is real.

We now look for solutions of (15.100) and (15.101) in the form of power series. It is not difficult to see that for small values of ζ, exponents of leading terms of F and G must coincide. One can thus write

$$F = \zeta^s \sum_{n=0}^{\infty} a_n \zeta^n \,, \qquad a_0 \neq 0 \,; \tag{15.102}$$

$$G = \zeta^s \sum_{n=0}^{\infty} b_n \zeta^n \,, \qquad b_0 \neq 0 \,. \tag{15.103}$$

Substituting these relations in (15.100) and (15.101) and equating to zero the coefficients of ζ^{n+s}, one finds that ($n \geq 0$)

$$(n+s-k+1)a_{n+1} - a_n - \gamma b_{n+1} - \frac{\beta_1}{\beta} b_n = 0 \,, \tag{15.104}$$

$$(n+s+k+1)b_{n+1} - b_n + \gamma a_{n+1} - \frac{\beta_2}{\beta} a_n = 0 \,. \tag{15.105}$$

On the other hand, terms in ζ^{s-1} yield

$$(s-k)a_0 - \gamma b_0 = 0 \,, \tag{15.106}$$

$$(s+k)b_0 + \gamma a_0 = 0 \,. \tag{15.107}$$

From these equations one concludes there is a nontrivial solution if and only if $s = \pm(k^2 - \gamma^2)^{1/2}$. If s is negative, functions f and g have a nonintegrable singularity at $\zeta = 0$. Hence we take s positive. Functions f and g can be singular, but the singularities are integrable.[9]

Let us now turn to the behavior of F and G if ζ is large. Multiplying (15.104) by β_2, (15.105) by β and subtracting the resulting equations, one gets

$$a_{n+1} \left\{ \beta_2(n+s-k+1) - \beta\gamma \right\} = b_{n+1} \left\{ \beta(n+s+k+1) + \beta_2\gamma \right\} . \tag{15.108}$$

If n is large we thus have $\beta_2 a_{n+1} = \beta b_{n+1}$. Substituting this relation in (15.104) and (15.105) evaluated at large n, one finds that

$$a_{n+1} = \frac{2}{n} a_n \,, \qquad b_{n+1} = \frac{2}{n} b_n \,.$$

Thus functions F and G have a factor $e^{2\zeta}$ if ζ is large. Therefore f and g diverge as $\zeta \to \infty$, unless the series terminate somewhere.

---

[9] The integral $\int d\zeta \, \zeta^2 f^2$ is finite around the origin if the function $\zeta^2 f^2 \sim \zeta^{2s}$ is less singular than ζ^{-1}. If s is negative this implies that $k^2 < \gamma^2 + 1/4$ or, since $|k| \geq 1$, that $\gamma^2 > 3/4$. The latter inequality is never satisfied for values of Z found in nature. If s is positive and $|k| = 1$, one can eliminate the singularity in f and g by taking into account the finite extension of the nucleus.

Suppose that the F series stops at $n = n'$, so that $a_{n'+1} = 0$. Then $b_{n'+1} = 0$ also, and (15.104) and (15.105) both yield

$$\beta a_{n'} = -\beta_1 b_{n'} . \tag{15.109}$$

Substituting this relation in (15.108) evaluated for $n + 1 = n'$, one gets

$$2\beta(n'+s) = \gamma(\beta_1 - \beta_2) = \frac{2E\gamma}{\hbar c} . \tag{15.110}$$

From this one concludes that $E > 0$. Take the square of (15.110) and substitute in it the value of β given in (15.98). Isolating E one finds that[10]

$$E = mc^2 \left\{ 1 + \frac{\gamma^2}{(n'+s)^2} \right\}^{-1/2} = mc^2 \left\{ 1 + \frac{\gamma^2}{\left(n' + \sqrt{k^2 - \gamma^2}\right)^2} \right\}^{-1/2}$$
$$= mc^2 \left\{ 1 + \frac{\gamma^2}{\left(n' + \sqrt{(j+1/2)^2 - \gamma^2}\right)^2} \right\}^{-1/2} . \tag{15.111}$$

The last equality comes form the fact that $|k| = j + 1/2$.

The index n' is a nonnegative integer. If $n' > 0$ the constant k can be either positive or negative. If $n' = 0$, on the other hand, (15.106) and (15.109) are compatible only if k is positive. This entails that $j = l + 1/2$.

Constant γ is equal to $Z\alpha$, where $\alpha = e^2/\hbar c \approx 1/137$ is the *fine-structure constant.* For hydrogen and light hydrogen-like ions, one can expand E in a series of powers of γ. Letting $n = n' + |k|$ one finds that

$$E = mc^2 \left\{ 1 - \frac{\gamma^2}{2n^2} - \frac{\gamma^4}{2n^3} \left(\frac{1}{|k|} - \frac{3}{4n} \right) + O(\gamma^6) \right\} . \tag{15.112}$$

The first term in the right-hand side of (15.112) represents the relativistic energy of the electron at rest. The second term coincides with the value calculated in Sect. 9.2 with the nonrelativistic Hamiltonian. The third term represents the so-called *fine structure* of levels: for a given value of n, levels corresponding to different values of j have different energies.

Figure 15.1 is similar to Fig. 9.1 on p. 188, except that the fine structure of the first three energy levels of hydrogen is now illustrated. Note that the fine structure is not shown to scale. The separation between levels with the same n is about 10^4 times smaller than the separation between levels with different n. Values of l shown in Fig. 15.1 correspond to those of the χ_1 component of the wave function which, for solutions with positive total energy, is much more important than χ_2.

---

[10] The explicit representation of functions f and g, properly normalized, can be found in [139], Sect. 36.

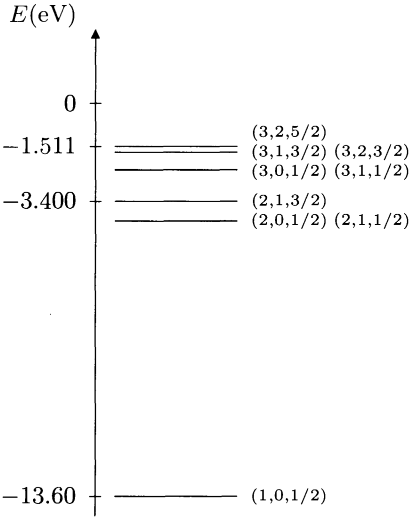

Fig. 15.1. The first three energy levels of the hydrogen atom and the corresponding quantum numbers (n, l, j) of associated states

Equation (15.112) and Fig. 15.1 show that even if the n degeneracy is partly removed, there remains a degeneracy in j. In the real world the latter is removed by quantum-field-theory effects that give rise to the *Lamb shift* of spectral lines [162]. Except for such small effects the hydrogen atom spectrum predicted by Dirac's equation is in excellent agreement with experiment.

Exercises

15.1. Obtain the commutation relations of matrices $\Sigma_{\kappa\lambda}$ from the expression for their element $(\Sigma_{\kappa\lambda})^{\mu}{}_{\nu}$.

15.2. Let $\{a^j,\ j = 1,\ 2,\ 3\}$ be three real constants and let $\{\Sigma_{0j}\}$ be the generators of special Lorentz transformations. Show that

$$\exp\left(-\mathrm{i}\sum_{j=1}^{3} a^j \Sigma_{0j}\right) = \begin{pmatrix} \gamma & -\gamma[v^k]^T \\ -\gamma[v^k] & I + [v^k v^l]\left\{\frac{\gamma-1}{\boldsymbol{v}^2}\right\} \end{pmatrix},$$

where square brackets denote a submatrix and where dimensionless constants γ and v^j are given by

$$\gamma v^j = \frac{\sinh|\boldsymbol{a}|}{|\boldsymbol{a}|}a^j, \qquad \gamma = \left(1 - \boldsymbol{v}^2\right)^{-1/2}.$$

This transformation changes a null momentum into a momentum $\boldsymbol{p} = -m\gamma c\boldsymbol{v}$.

15.3. Let $\{a^j\}$, $\{v^j\}$ and γ be defined as in Exercise (15.2) and let $\{\sigma_{0j}\}$ be generators associated with the spinor representation.

a) Show that

$$\exp\left(-\frac{\mathrm{i}}{2}\sum_{j=1}^{3} a^j\sigma_{0j}\right)$$

$$= \cosh\frac{|\boldsymbol{a}|}{2}\begin{pmatrix} 1 & -\left\{\frac{\gamma-1}{\gamma+1}\right\}^{1/2}\frac{\boldsymbol{v}\cdot\boldsymbol{\sigma}}{|\boldsymbol{v}|} \\ -\left\{\frac{\gamma-1}{\gamma+1}\right\}^{1/2}\frac{\boldsymbol{v}\cdot\boldsymbol{\sigma}}{|\boldsymbol{v}|} & 1 \end{pmatrix} \equiv \cosh\frac{|\boldsymbol{a}|}{2}\bar{M}\,.$$

b) Let $\boldsymbol{p} = -m\gamma c\boldsymbol{v}$. Check that $u^{(i)}(\boldsymbol{p}) = \bar{M}u^{(i)}(0)$ if $i = 1$ or 2 and that $u^{(i)}(-\boldsymbol{p}) = \bar{M}u^{(i)}(0)$ if $i = 3$ or 4. Interpret this result.

15.4. Use results of Exercises (15.2) and (15.3) to check the lemma on p. 343, in the case where the $a^{\kappa\lambda}$ correspond to a special Lorentz transformation.

15.5. Show that Dirac's equation is invariant under space inversion.

15.6. Show that Dirac's equation is invariant under time reversal.

15.7. Show that Dirac's equation is invariant under charge conjugation.

15.8. Show that Dirac's equation is invariant under a gauge transformation.

15.9. Let $(\boldsymbol{E}, \boldsymbol{B})$ be a time-dependent electromagnetic field and suppose that a gauge where $\phi = 0$ is used. Let

$$\psi(\boldsymbol{r}, t) = \exp\left\{-\frac{\mathrm{i}}{\hbar}mc^2t\right\}\begin{pmatrix}\chi_1(\boldsymbol{r}, t)\\ \chi_2(\boldsymbol{r}, t)\end{pmatrix}.$$

Show that in the nonrelativistic limit $\chi_1(\boldsymbol{r}, t)$ satisfies the equation

$$\mathrm{i}\hbar\frac{\partial\chi_1}{\partial t} = \frac{1}{2m}\left[\mathrm{i}\hbar\boldsymbol{\nabla} + q_e\boldsymbol{A}\right]\cdot\left[\mathrm{i}\hbar\boldsymbol{\nabla} + q_e\boldsymbol{A}\right]\chi_1 - \frac{q_e\hbar}{2m}\boldsymbol{\sigma}\cdot\boldsymbol{B}\chi_1.$$

15.10. Check (15.73).

15.11. Check (15.88) in the case where $j = 3/2$, $l = 1$ and $m = 1/2$.

15.12. Let $\boldsymbol{\tau}$ be the operator in (15.85) and let $\Omega_{jlm}(\boldsymbol{r})$ be the function defined in (15.87).

a) Show that $[\boldsymbol{\tau}, \boldsymbol{\sigma}\cdot\hat{r}] = 0$. From this derive that

$$(\sigma\cdot\hat{r})\,\Omega_{jlm} = a\Omega_{jlm} + b\Omega_{j\bar{l}m}\,, \tag{15.113}$$

where $\bar{l} = 2j - l$ and where a and b do not depend on m.

b) Show that Ω_{jlm} and $(\boldsymbol{\sigma}\cdot\hat{r})\Omega_{jlm}$ are two eigenvectors of the space-inversion operator with opposite eigenvalues. Deduce that coefficient a vanishes.

c) Evaluate both sides of (15.113) if $\hat{r} = \hat{z}$ and $m = 1/2$, and check that $b = -1$.

15.13. Write $\chi_1(\boldsymbol{r}) = f(r)\Omega_{jlm}$ in (15.82) and derive an equation for $f(r)$.

16 The Path Integral

The importance of the evolution operator in quantum mechanics has been emphasized several times. We now obtain an explicit formula for its matrix elements in terms of a path integral. Next we evaluate that integral in the semiclassical case, that is, when the action associated with the classical trajectory is much larger than Planck's constant. This will lead to an approximation for the eigenvalues and eigenfunctions of the Hamiltonian of a particle in one dimension, valid for high energy levels. The relation between quantum and classical properties will also be investigated more closely.

16.1 Propagator and Path Integral

Consider an isolated quantum system whose Hamiltonian may depend on time. The Schrödinger equation

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = H(t)|\psi(t)\rangle ,$$

can formally be integrated as

$$|\psi(t)\rangle = U(t,t_0)|\psi(t_0)\rangle ,$$

where $U(t,t_0)$ is the evolution operator.

More specifically, let the quantum system be a spinless particle with no internal degrees of freedom and constrained to one spatial dimension. The state space has a basis made of kets $|x\rangle$. Matrix elements of the evolution operator in that basis, that is,

$$U(x,t;x_0,t_0) = \langle x|U(t,t_0)|x_0\rangle , \tag{16.1}$$

make up the *propagator*. They yield, through (5.54), the time evolution of the wave function.

Let δt be an infinitesimal time internal and let $t = t_0 + \delta t$. The evolution operator is then essentially given by

$$U(t_0+\delta t,t_0) = I - \frac{\mathrm{i}}{\hbar}\delta t H(t_0) . \tag{16.2}$$

L. Marchildon, *Quantum Mechanics*

Thus for the propagator

$$U(x,t_0+\delta t;x_0,t_0)=\delta(x-x_0)-\frac{\mathrm{i}\delta t}{\hbar}\langle x|H(t_0)|x_0\rangle\,. \tag{16.3}$$

These results are valid in the general case where H depends on time.

Knowledge of the propagator $U(x,t;x_0,t_0)$ is sufficient to determine the wave function at any time t if it is known at a given time t_0. We will write the propagator in terms of a path integral, a mathematical object that we now define.

Let (t_0,t) be a finite interval. Let us subdivide it in $n+1$ subintervals of equal length $\varepsilon=(t-t_0)/(n+1)$ and denote by t_l ($l=1,2,\dots,n$) the points that separate these subintervals. One can write

$$\begin{aligned}U(x,t;x_0,t_0)=\int \mathrm{d}x_1\dots\mathrm{d}x_n\,&U(x,t;x_n,t_n)\\ &\times U(x_n,t_n;x_{n-1},t_{n-1})\dots U(x_1,t_1;x_0,t_0)\,.\end{aligned} \tag{16.4}$$

Consider the factor $U(x_l,t_l;x_{l-1},t_{l-1})$. For n large enough, hence for small ε,

$$U(x_l,t_l;x_{l-1},t_{l-1})=\delta(x_l-x_{l-1})-\frac{\mathrm{i}\varepsilon}{\hbar}\langle x_l|H(t_{l-1})|x_{l-1}\rangle\,. \tag{16.5}$$

Suppose that

$$H(t)=\frac{1}{2m}P^2+V(X,t)\,. \tag{16.6}$$

One then finds

$$\begin{aligned}\langle x_l|\frac{1}{2m}P^2|x_{l-1}\rangle&=\int_{-\infty}^{\infty}\mathrm{d}p\,\frac{1}{2m}\langle x_l|p\rangle\langle p|P^2|x_{l-1}\rangle\\&=\int_{-\infty}^{\infty}\mathrm{d}p\,\frac{p^2}{2m}\langle x_l|p\rangle\langle p|x_{l-1}\rangle\\&=\int_{-\infty}^{\infty}\mathrm{d}p\,\frac{1}{2\pi\hbar}\frac{p^2}{2m}\exp\left\{\frac{\mathrm{i}p}{\hbar}(x_l-x_{l-1})\right\}\end{aligned}$$

and

$$\langle x_l|V(X,t_{l-1})|x_{l-1}\rangle=V\left(\frac{x_l+x_{l-1}}{2},t_{l-1}\right)\delta(x_l-x_{l-1})\,.$$

Thus we get

$$\begin{aligned}&U(x_l,t_l;x_{l-1},t_{l-1})\\&\quad=\delta(x_l-x_{l-1})\left\{1-\frac{\mathrm{i}\varepsilon}{\hbar}V\left(\frac{x_l+x_{l-1}}{2},t_{l-1}\right)\right\}\\&\qquad-\frac{\mathrm{i}\varepsilon}{\hbar}\int_{-\infty}^{\infty}\mathrm{d}p\,\frac{1}{2\pi\hbar}\frac{p^2}{2m}\exp\left\{\frac{\mathrm{i}p}{\hbar}(x_l-x_{l-1})\right\}\end{aligned}$$

$$
\begin{aligned}
&= \int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi\hbar} \exp\left\{ \frac{\mathrm{i}p}{\hbar}(x_l - x_{l-1}) \right\} \\
&\quad \times \left[1 - \frac{\mathrm{i}\varepsilon}{\hbar} V\left(\frac{x_l + x_{l-1}}{2}, t_{l-1} \right) - \frac{\mathrm{i}\varepsilon}{\hbar} \frac{p^2}{2m} \right] \\
&= \int_{-\infty}^{\infty} \frac{\mathrm{d}p}{2\pi\hbar} \exp\left\{ \frac{\mathrm{i}\varepsilon}{\hbar} \left[p\left(\frac{x_l - x_{l-1}}{\varepsilon} \right) - H\left(p, \frac{x_l + x_{l-1}}{2}, t_{l-1} \right) \right] \right\},
\end{aligned}
\tag{16.7}
$$

where $H\left(p, \frac{1}{2}(x_l + x_{l-1}), t_{l-1}\right)$ is the *classical Hamiltonian* and where terms of order ε^2 have been neglected.

Let us now substitute this result in the expression (16.4) for the propagator. Writing x_{n+1} for x and t_{n+1} for t one gets

$$
\begin{aligned}
U(x, t; x_0, t_0) &= \int \left[\prod_{k=1}^{n} \mathrm{d}x_k \right] \prod_{l=1}^{n+1} \int_{-\infty}^{\infty} \frac{\mathrm{d}p_l}{2\pi\hbar} \exp\left\{ \frac{\mathrm{i}\varepsilon}{\hbar} \left[p_l \left(\frac{x_l - x_{l-1}}{\varepsilon} \right) \right.\right. \\
&\qquad \left.\left. - H\left(p_l, \frac{x_l + x_{l-1}}{2}, t_{l-1} \right) \right] \right\} \\
&= \int \left[\prod_{k=1}^{n} \mathrm{d}x_k \right] \left[\prod_{m=1}^{n+1} \frac{\mathrm{d}p_m}{2\pi\hbar} \right] \exp\left\{ \mathrm{i} \sum_{l=1}^{n+1} \frac{\varepsilon}{\hbar} \left[p_l \left(\frac{x_l - x_{l-1}}{\varepsilon} \right) \right.\right. \\
&\qquad \left.\left. - H\left(p_l, \frac{x_l + x_{l-1}}{2}, t_{l-1} \right) \right] \right\}.
\end{aligned}
$$

We are interested in the limit of that expression as $n \to \infty$ and $\varepsilon \to 0$. The argument of the exponential then becomes

$$
\begin{aligned}
&\lim_{n\to\infty, \varepsilon\to 0} \mathrm{i} \sum_{l=1}^{n+1} \frac{\varepsilon}{\hbar} \left[p_l \left(\frac{x_l - x_{l-1}}{\varepsilon} \right) - H\left(p_l, \frac{x_l + x_{l-1}}{2}, t_{l-1} \right) \right] \\
&= \frac{\mathrm{i}}{\hbar} \int_{t_0}^{t} \mathrm{d}\tau \left[p \frac{\mathrm{d}x}{\mathrm{d}\tau} - H(p, x, \tau) \right] \\
&= \frac{\mathrm{i}}{\hbar} S\left\{ p(\tau), x(\tau) \right\},
\end{aligned}
\tag{16.8}
$$

where $S\{p(\tau), x(\tau)\}$ is the *action functional* evaluated on the trajectory $\{p(\tau), x(\tau)\}$ in *phase space* ([93], Sect. 8.5; [136], Sect. 43). Except for the fact that $x(t_0) = x_0$ and $x(t) = x$, this trajectory is completely arbitrary. It does not in general coincide with the classical trajectory.

Introducing notations

$$
\mathcal{D}x = \lim_{n\to\infty} \prod_{k=1}^{n} \mathrm{d}x_k, \qquad \frac{\mathcal{D}p}{2\pi\hbar} = \lim_{n\to\infty} \prod_{m=1}^{n+1} \frac{\mathrm{d}p_m}{2\pi\hbar},
\tag{16.9}
$$

one can write

$$U(x,t;x_0,t_0)=\int\frac{\bar{\mathcal{D}}x\mathcal{D}p}{2\pi\hbar}\exp\left\{\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t}\mathrm{d}\tau\left[p\frac{\mathrm{d}x}{\mathrm{d}\tau}-H(p,x,\tau)\right]\right\}. \tag{16.10}$$

This is the *path integral.* It is defined by the limiting process we have explained.[1]

The propagator $U(x,t;x_0,t_0)$ is thus obtained as an integral on all phase-space trajectories going from x_0 to x. There are no restrictions on position or momentum. The function integrated is the exponential of $\mathrm{i}/\hbar$ times the action associated with each trajectory.

The expression just obtained is valid in a more general context than the one in which it was derived, i.e. a Hamiltonian given by (16.6). In the present case, however, the expression for the propagator considerably simplifies if the *dp* integrals are carried out. Going back to (16.7) one sees that

$$\begin{aligned}
&U(x_l,t_l;x_{l-1},t_{l-1})\\
&=\int_{-\infty}^{\infty}\frac{\mathrm{d}p}{2\pi\hbar}\exp\left\{\frac{\mathrm{i}\varepsilon}{\hbar}\left[p\left(\frac{x_l-x_{l-1}}{\varepsilon}\right)-\frac{p^2}{2m}-V\left(\frac{x_l+x_{l-1}}{2},t_{l-1}\right)\right]\right\}\\
&=\int_{-\infty}^{\infty}\frac{\mathrm{d}p}{2\pi\hbar}\exp\left\{-\frac{\mathrm{i}\varepsilon}{2m\hbar}\left[p-\frac{m(x_l-x_{l-1})}{\varepsilon}\right]^2\right.\\
&\qquad\left.+\frac{\mathrm{i}m}{2\varepsilon\hbar}(x_l-x_{l-1})^2-\frac{\mathrm{i}\varepsilon}{\hbar}V\left(\frac{x_l+x_{l-1}}{2},t_{l-1}\right)\right\}\\
&=\exp\left\{\frac{\mathrm{i}m}{2\varepsilon\hbar}(x_l-x_{l-1})^2-\frac{\mathrm{i}\varepsilon}{\hbar}V\left(\frac{x_l+x_{l-1}}{2},t_{l-1}\right)\right\}\\
&\quad\times\int_{-\infty}^{\infty}\frac{\mathrm{d}q}{2\pi\hbar}\exp\left\{-\frac{\mathrm{i}\varepsilon}{2m\hbar}q^2\right\}.
\end{aligned}\tag{16.11}$$

There remains a Gaussian integral with purely imaginary exponential argument. In Sect. 16.2 we will show that

$$\int_{-\infty}^{\infty}\frac{\mathrm{d}q}{2\pi\hbar}\exp\left\{-\frac{\mathrm{i}\varepsilon}{2m\hbar}q^2\right\}=\left\{\frac{m}{2\pi\mathrm{i}\varepsilon\hbar}\right\}^{1/2}. \tag{16.12}$$

Substituting (16.12) and (16.11) in (16.4) one finds that

$$\begin{aligned}
U(x,t;x_0,t_0)&=\int\left[\prod_{k=1}^{n}\mathrm{d}x_k\right]\left(\frac{m}{2\pi\mathrm{i}\varepsilon\hbar}\right)^{\frac{n+1}{2}}\\
&\times\exp\left\{\frac{\mathrm{i}}{\hbar}\sum_{l=1}^{n+1}\varepsilon\left[\frac{m}{2}\left(\frac{x_l-x_{l-1}}{\varepsilon}\right)^2-V\left(\frac{x_l+x_{l-1}}{2},t_{l-1}\right)\right]\right\}.
\end{aligned}\tag{16.13}$$

---

[1] The path-integral formulation of quantum mechanics is developed in [84]. See also [1], Sects. 11 and 12, [83] and, for a more formal exposition, [195]. The numerical computation of path integrals is developed in [60], [89] and [150].

Taking the limit of that expression as $n \to \infty$ and $\varepsilon \to 0$, one gets

$$\begin{aligned} U(x,t;x_0,t_0) &= \int \mathcal{D}x \exp\left\{\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} \mathrm{d}\tau \left[\frac{m}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}\tau}\right)^2 - V(x,\tau)\right]\right\} \\ &= \int \mathcal{D}x \exp\left\{\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} \mathrm{d}\tau\, \mathcal{L}\left(x, \frac{\mathrm{d}x}{\mathrm{d}\tau}, \tau\right)\right\} \\ &= \int \mathcal{D}x \exp\left\{\frac{\mathrm{i}}{\hbar} S\left[x(\tau)\right]\right\}. \end{aligned} \tag{16.14}$$

The symbol $\mathcal{L}$ stands for the particle's *Lagrangian* and the functional $S(\tau)$ is evaluated on trajectory $x(\tau)$ in *configuration space.* The element $\mathcal{D}x$ is defined as

$$\mathcal{D}x = \lim_{n\to\infty} \left(\frac{m}{2\pi \mathrm{i}\varepsilon\hbar}\right)^{\frac{n+1}{2}} \prod_{k=1}^{n} \mathrm{d}x_k . \tag{16.15}$$

Equation (16.14) gives the propagator as a path integral on all trajectories in configuration space that start at x_0 and end at x. Some of these trajectories are shown in Fig. 16.1.

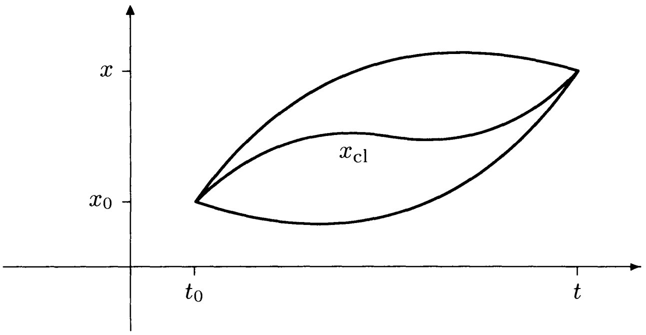

Fig. 16.1. The classical trajectory and two other trajectories

We close this section with a few remarks.

i) If $S \gg \hbar$ the path integral is largely dominated by the *classical trajectory* and the ones in its immediate neighborhood. Indeed the action is stationary around the classical trajectory. Everywhere else the integral essentially vanishes due to rapid variations in the exponential's phase. This point will be developed further in later sections.
ii) The path integral expression of the propagator directly reveals the effect on the propagator of invariance properties of the action.
iii) The above results immediately generalize to a quantum system of several spinless particles with no internal degrees of freedom, in three-dimensional space. Thus

$$U(\boldsymbol{r}_1,\dots,\boldsymbol{r}_N,t;\boldsymbol{r}_{10},\dots,\boldsymbol{r}_{N0},t_0) = \int \prod_{r=1}^{N} \left[\frac{\bar{\mathcal{D}}\boldsymbol{r}_r \mathcal{D}\boldsymbol{p}_r}{(2\pi\hbar)^3}\right]$$
$$\times \exp\left\{\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} \mathrm{d}\tau \left[\sum_{s=1}^{N} \boldsymbol{p}_s \cdot \frac{\mathrm{d}\boldsymbol{r}_s}{\mathrm{d}\tau} - H(\boldsymbol{p}_1,\dots,\boldsymbol{p}_N;\boldsymbol{r}_1,\dots,\boldsymbol{r}_N;\tau)\right]\right\}, \tag{16.16}$$

where all phase space trajectories going from a given initial point to a given final point are included. In the case where

$$H = \sum_{s=1}^{N} \frac{1}{2m_s} \boldsymbol{P}_s \cdot \boldsymbol{P}_s + V(\boldsymbol{R}_1,\dots,\boldsymbol{R}_N;t), \tag{16.17}$$

one also obtains

$$U(\boldsymbol{r}_1,\dots,\boldsymbol{r}_N,t;\boldsymbol{r}_{10},\dots,\boldsymbol{r}_{1N},t_0)$$
$$= \int \prod_{s=1}^{N} [\mathcal{D}\boldsymbol{r}_s] \exp\left\{\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} \mathrm{d}\tau\, \mathcal{L}\left[\boldsymbol{r}_1,\dots,\boldsymbol{r}_N;\frac{\mathrm{d}\boldsymbol{r}_1}{\mathrm{d}\tau},\dots,\frac{\mathrm{d}\boldsymbol{r}_N}{\mathrm{d}\tau};\tau\right]\right\}, \tag{16.18}$$

where $\mathcal{L}$ stands for the Lagrangian of the N-particle system.

16.2 Convergence of Path Integrals

Let us first compute the integral left out in Sect. 16.1, i.e. for real and positive α:

$$I(\alpha) = \int_{-\infty}^{\infty} \mathrm{d}q \exp\left\{-\mathrm{i}\alpha q^2\right\}. \tag{16.19}$$

Let $q = \mathrm{e}^{-\mathrm{i}\pi/4} z$. Since the function to be integrated is everywhere analytic one can use Cauchy's theorem and write

$$I(\alpha) = \lim_{R\to\infty} \mathrm{e}^{-\mathrm{i}\pi/4} \int_{C(R)} \mathrm{d}z \exp\left\{-\alpha z^2\right\}, \tag{16.20}$$

where $C(R)$ is the open contour shown in Fig. 16.2 (solid lines).

Consider the integral on the circular arc in the first quadrant. One has

$$\left|\int_{\mathrm{arc}} \mathrm{d}z \exp\left\{-\alpha z^2\right\}\right| = \left|\int_0^{\pi/4} R\mathrm{e}^{\mathrm{i}\theta}\mathrm{i}\mathrm{d}\theta \exp\left\{-\alpha R^2 \mathrm{e}^{2\mathrm{i}\theta}\right\}\right|$$
$$\leq R\int_0^{\pi/4} \mathrm{d}\theta \exp\left\{-\alpha R^2 \cos 2\theta\right\}$$

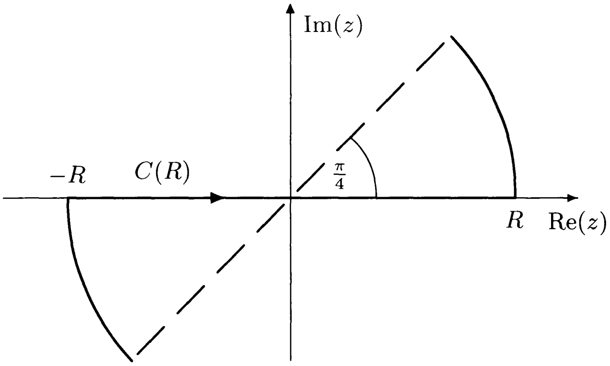

Fig. 16.2. Contour used for the evaluation of $I(\alpha)$

$$
\begin{aligned}
&\leq R \int_0^{\pi/4-(\alpha R^2)^{-3/4}} \mathrm{d}\theta \exp\left\{-\alpha R^2 \cos 2\theta\right\} \\
&\quad + R \int_{\pi/4-(\alpha R^2)^{-3/4}}^{\pi/4} \mathrm{d}\theta \, \exp\left\{-\alpha R^2 \cos 2\theta\right\} .
\end{aligned}
$$

It is not difficult to show that both terms on the right-hand side vanish as R goes to infinity. Hence the integral on the circular arc vanishes. The same holds for the integral on the circular arc in the third quadrant. Only the integral on the Re(z) axis remains, and therefore

$$
I(\alpha) = \mathrm{e}^{-\mathrm{i}\pi/4} \int_{-\infty}^{\infty} \mathrm{d}z \, \exp\left\{-\alpha z^2\right\} = \mathrm{e}^{-\mathrm{i}\pi/4} \sqrt{\frac{\pi}{\alpha}} = \sqrt{\frac{\pi}{\mathrm{i}\alpha}} . \tag{16.21}
$$

♣

The formula just obtained can be interpreted as the analytic continuation to values of β such that $0 \leq \arg(\beta) \leq \pi/2$ of the following identity:

$$
\int_{-\infty}^{\infty} \mathrm{d}q \, \exp\left\{-\beta q^2\right\} = \sqrt{\frac{\pi}{\beta}} . \tag{16.22}
$$

In particular

$$
\begin{aligned}
I(\alpha) &= \lim_{\eta \to 0^+} \int_{-\infty}^{\infty} \mathrm{d}q \, \exp\left\{-\mathrm{i}(\alpha - \mathrm{i}\eta) q^2\right\} \\
&= \lim_{\eta \to 0^+} \int_{-\infty}^{\infty} \mathrm{d}q \, \exp\left\{-\mathrm{i}\alpha q^2 - \eta q^2\right\} .
\end{aligned} \tag{16.23}
$$

The last equality is useful in that the integral is often easier to evaluate for $\eta \neq 0$ than for $\eta = 0$. The analytic continuation can also be carried to values of β such that $-\pi/2 \leq \arg(\beta) \leq 0$.

The convergence of path integrals is a delicate question. In the expression (16.14) of the propagator,

$$U(x,t;x_0,t_0)=\int \mathcal{D}x \exp\left\{\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t}\mathrm{d}\tau\left[\frac{m}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}\tau}\right)^2-V(x,\tau)\right]\right\}, \tag{16.24}$$

the integral on Dx is interpreted as an integral on all trajectories going from x_0 to x. The exponential's argument is purely imaginary. Thus all trajectories, even those very far from the classical one, give the integral a complex contribution of unit norm. But the convergence of path integrals can be accelerated by the introduction of a *convergence factor* similar to the one used in (16.23). We can write

$$U(x,t;x_0,t_0)$$
$$=\lim_{\eta\to 0^+}\int \mathcal{D}x \exp\left\{\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t}\mathrm{d}\tau\left[\frac{m}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}\tau}\right)^2-V(x,\tau)+\mathrm{i}\eta x^2\right]\right\}. \tag{16.25}$$

The additional term only influences trajectories where x is very large, that is, the ones very far from the classical trajectory.

As in the evaluation of $I(\alpha)$, the introduction of the convergence factor is related to the analytic continuation of the path integral to imaginary values of time. Let $t=-\mathrm{i}s$ and $\mathrm{d}\tau=-\mathrm{i}\mathrm{d}\sigma$ in (16.24), where s and $\mathrm{d}\sigma$ are real variables. One then has

$$U(x,-\mathrm{i}s;x_0,-\mathrm{i}s_0)=\int \mathcal{D}x \exp\left\{-\frac{1}{\hbar}\int_{s_0}^{s}\mathrm{d}\sigma\left[\frac{m}{2}\left(\frac{\mathrm{d}x}{\mathrm{d}\sigma}\right)^2+V(x,-\mathrm{i}\sigma)\right]\right\}. \tag{16.26}$$

For a sufficiently well-behaved potential (for instance a time-independent one), the exponential's argument goes to $-\infty$ as x gets very large. Hence the right-hand side converges rapidly and is a well-defined function of s_0 and s, say $F(x,s;x_0,s_0)$. For t and t_0 real the propagator is then given by

$$U(x,t;x_0,t_0)=F(x,\mathrm{i}t;x_0,\mathrm{i}t_0). \tag{16.27}$$

The convergence of functional integrals can also be enforced by carrying out explicitly the limiting process indicated in Sect. 16.1. We illustrate this process with the case of a free particle, where $V(x,t)=0$. The propagator is then given by

$$U_{\text{free}}(x,t;x_0,t_0)=\lim_{\eta\to\infty,\varepsilon\to 0}\int\left[\prod_{k=1}^{n}\mathrm{d}x_k\right]\left(\frac{m}{2\pi\mathrm{i}\hbar\varepsilon}\right)^{\frac{n+1}{2}}$$
$$\times\exp\left\{\frac{\mathrm{i}}{\hbar}\sum_{l=1}^{n+1}\varepsilon\frac{m}{2}\left(\frac{x_l-x_{l-1}}{\varepsilon}\right)^2\right\}. \tag{16.28}$$

Let us compute the integral on $\mathrm{d}x_1$. Using one factor $(m/2\pi i\hbar\varepsilon)^{1/2}$ one gets

$$\int_{-\infty}^{\infty} \mathrm{d}x_1 \left(\frac{m}{2\pi \mathrm{i}\hbar\varepsilon}\right)^{1/2} \exp\left\{\frac{\mathrm{i}m}{2\varepsilon\hbar}\left[(x_2-x_1)^2+(x_1-x_0)^2\right]\right\}$$

$$= \int_{-\infty}^{\infty} \mathrm{d}x_1 \left(\frac{m}{2\pi \mathrm{i}\hbar\varepsilon}\right)^{1/2} \exp\left\{\frac{\mathrm{i}m}{\varepsilon\hbar}\left[\left(x_1-\frac{x_2+x_0}{2}\right)^2+\frac{1}{4}(x_2-x_0)^2\right]\right\}$$

$$= \frac{1}{\sqrt{2}}\exp\left\{\frac{\mathrm{i}m}{2\hbar(2\varepsilon)}(x_2-x_0)^2\right\}.$$

By induction one can show that (Exercise 16.3)

$$\int \left[\prod_{k=1}^{p}\mathrm{d}x_k\right]\left(\frac{m}{2\pi\mathrm{i}\hbar\varepsilon}\right)^{p/2}\exp\left\{\frac{\mathrm{i}}{\hbar}\sum_{l=1}^{p+1}\varepsilon\frac{m}{2}\left(\frac{x_l-x_{l-1}}{\varepsilon}\right)^2\right\}$$

$$= \frac{1}{\sqrt{p+1}}\exp\left\{\frac{\mathrm{i}m}{2\hbar(p+1)\varepsilon}(x_{p+1}-x_0)^2\right\}. \tag{16.29}$$

Hence one gets for the propagator

$$U_{\text{free}}(x,t;x_0,t_0)$$

$$= \lim_{\eta\to\infty,\varepsilon\to 0}\left(\frac{m}{2\pi\mathrm{i}\hbar\varepsilon}\right)^{1/2}\frac{1}{\sqrt{n+1}}\exp\left\{\frac{\mathrm{i}m}{2\hbar(n+1)\varepsilon}(x_{n+1}-x_0)^2\right\}$$

$$= \left\{\frac{m}{2\pi\mathrm{i}\hbar(t-t_0)}\right\}^{1/2}\exp\left\{\frac{\mathrm{i}m(x-x_0)^2}{2\hbar(t-t_0)}\right\}. \tag{16.30}$$

This result coincides with (5.99).

16.3 The Semiclassical Approximation

In the remainder of this chapter, we will consider potentials (hence Hamiltonians) that do not explicitly depend on time. The action is thus given by

$$S\left[x(\tau)\right] = \int_{t_0}^{t}\mathrm{d}\tau\,\mathcal{L}(x,\dot{x}) = \int_{t_0}^{t}\mathrm{d}\tau\,\left[\frac{m}{2}\dot{x}^2 - V(x)\right], \tag{16.31}$$

where $\dot{x} = \mathrm{d}x/\mathrm{d}t$. The action is stationary around the classical trajectory. The path integral is carried out on all trajectories going from x_0 to x.

Let us write $x(\tau) = x_{\text{cl}}(\tau) + \delta x(\tau)$ and expand the action functional up to terms quadratic in δx. One has

$$S\left[x_{\text{cl}}(\tau)+\delta x(\tau)\right] = \int_{t_0}^{t}\mathrm{d}\tau\left[\frac{m}{2}\left(\dot{x}_{\text{cl}}+\delta\dot{x}\right)^2 - V\left(x_{\text{cl}}+\delta x\right)\right]$$

$$= \int_{t_0}^{t}\mathrm{d}\tau\left[\frac{m}{2}\dot{x}_{\text{cl}}^2 + m\dot{x}_{\text{cl}}\delta\dot{x} + \frac{m}{2}(\delta\dot{x})^2 - V(x_{\text{cl}}) - \left(\frac{\partial V}{\partial x}\right)_{x_{\text{cl}}}\delta x \right.$$

$$\left. - \frac{1}{2}\left(\frac{\partial^2 V}{\partial x^2}\right)_{x_{\text{cl}}}(\delta x)^2\right]$$

$$= \int_{t_0}^{t} \mathrm{d}\tau \left[\frac{m}{2}\dot{x}_{\mathrm{cl}}^2 - V(x_{\mathrm{cl}})\right] + \frac{1}{2}\int_{t_0}^{t} \mathrm{d}\tau \left[m(\delta\dot{x})^2 - \left(\frac{\partial^2 V}{\partial x^2}\right)_{x_{\mathrm{cl}}} (\delta x)^2\right]$$
$$+ \int_{t_0}^{t} \mathrm{d}\tau \left[\frac{\mathrm{d}}{\mathrm{d}\tau}\left(m\dot{x}_{\mathrm{cl}}\delta x\right) - \delta x \left\{m\ddot{x}_{\mathrm{cl}} + \left(\frac{\partial V}{\partial x}\right)_{x_{\mathrm{cl}}}\right\}\right] .$$

The last term in the right-hand side vanishes owing to the classical equation of motion and to the fact that $\delta x = 0$ at $\tau = t_0$ and at $\tau = t$. The first term is the action evaluated on the classical trajectory. One can thus write

$$S\left[x_{\mathrm{cl}} + \delta x\right] = S_{\mathrm{cl}} + \frac{1}{2}\int_{t_0}^{t} \mathrm{d}\tau \left[m(\delta\dot{x})^2 - \left(\frac{\partial^2 V}{\partial x^2}\right)_{x_{\mathrm{cl}}} (\delta x)^2\right] , \tag{16.32}$$

where $(\partial^2 V/\partial x^2)_{x_{\mathrm{cl}}}$ is a function of τ.

Equation (16.32) is an approximation of the action functional. For it to be used in the path integral it is clear that variation of the action, in units of $\hbar$, should be large for small variations of the trajectory. Indeed third-order terms in δx will then be significant only when the exponential's phase will change very quickly. This means they will have little influence, since trajectories with rapid phase changes little contribute to the path integral. The approximation that consists in expanding the action up to quadratic terms in δx is called *semiclassical.* The propagator is then given by

$$U_{\mathrm{scl}}(x, t; x_0, t_0) = \exp\left\{\frac{\mathrm{i}}{\hbar} S_{\mathrm{cl}}\right\} \int \mathcal{D}(\delta x)$$
$$\times \exp\left\{\frac{\mathrm{i}}{2\hbar}\int_{t_0}^{t} \mathrm{d}\tau \left[m(\delta\dot{x})^2 - \left(\frac{\partial^2 V}{dx^2}\right)_{x_{\mathrm{cl}}} (\delta x)^2\right]\right\} . \tag{16.33}$$

Note that the path integral here is carried out on $\mathcal{D}(\delta x)$. Owing to its definition δx vanishes at $\tau = t_0$ and at $\tau = t$.[2]

We now compute the path integral on $\mathcal{D}(\delta x)$. Let us introduce, in the interval (t_0, t), a complete set of real functions $x_j(\tau)$ such that

$$\int_{t_0}^{t} \mathrm{d}\tau\, x_j(\tau) x_k(\tau) = \delta_{jk} , \qquad x_j(t_0) = 0 = x_j(t) . \tag{16.34}$$

These functions will soon be specified further. At any rate one can write

$$\delta x(\tau) = \sum_{j=1}^{\infty} c_j x_j(\tau) , \tag{16.35}$$

---

[2] Equation (16.33) is valid if only one classical trajectory goes from (x_0, t_0) to (x, t). If there are many, the right-hand side should be summed over all classical trajectories.

where the c_j are suitable coefficients. The set of all trajectories $\delta x(\tau)$ corresponds to the set of all choices of coefficients $\{c_j\}$. Thus one can write

$$\mathcal{D}(\delta x) = \lim_{n\to\infty} J_n \prod_{j=1}^{n} \mathrm{d}c_j \,. \tag{16.36}$$

Here J_n is a Jacobian which, in general, depends on n and $t - t_0$ but not on the c_j, to which δx is related linearly.

Let us substitute the expression of $\delta x(\tau)$ in the exponential in (16.33). After integration by parts one gets

$$\begin{aligned}
&\frac{\mathrm{i}}{2\hbar}\int_{t_0}^{t} \mathrm{d}\tau \left[m(\delta\dot{x})^2 - \left(\frac{\partial^2 V}{\partial x^2}\right)_{x_{\mathrm{cl}}} (\delta x)^2 \right] \\
&\qquad = -\frac{\mathrm{i}}{2\hbar}\sum_{i,j=1}^{\infty} c_i c_j \int_{t_0}^{t} \mathrm{d}\tau \left[m\ddot{x}_i + \left(\frac{\partial^2 V}{\partial x^2}\right)_{x_{\mathrm{cl}}} x_i \right] x_j \,.
\end{aligned}$$

We now specify that the $x_i(\tau)$ are eigenfunctions of the operator

$$\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} + \frac{1}{m}\left(\frac{\partial^2 V}{\partial x^2}\right)_{x_{\mathrm{cl}}} \,. \tag{16.37}$$

Then

$$\ddot{x}_i + \frac{1}{m}\left(\frac{\partial^2 V}{\partial x^2}\right)_{x_{\mathrm{cl}}} x_i = \lambda_i x_i \,, \tag{16.38}$$

where the λ_i are eigenvalues. Orthonormality of the functions $x_i(\tau)$ implies that

$$\begin{aligned}
&U_{\mathrm{scl}}(x, t; x_0, t_0) \\
&= \exp\left\{\frac{\mathrm{i}}{\hbar} S_{\mathrm{cl}}\right\} \lim_{n\to\infty} J_n \int \left[\prod_{k=1}^{n} \mathrm{d}c_k\right] \exp\left\{ -\frac{\mathrm{i}m}{2\hbar} \sum_{i,j=1}^{n} c_i c_j \int_{t_0}^{t} \mathrm{d}\tau\, \lambda_i x_i x_j \right\} \\
&= \exp\left\{\frac{\mathrm{i}}{\hbar} S_{\mathrm{cl}}\right\} \lim_{n\to\infty} J_n \int \left[\prod_{k=1}^{n} \mathrm{d}c_k\right] \exp\left\{ -\frac{\mathrm{i}m}{2\hbar} \sum_{j=1}^{n} \lambda_j c_j^2 \right\} \\
&= \exp\left\{\frac{\mathrm{i}}{\hbar} S_{\mathrm{cl}}\right\} \lim_{n\to\infty} J_n \prod_{j=1}^{n} \left(\frac{2\pi\hbar}{\mathrm{i}m\lambda_j}\right)^{1/2} \\
&= \exp\left\{\frac{\mathrm{i}}{\hbar} S_{\mathrm{cl}}\right\} J \prod_{j=1}^{\infty} \lambda_j^{-1/2} \,,
\end{aligned} \tag{16.39}$$

where

$$J = \lim_{n\to\infty} J_n \left(\frac{2\pi\hbar}{\mathrm{i}m}\right)^{n/2} \tag{16.40}$$

is a normalization factor which, in general, depends on m and on $t - t_0$.

There remains to evaluate the product of the $\lambda_j^{-1/2}$, where λ_j is an eigenvalue of operator (16.37). One shows in Exercises (16.4) and (16.5) that

$$\prod_{j=1}^{\infty} \lambda_j^{-1/2} = C\left\{\psi(t)\right\}^{-1/2}, \tag{16.41}$$

where C is a factor independent of the potential V and where $\psi(\tau)$ satisfies the following differential equation and initial conditions:

$$\left\{\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} + \frac{1}{m}\left(\frac{\mathrm{d}^2V}{\mathrm{d}x^2}\right)_{x_{\mathrm{cl}}}\right\}\psi = 0; \qquad \psi(t_0) = 0, \qquad \left(\frac{\mathrm{d}\psi}{\mathrm{d}\tau}\right)_{t_0} = 1. \tag{16.42}$$

From (16.42) and the classical trajectory equation one easily checks that

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left\{\dot{x}_{\mathrm{cl}}\frac{\mathrm{d}\psi}{\mathrm{d}\tau} - \ddot{x}_{\mathrm{cl}}\psi\right\} = 0. \tag{16.43}$$

Using initial conditions on $\psi(\tau)$ one gets

$$\dot{x}_{\mathrm{cl}}\frac{\mathrm{d}\psi}{\mathrm{d}\tau} - \ddot{x}_{\mathrm{cl}}\psi = \dot{x}_{\mathrm{cl}}(t_0). \tag{16.44}$$

Equation (16.44) is equivalent to

$$\dot{x}_{\mathrm{cl}}^2\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\frac{\psi}{\dot{x}_{\mathrm{cl}}}\right) = \dot{x}_{\mathrm{cl}}(t_0),$$

from which we find that

$$\psi(t) = \dot{x}_{\mathrm{cl}}(t_0)\dot{x}_{\mathrm{cl}}(t)\int_{t_0}^{t} \mathrm{d}\tau\, \frac{1}{\dot{x}_{\mathrm{cl}}^2(\tau)}.$$

Substituting this result in (16.41) and (16.39), one gets for the propagator

$$U_{\mathrm{scl}}(x,t;x_0,t_0) = \exp\left\{\frac{\mathrm{i}}{\hbar}S_{\mathrm{cl}}\right\} JC\left[\dot{x}_{\mathrm{cl}}(t_0)\dot{x}_{\mathrm{cl}}(t)\int_{t_0}^{t}\mathrm{d}\tau\,\frac{1}{\dot{x}_{\mathrm{cl}}^2(\tau)}\right]^{-1/2}. \tag{16.45}$$

There remains to evaluate the constant JC. Since it does not depend on the potential one can use the free-particle propagator that was calculated earlier. If $V(x) = 0$,

$$\dot{x}_{\mathrm{cl}} = \frac{x - x_0}{t - t_0}, \qquad S_{\mathrm{cl}} = \frac{m}{2} \frac{(x - x_0)^2}{t - t_0}. \tag{16.46}$$

Substituting (16.46) in (16.45) and comparing the result with (16.30), one sees that

$$JC = \left\{ \frac{m}{2\pi \mathrm{i} \hbar} \right\}^{1/2}. \tag{16.47}$$

Hence

$$U_{\mathrm{scl}}(x, t; x_0, t_0) = \exp\left\{ \frac{\mathrm{i}}{\hbar} S_{\mathrm{cl}} \right\} \left[\frac{2\pi \mathrm{i} \hbar}{m} \dot{x}_{\mathrm{cl}}(t_0) \dot{x}_{\mathrm{cl}}(t) \int_{t_0}^{t} \mathrm{d}\tau \, \frac{1}{\dot{x}_{\mathrm{cl}}^2(\tau)} \right]^{-1/2}. \tag{16.48}$$

Expression (16.48) is the semiclassical approximation of the propagator. The only approximation made (in the case where the classical trajectory is unique) was to keep terms quadratic in δx only, in the expression of the action.

For a harmonic oscillator

$$V(x) = \frac{1}{2} m \omega^2 x^2 .$$

The semiclassical expression of the propagator is then exact. The classical trajectory is given by

$$x_{\mathrm{cl}}(\tau) = x_m \cos\left[\omega\left(\tau - \tau_m\right)\right], \tag{16.49}$$

where x_m is the maximum displacement of the particle and τ_m the corresponding time. Substituting this equation in the expression of the propagator one gets

$$U_{\mathrm{osc}}(x, t; x_0, t_0) = \exp\left\{ \frac{\mathrm{i}}{\hbar} S_{\mathrm{cl}} \right\} \left[\frac{\omega m}{2\pi \mathrm{i} \hbar \sin\left[\omega(t - t_0)\right]} \right]^{1/2}. \tag{16.50}$$

The classical action of the harmonic oscillator can be written in terms of x, t, x_0 and t_0 as (Exercise 16.6)

$$\begin{aligned} S_{\mathrm{cl}} &= \frac{m}{2} \int_{t_0}^{t} \mathrm{d}\tau \, \left(\dot{x}_{\mathrm{cl}}^2 - \omega^2 x_{\mathrm{cl}}^2 \right) \\ &= \frac{m}{2} \left[x_{\mathrm{cl}} \dot{x}_{\mathrm{cl}} \right]_{t_0}^{t} \\ &= \frac{m\omega}{2 \sin\left[\omega(t - t_0)\right]} \left\{ (x^2 + x_0^2) \cos\left[\omega(t - t_0)\right] - 2 x x_0 \right\}. \end{aligned} \tag{16.51}$$

We close this section by rewriting in a different way the semiclassical expression of the propagator. The energy E_{cl} associated with the classical trajectory is a function of x_0, x and $t - t_0$ implicitly defined as

$$t-t_0=\int_{x_0}^{x}\mathrm{d}x'\,\frac{1}{\dot{x}_{\mathrm{cl}}(x')}=\int_{x_0}^{x}\mathrm{d}x'(\pm 1)\left\{\frac{m}{2\left[E_{\mathrm{cl}}-V(x')\right]}\right\}^{1/2}. \tag{16.52}$$

The ±1 factor takes the sign of $\dot{x}_{\mathrm{cl}}$ into account. The action evaluated on the classical trajectory is equal to

$$\begin{aligned}S_{\mathrm{cl}}&=\int_{t_0}^{t}\mathrm{d}\tau\,\left[m\dot{x}_{\mathrm{cl}}^2-E_{\mathrm{cl}}\right]\\&=-E_{\mathrm{cl}}(t-t_0)+\int_{x_0}^{x}\mathrm{d}x' m\dot{x}_{\mathrm{cl}}(x')\\&=-E_{\mathrm{cl}}(t-t_0)+\hbar\int_{x_0}^{x}\mathrm{d}x' k_{\mathrm{cl}}(x'),\end{aligned} \tag{16.53}$$

where we used the fact that energy is constant on the classical trajectory and where a function $k_{\mathrm{cl}}(x)$ was defined as

$$k_{\mathrm{cl}}(x)=\frac{m\dot{x}_{\mathrm{cl}}}{\hbar}=\pm\frac{1}{\hbar}\left\{2m\left[E_{\mathrm{cl}}-V(x)\right]\right\}^{1/2}. \tag{16.54}$$

Substituting (16.53) and (16.54) in (16.48), one obtains for the semiclassical propagator

$$\begin{aligned}U_{\mathrm{scl}}(x,t;x_0,t_0)&=\exp\left\{-\frac{\mathrm{i}E_{\mathrm{cl}}}{\hbar}(t-t_0)+\mathrm{i}\int_{x_0}^{x}\mathrm{d}x' k_{\mathrm{cl}}(x')\right\}\\&\quad\cdot\left[2\pi\mathrm{i}k_{\mathrm{cl}}(x_0)k_{\mathrm{cl}}(x)\int_{x_0}^{x}\mathrm{d}x'\,\left[k_{\mathrm{cl}}(x')\right]^{-3}\right]^{-1/2}.\end{aligned} \tag{16.55}$$

16.4 WKB Wave Functions

Let $V(x)$ be a potential like the one shown in Fig. 16.3. Let $\psi(x,t)$ be the wave function associated with a stationary state of energy E, that is,

$$\psi(x,t)=\psi(x)\mathrm{e}^{-\mathrm{i}Et/\hbar}. \tag{16.56}$$

Let x_1 and x_2 be the points where $V(x)=E$. We will show that in the region where $x_1 \ll x \ll x_2$, wave functions like

$$\psi(x)=\frac{1}{\sqrt{k(x)}}\exp\left\{\mathrm{i}\int_{a}^{x}\mathrm{d}x' k(x')\right\} \tag{16.57}$$

follow from the semiclassical expression of the propagator. Here $x_1 \ll a \ll x_2$ and

$$k(x)=\pm\frac{1}{\hbar}\left\{2m\left[E-V(x)\right]\right\}^{1/2}. \tag{16.58}$$

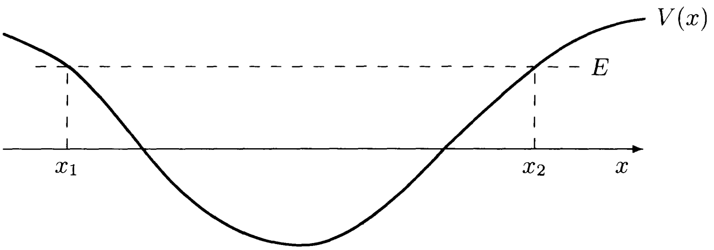

Fig. 16.3. Potential and energy of a stationary state

Note that at this stage E is not associated with a specific classical energy. Therefore the function $k(x)$ does not coincide with $k_{\mathrm{cl}}(x)$ defined in Sect. 16.3. Note also that expression (16.57) diverges around x_1 or x_2, so that it is certainly not valid in these regions.

The proof consists in showing that expression (16.56) for $\psi(x,t)$ approximately satisfies the equation

$$\psi(x,t)=\int_{-\infty}^{\infty}\mathrm{d}x_0\,U_{\mathrm{scl}}(x,t;x_0,t_0)\psi(x_0,t_0)\,. \tag{16.59}$$

For this purpose we will use the *stationary-phase approximation.* This can be formulated as follows. Let $f(x)$ and $g(x)$ be two real functions. Then

$$\begin{aligned}&\int_{-\infty}^{\infty}\mathrm{d}x\,f(x)\exp\{\mathrm{i}g(x)\}\\&\qquad\approx f(\bar{x})\exp\{\mathrm{i}g(\bar{x})\}\int_{-\infty}^{\infty}\mathrm{d}x\,\exp\left\{\frac{\mathrm{i}(x-\bar{x})^2}{2}\left(\frac{\mathrm{d}^2g}{\mathrm{d}x^2}\right)_{\bar{x}}\right\}\\&\qquad\approx f(\bar{x})\exp\{\mathrm{i}g(\bar{x})\}\left[\frac{2\pi\mathrm{i}}{\left(\frac{\mathrm{d}^2g}{\mathrm{d}x^2}\right)_{\bar{x}}}\right]^{1/2}.\end{aligned} \tag{16.60}$$

Here $\bar{x}$ is the point where $\mathrm{d}g/\mathrm{d}x$ vanishes, in other words the exponential's phase is stationary around $\bar{x}$. Equation (16.60) is valid provided that $f(x)$ does not change too much around $\bar{x}$ and that $g(x)$ is well represented by quadratic terms in the Taylor series. If $\mathrm{d}g/\mathrm{d}x$ has several zeros, the right-hand side of (16.60) must be summed over these.

In the right-hand side of (16.59), let us substitute the semiclassical expression (16.55) of the propagator and the form (16.56)–(16.57) proposed for $\psi(x_0,t_0)$. The result is an integral like

$$\int_{-\infty}^{\infty}\mathrm{d}x_0\,f(x_0)\exp\{\mathrm{i}g(x_0)\}\,, \tag{16.61}$$

where

$$f(x_0) = \left\{2\pi \mathrm{i} k(x_0) k_{\mathrm{cl}}(x_0) k_{\mathrm{cl}}(x) \int_{x_0}^{x} \mathrm{d}x' \left[k_{\mathrm{cl}}(x')\right]^{-3}\right\}^{-1/2}, \tag{16.62}$$

$$g(x_0) = -\frac{E_{\mathrm{cl}}}{\hbar}(t-t_0) - \frac{E}{\hbar}t_0 + \int_{x_0}^{x} \mathrm{d}x' k_{\mathrm{cl}}(x') + \int_{a}^{x_0} \mathrm{d}x' k(x') . \tag{16.63}$$

Keep in mind, here, that E_{cl} and k_{cl} depend on x_0. One has

$$\begin{aligned} \frac{\mathrm{d}g}{\mathrm{d}x_0} &= \left\{-\frac{t-t_0}{\hbar} + \int_{x_0}^{x} \mathrm{d}x' \frac{(\pm 1)}{2\hbar}\left[\frac{2m}{E_{\mathrm{cl}} - V(x')}\right]^{1/2}\right\} \frac{\partial E_{\mathrm{cl}}}{\partial x_0} \\ &\quad - k_{\mathrm{cl}}(x_0) + k(x_0) \\ &= -k_{\mathrm{cl}}(x_0) + k(x_0) , \end{aligned} \tag{16.64}$$

because, owing to (16.52), the expression inside curly brackets vanishes identically. The point $\bar{x}$ where $\mathrm{d}g/\mathrm{d}x_0$ vanishes is determined by the equation $k_{\mathrm{cl}}(x_0) = k(x_0)$, that is,

$$E_{\mathrm{cl}}\left(x_0 = \bar{x}, x, t-t_0\right) = E . \tag{16.65}$$

The second derivative of $g(x_0)$ at $\bar{x}$ is equal to

$$\begin{aligned} \left(\frac{\mathrm{d}^2 g}{\mathrm{d}x_0^2}\right)_{\bar{x}} &= \frac{\mathrm{d}}{\mathrm{d}x_0}\left\{k(x_0) - k_{\mathrm{cl}}(x_0)\right\}_{\bar{x}} \\ &= \pm\frac{\sqrt{2m}}{\hbar}\frac{\mathrm{d}}{\mathrm{d}x_0}\left\{[E - V(x_0)]^{1/2} - [E_{\mathrm{cl}} - V(x_0)]^{1/2}\right\}_{\bar{x}} \\ &= -\frac{m}{\hbar^2 k(\bar{x})}\left(\frac{\partial E_{\mathrm{cl}}}{\partial x_0}\right)_{\bar{x}} \\ &= \frac{1}{[k(\bar{x})]^2}\left\{\int_{\bar{x}}^{x} \mathrm{d}x' \left[k(x')\right]^{-3}\right\}^{-1} , \end{aligned}$$

where the derivative of E_{cl} is calculated by means of (16.52). Substituting this result in (16.60) one gets with some manipulations

$$\begin{aligned} \psi(x,t) &= \int_{-\infty}^{\infty} \mathrm{d}x_0\, U(x,t;x_0,t_0)\psi(x_0,t_0) \\ &= \frac{1}{\sqrt{k(x)}} \exp\left\{\mathrm{i}\int_{a}^{x} \mathrm{d}x' k(x') - \frac{\mathrm{i}Et}{\hbar}\right\} , \end{aligned} \tag{16.66}$$

which ends the proof. We have shown that, in the stationary-phase approximation, wave functions (16.56)–(16.57) satisfy the propagation equation.

The result just obtained could have been proved, in fact more simply, directly from Schrödinger's equation.[3] Our proof was meant to illustrate the

---

[3] See for instance [138], Sect. 46, [159], Sect. 7.1 and [200], Sect. 34.

use of the semiclassical expression of the propagator. The expression we found for the wave functions is called the *WKB approximation.*[4]

From now on it will be convenient to set

$$k(x) = \frac{1}{\hbar}\left\{2m\left[E - V(x)\right]\right\}^{1/2} \tag{16.67}$$

and to include the ± signs in the expression for wave functions. Thus in a region where $E - V(x) > 0$, the most general WKB wave function is given by

$$\psi(x) = \frac{c_1}{\sqrt{k(x)}}\exp\left\{\mathrm{i}\int_a^x \mathrm{d}x' k(x')\right\} + \frac{c_2}{\sqrt{k(x)}}\exp\left\{-\mathrm{i}\int_a^x \mathrm{d}x' k(x')\right\}. \tag{16.68}$$

In a region where $E - V(x) < 0$ we let

$$\kappa(x) = \frac{1}{\hbar}\left\{2m\left[V(x) - E\right]\right\}^{1/2}. \tag{16.69}$$

One can show that the WKB wave function is then given by

$$\psi(x) = \frac{c_1'}{\sqrt{\kappa(x)}}\exp\left\{-\int_{a'}^x \mathrm{d}x' \kappa(x')\right\} + \frac{c_2'}{\sqrt{\kappa(x)}}\exp\left\{\int_{a'}^x \mathrm{d}x' \kappa(x')\right\}, \tag{16.70}$$

where a' is a point in the region. Note that none of these expressions is valid in a region where $E - V(x) \approx 0$. The way to connect solutions for $E - V(x) < 0$ and $E - V(x) > 0$ is subtle, and will now be dealt with.

16.5 Turning Point

A *turning point* (beyond which a classical particle cannot reach) is a point where $E = V(x)$. Examples are points x_1 and x_2 in Fig. 16.3. WKB wave functions are valid in a region far from turning points.

The Hamiltonian's eigenvalue equation can be written as

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + \left[k(x)\right]^2\psi = 0, \tag{16.71}$$

where $k(x)$ is given in (16.67). Let x_1 be the turning point shown in Fig. 16.3. Assume that around x_1 one has approximately

---

[4] The WKB approximation gets its name from G. Wentzel, H. A. Kramers and L. Brillouin who independently developed it in 1926. A good exposition of the relation between path integrals and the WKB approximation can be found in [116]. Reference [54] applies semiclassical methods to many problems in quantum mechanics and quantum field theory.

$$k^2(x) = c(x - x_1)\,, \tag{16.72}$$

where c is a positive constant. We will take x as a complex variable. The Hamiltonian's eigenvalue equation then has two linearly independent exact solutions, that are everywhere analytic functions of x. They are (Exercise 16.8)

$$\psi_\pm(x) = a_\pm\, \xi^{1/2} k^{-1/2} J_{\pm 1/3}(\xi)\,, \tag{16.73}$$

where $J_{\pm 1/3}$ are Bessel functions and where

$$\xi(x) = \int_{x_1}^{x} \mathrm{d}x' k(x')\,. \tag{16.74}$$

We emphasize that functions $\psi_\pm(x)$ are everywhere analytic, even though the factors that make them up have cuts, for instance along negative values of $x - x_1$.

We will obtain the asymptotic form of solutions $\psi_\pm(x)$ for large (positive or negative) values of $x - x_1$. We will see that it coincides with WKB solutions. From this we will find a way to connect WKB solutions on both sides of a turning point.[5]

The asymptotic form of Bessel functions $J_\nu(\xi)$, if $|\xi|$ is large and $|\arg \xi| < \pi$, is given by (7.190):

$$J_\nu(\xi) \to \left(\frac{2}{\pi\xi}\right)^{1/2} \cos\left\{\xi - \pi\left(\frac{1}{4} + \frac{\nu}{2}\right)\right\}. \tag{16.75}$$

The asymptotic form of $\psi_\pm(x)$ for $x - x_1 \gg 0$ is therefore given by

$$\psi_\pm(x) \to a_\pm \left(\frac{2}{\pi k}\right)^{1/2} \cos\left\{\xi - \frac{\pi}{12}(3 \pm 2)\right\}, \tag{16.76}$$

which essentially coincides with WKB wave functions for $E - V(x) > 0$.

The situation is more delicate for negative values of $x - x_1$. For any complex x let $x - x_1 = |x - x_1| \mathrm{e}^{\mathrm{i}\theta}$. One has

$$k(x) = [c(x - x_1)]^{1/2} = [c\,|x - x_1|]^{1/2}\, \mathrm{e}^{\mathrm{i}\theta/2} = \kappa(x) \mathrm{e}^{\mathrm{i}\theta/2}\,, \tag{16.77}$$

$$\begin{aligned}\xi(x) &= \int_{x_1}^{x} \mathrm{d}x' k(x') = \int_{x_1}^{x} \mathrm{d}x'\, c^{1/2}(x' - x_1)^{1/2} = \frac{2}{3} c^{1/2} (x - x_1)^{3/2} \\ &= \zeta(x) \mathrm{e}^{3\mathrm{i}\theta/2}\,.\end{aligned} \tag{16.78}$$

Here κ and ζ are positive functions and

---

[5] The interval $x - x_1$ should be large enough for the solutions' asymptotic form to be used, and small enough for approximation (16.72) to be valid. This is not inconsistent if the energy level shown in Fig. 16.3 is sufficiently high.

$$\zeta(x) = \int_{x_1}^{x_1+|x-x_1|} \mathrm{d}x'\kappa(x') = \int_{x}^{x_1} \mathrm{d}x'\kappa(x') . \tag{16.79}$$

From the series expansion (7.187) of Bessel functions one can see that for any real ν and any α that is an integral multiple of π,

$$J_\nu\left(z\mathrm{e}^{\mathrm{i}\alpha}\right) = \mathrm{e}^{\mathrm{i}\nu\alpha} J_\nu(z) . \tag{16.80}$$

Thus for $\theta = \pi$,

$$J_\nu(\xi) = J_\nu\left\{\zeta\mathrm{e}^{3\mathrm{i}\pi/2}\right\} = \mathrm{e}^{\mathrm{i}\pi\nu} J_\nu\left\{\zeta\mathrm{e}^{\mathrm{i}\pi/2}\right\} . \tag{16.81}$$

The asymptotic form of $\psi_\pm(x)$ for $x - x_1 \ll 0$ is therefore given by

$$\begin{aligned}
\psi_\pm(x) &= a_\pm \zeta^{1/2}\mathrm{e}^{3\mathrm{i}\pi/4}\kappa^{-1/2}\mathrm{e}^{-\mathrm{i}\pi/4}\mathrm{e}^{\pm\mathrm{i}\pi/3} J_{\pm 1/3}\left\{\zeta\mathrm{e}^{\mathrm{i}\pi/2}\right\} \\
&\to a_\pm \kappa^{-1/2}\mathrm{e}^{\mathrm{i}\pi(1/2\pm 1/3)}\zeta^{1/2}\left\{\frac{2}{\pi\zeta\mathrm{e}^{\mathrm{i}\pi/2}}\right\}^{1/2} \cos\left\{\mathrm{i}\zeta - \frac{\pi}{12}(3\pm 2)\right\} \\
&\to a_\pm \left(\frac{2}{\pi\kappa}\right)^{1/2} \frac{1}{2}\left\{\mp\mathrm{e}^{\zeta} + \mathrm{e}^{-\zeta}\mathrm{e}^{\pm\mathrm{i}\pi/6}\right\} .
\end{aligned} \tag{16.82}$$

We now let

$$\left(\frac{2}{\pi}\right)^{1/2} a_\pm = a , \tag{16.83}$$

and take the sum of ψ_- and ψ_+. We get

$$\psi_- + \psi_+ \to \begin{cases} ak^{-1/2}\left\{\cos\left(\xi - \frac{\pi}{12}\right) + \cos\left(\xi - \frac{5\pi}{12}\right)\right\} & \\ \qquad = \sqrt{3}ak^{-1/2}\cos\left(\xi - \frac{\pi}{4}\right) & \text{if } x \gg x_1 , \\ \frac{\sqrt{3}}{2}a\kappa^{-1/2}\mathrm{e}^{-\zeta} & \text{if } x \ll x_1 . \end{cases}$$

If, on the other hand, we take the difference between ψ_- and ψ_+, we obtain

$$\psi_- - \psi_+ \to \begin{cases} ak^{-1/2}\left\{\cos\left(\xi - \frac{\pi}{12}\right) - \cos\left(\xi - \frac{5\pi}{12}\right)\right\} & \\ \qquad = ak^{-1/2}\cos\left(\xi + \frac{\pi}{4}\right) & \text{if } x \gg x_1 , \\ a\kappa^{-1/2}\mathrm{e}^{\zeta} + O\left(\mathrm{e}^{-\zeta}\right) \approx a\kappa^{-1/2}\mathrm{e}^{\zeta} & \text{if } x \ll x_1 . \end{cases}$$

Table 16.1 summarizes how to connect, on both sides of a turning point, different solutions obtained by the WKB approximation.

It is very important to point out that in Table 16.1 arrows are drawn in one direction only. This restriction is related to the fact that the potential does not satisfy (16.72) exactly. Hence the relative coefficients we found for e^{ζ} and $\mathrm{e}^{-\zeta}$ terms are only approximate. Thus suppose one knows that $\psi(x) \sim k^{-1/2}\cos(\xi - \pi/4)$ for $x \gg x_1$. One cannot then eliminate the presence of a small term proportional to e^{ζ} for $x \ll x_1$, which finally will overwhelm $(1/2)\kappa^{-1/2}\mathrm{e}^{-\zeta}$. Likewise, suppose the asymptotic form of $\psi(x)$ is given

Table 16.1. Connection formulas at a turning point

$x \ll x_1$		$x \gg x_1$
$\frac{1}{2}\kappa^{-1/2}e^{-\zeta}$	$\rightarrow$	$k^{-1/2}\cos\left(\xi - \frac{\pi}{4}\right)$
$\kappa^{-1/2}e^{+\zeta}$	$\leftarrow$	$k^{-1/2}\cos\left(\xi + \frac{\pi}{4}\right)$

by $\kappa^{-1/2}\mathrm{e}^{\zeta}$ for $x \ll x_1$. The coefficient of the $\mathrm{e}^{-\zeta}$ term is then completely unknown. One cannot conclude that $\psi(x) \sim k^{-1/2}\cos(\xi + \pi/4)$ for $x \gg x_1$.

The connection formulas are easily adapted to the case where the potential around the turning point is an increasing function, i.e. when constant c in (16.72) is negative. This corresponds to point x_2 in Fig. 16.3. Provided functions ζ and ξ are defined so as to be positive, Table 16.1 stays the same, except that headlines $x \ll x_1$ and $x \gg x_1$ must be interchanged.

The connection formulas, when used in the appropriate direction, link WKB wave functions on both sides of a turning point. We now apply these results to two particularly relevant situations.

16.6 The Bohr–Sommerfeld Rule

We first use WKB wave functions for the approximate computation of eigenvalues of the Hamiltonian associated with the potential shown in Fig. 16.3. Here E is an eigenvalue while x_1 and x_2 are turning points. We assume that for this value of E there are no other turning points. Furthermore we assume that E is sufficiently high for the WKB approximation to be valid, far enough of course from turning points.

Wave functions in the discrete spectrum vanish as $x \to \pm\infty$. WKB solutions in regions where $x \to \pm\infty$ will thus involve only the exponentially decreasing term, so that

$$\psi(x) = \begin{cases} \frac{c_1}{2\sqrt{\kappa(x)}} \exp\left\{-\int_x^{x_1} \mathrm{d}x'\kappa(x')\right\} & \text{if } x \ll x_1\,, \\ \frac{c_2}{2\sqrt{\kappa(x)}} \exp\left\{-\int_{x_2}^{x} \mathrm{d}x'\kappa(x')\right\} & \text{if } x \gg x_2\,. \end{cases} \tag{16.84}$$

Connection formulas can be used from a region of exponential decrease to a region of oscillatory behavior. In the present case the connection can be established through x_1 and through x_2. If we first connect through x_1, we obtain the wave function in region $x_1 \ll x \ll x_2$ as

$$\psi(x) = \frac{c_1}{\sqrt{k(x)}} \cos\left\{\int_{x_1}^{x} \mathrm{d}x' k(x') - \frac{\pi}{4}\right\}, \qquad x_1 \ll x \ll x_2. \tag{16.85}$$

Connecting now through x_2, we once more obtain the wave function between turning points, this time as

$$\psi(x) = \frac{c_2}{\sqrt{k(x)}} \cos\left\{\int_x^{x_2} \mathrm{d}x' k(x') - \frac{\pi}{4}\right\}, \qquad x_1 \ll x \ll x_2. \tag{16.86}$$

Thus we have two different expressions of the wave function in region $x_1 \ll x \ll x_2$, which must coincide. To see whether this is the case it is useful to rewrite (16.85) as

$$\psi(x) = \frac{c_1}{\sqrt{k(x)}} \cos\left\{\int_x^{x_2} \mathrm{d}x' k(x') - \frac{\pi}{4} + \left[\frac{\pi}{2} - \int_{x_1}^{x_2} \mathrm{d}x' k(x')\right]\right\}. \tag{16.87}$$

Clearly (16.87) and (16.86) will coincide only if $|c_1| = |c_2|$ and if the cosine's arguments differ by an integral multiple of π. This entails that

$$\int_{x_1}^{x_2} \mathrm{d}x' k(x') = \left(n + \frac{1}{2}\right)\pi, \qquad n = 0, 1, 2, \ldots \tag{16.88}$$

To interpret (16.88) one should recall that x_1, x_2 and k are functions of energy. One has

$$\int_{x_1(E)}^{x_2(E)} \mathrm{d}x' \left[\frac{2m}{\hbar^2}\left[E - V(x')\right]\right]^{1/2} = \left(n + \frac{1}{2}\right)\pi. \tag{16.89}$$

This, in fact, is an equation for E, whose solutions are the semiclassical approximation of energy eigenvalues. These values of E correspond to the *Bohr–Sommerfeld quantization rule.*[6] The approximations we made suggest that this rule may be the more valid as n gets larger. But surprisingly, (16.89) is sometimes a good approximation even for the first few excited states.

The quantization equation for E can also be interpreted in a different way, where it finds an application in statistical physics and the study of macroscopic bodies. The term $\{2m[E - V(x)]\}^{1/2}$ represents the classical momentum of a particle of energy E in a potential $V(x)$. Consider the phase space (p, x) associated with the classical problem. The integral $2\int_{x_1}^{x_2} \mathrm{d}x' p(x')$ is equal to the area in phase space associated with all classical trajectories (tracked in both directions) with energy lower than E. On the other hand, n represents the number of quantum states with energy lower than E. An area element $\Delta p \Delta x$ in classical phase space therefore corresponds to Δn quantum states, where $\Delta n = \Delta x \Delta p / 2\pi\hbar$. This result can be generalized to a system

---

[6] Equation (16.89) can also be obtained directly from the semiclassical propagator, without using connection formulas. One makes use of (11.14), where $U(x, t; x, 0)$ is replaced by $U_{\mathrm{scl}}(x, t; x, 0)$. Note that all trajectories in the path integral are then closed loops and that each of these loops can be tracked n times. The interested reader is referred to [62] and [191], Sect. 4. Reference [161] applies the WKB approximation to multidimensional systems.

with s degrees of freedom. In the volume element $\Delta q_1 \dots \Delta q_s \Delta p_1 \dots \Delta p_s$ in phase space there are about Δn quantum states, where

$$\Delta n = \frac{\Delta q_1 \dots \Delta q_s \Delta p_1 \dots \Delta p_s}{(2\pi\hbar)^s} . \tag{16.90}$$

This number quickly reaches beyond imagination. Let 10^{23} particles fill a one-meter-long cube. Take momentum intervals $\Delta p_i = m_p \times$ 1 m s^{-1}, where m_p is the proton's mass. Then $\Delta n \approx \exp(5 \times 10^{24})$.

16.7 Potential Barrier

Consider a beam of particles of energy E coming on the potential barrier shown in Fig. 16.4. We assume the barrier is high and wide so that only a small fraction of particles is transmitted. We will compute the probability of tunneling through the barrier, with the help of the WKB approximation.

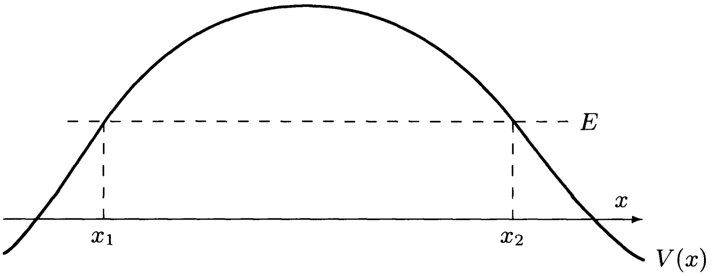

Fig. 16.4. Potential barrier

If the incoming beam comes from the region where $x \ll x_1$, we expect in that region a reflected as well as an incoming wave, and a transmitted wave in the region where $x \gg x_2$. Thus we look for a WKB solution of the form

$$\psi(x) = \begin{cases} \frac{c_1}{\sqrt{k(x)}} \exp\left\{ \mathrm{i} \int_x^{x_1} \mathrm{d}x' k(x') \right\} \\ \quad + \frac{c_2}{\sqrt{k(x)}} \exp\left\{ -\mathrm{i} \int_x^{x_1} \mathrm{d}x' k(x') \right\} & \text{if } x \ll x_1 , \\ \frac{c}{\sqrt{k(x)}} \exp\left\{ \mathrm{i} \int_{x_2}^{x} \mathrm{d}x' k(x') \right\} & \text{if } x \gg x_2 . \end{cases} \tag{16.91}$$

Suppose that $V(x) \to V_\infty$ (a constant) as $x \to \pm\infty$. Evaluating the probability current on both sides of the barrier, one easily shows that the reflection coefficient $\mathcal{R}$ and the transmission coefficient $\mathcal{T}$ are given by

$$\mathcal{R} = \left| \frac{c_1}{c_2} \right|^2 , \qquad \mathcal{T} = \left| \frac{c}{c_2} \right|^2 . \tag{16.92}$$

Let us now make the connections through points x_2 and x_1. In the region where $x \gg x_2$ the wave function can be written as

$$\psi(x) = \frac{c\mathrm{e}^{-\mathrm{i}\pi/4}}{\sqrt{k(x)}}\left[\cos\left\{\int_{x_2}^{x}\mathrm{d}x' k(x') + \frac{\pi}{4}\right\} + \mathrm{i}\cos\left\{\int_{x_2}^{x}\mathrm{d}x' k(x') - \frac{\pi}{4}\right\}\right]. \tag{16.93}$$

The first term connects to an increasing exponential in the middle region (increasing from x_2, of course). The second term, for all practical purposes, connects to a decreasing exponential, and thus carries a negligible contribution. This means that in the middle region, the WKB solution has the form

$$\begin{aligned}\psi(x) &= \frac{c\mathrm{e}^{-\mathrm{i}\pi/4}}{\sqrt{\kappa(x)}}\exp\left\{\int_{x}^{x_2}\mathrm{d}x'\kappa(x')\right\} \\ &= \left[c\mathrm{e}^{-\mathrm{i}\pi/4}\exp\left\{\int_{x_1}^{x_2}\mathrm{d}x'\kappa(x')\right\}\right] \\ &\quad\times\frac{1}{\sqrt{\kappa(x)}}\exp\left\{-\int_{x_1}^{x}\mathrm{d}x'\kappa(x')\right\}, \qquad x_1 \ll x \ll x_2\,. \end{aligned} \tag{16.94}$$

The second way to write $\psi(x)$ suggests how to make the connection through x_1. Letting

$$P^{-1/2} = \exp\left\{\int_{x_1}^{x_2}\mathrm{d}x'\kappa(x')\right\}, \tag{16.95}$$

one gets for the WKB wave function in the region $x \ll x_1$:

$$\begin{aligned}\psi(x) &= \frac{c\mathrm{e}^{-\mathrm{i}\pi/4}P^{-1/2}}{\sqrt{k(x)}}2\cos\left\{\int_{x}^{x_1}\mathrm{d}x'k(x') - \frac{\pi}{4}\right\} \\ &= \frac{cP^{-1/2}}{\sqrt{k(x)}}\left[-\mathrm{i}\exp\left\{\mathrm{i}\int_{x}^{x_1}\mathrm{d}x'k(x')\right\} + \exp\left\{-\mathrm{i}\int_{x}^{x_1}\mathrm{d}x'k(x')\right\}\right].\end{aligned}$$

Thus

$$c_1 = -\mathrm{i}cP^{-1/2}, \qquad c_2 = cP^{-1/2}, \tag{16.96}$$

from which one finds that

$$\mathcal{R} = 1, \qquad \mathcal{T} = P. \tag{16.97}$$

The quantity P is called the *penetration factor* and is very small for a wide and high barrier.[7] Note that $\mathcal{R} + \mathcal{T} = 1 + P \approx 1$. The fact that $\mathcal{R} + \mathcal{T}$ is not exactly equal to 1 is due to approximations inherent in the WKB method.

---

[7] References [48] and [117] derive the penetration factor from the semiclassical propagator, without using connection formulas.

The evaluation of the penetration factor can be adapted to the case of a spherically symmetric barrier in three dimensions. We know that $ru(r)$, with $u(r)$ the radial function, then satisfies the same equation as the function $\psi(x)$ in one dimension, except that the effective potential of (7.91) must be used. Insofar as radial functions have an appropriate asymptotic form, the penetration factor is given by (16.95), provided that the effective potential is used and the integral is carried out on the radial coordinate. This calculation has an important application in the analysis of α decay ([32], Sect. 100; [78], Chap. 10; [176], Chap. 12). This is the emission of a helium nucleus by a heavy nucleus, a process that is energetically favorable in many nuclear isotopes. The potential barrier comes from the combined effect of the protons' electrostatic repulsion and the attractive nuclear force. Reference [32] (Sects. 98 and 101) also illustrates the passage through a potential barrier by analyzing the "cold emission" of electrons by metals and the ionization of atoms in strong electric fields.

16.8 Classical Evolution

The semiclassical approximation is valid when the action associated with the classical trajectory is a high multiple of $\hbar$. This multiple corresponds to the variable n in (16.89). In typical atomic problems involving potential wells or barriers the value of n can be of the order of 10 or 100.

That value, obviously, is nothing compared with the action associated with genuine classical objects. One J s, for instance, is $10^{34}\,\hbar$. In this section we will inquire whether it is possible to describe such objects by the quantum-mechanical formalism. In other words, to what extent does the quantum-mechanical formalism reduce to the classical one when the action is large enough?

Ehrenfest's relations, developed in Sect. 6.6, go some way toward answering that question. So does the path integral. Indeed suppose that the action associated with the classical trajectory is large, say of order $10^{34}\,\hbar$. Clearly then, only trajectories which at all time are extremely close to the classical trajectory make a significant contribution to the path integral. This suggests that in this situation, the classical trajectory by itself characterizes the system's evolution.

To make this more precise, let us first point out that a classical object is not just any macroscopic aggregate of microobjects.[8] It is a collection of particles linked together in all kinds of ways. That collection is characterized by a large number of commuting dynamical variables called *collective coordinates*, and by conjugate momenta. The center-of-mass position is an example

---

[8] This section owes much to Omnès ([172]; [173], Chap. 6), where one can find precise formulations of results that we state more qualitatively, as well as a number of proofs.

of a collective coordinate. The number of these coordinates can be rather large, but it is a very small fraction of the number of microscopic coordinates of the object. We denote by Q a generic collective coordinate and by P the conjugate momentum.

We assume that the object's exact Hamiltonian can be written in the form

$$H = H_{\mathrm{c}} + H_{\mathrm{m}} + H_{\mathrm{cm}} \, . \tag{16.98}$$

Here H_{c} only depends on collective coordinates, and H_{m} only depends on microscopic coordinates other than collective coordinates. For the time being we neglect the interaction term H_{cm}.

A *classical property* consists in specifying, up to some uncertainty, values of the collective coordinates and conjugate momenta. The uncertainty is small on a macroscopic scale but large on a microscopic scale (so as to easily satisfy Heisenberg's principle). A classical property corresponds to a projector F. Let ϱ be the object's exact density operator. We say the object is in a *classical state* if there exists an F (associated with a classical property) such that $F\varrho F = \varrho$ (up to small corrections). Inasmuch as H_{cm} can be neglected, one can assume that F and ϱ are restricted to the Hilbert space associated with collective coordinates. The time evolution of ϱ is given by (12.9) where

$$U(t, t_0) = \exp\left\{ -\frac{\mathrm{i}}{\hbar} H_{\mathrm{c}}(t - t_0) \right\} . \tag{16.99}$$

One should note that even though the expressions classical property and classical state were introduced, the object's description is at this stage entirely made within the quantum-mechanical formalism. We now introduce the classical dynamical variables. Let Ω be a Hermitian operator acting in the Hilbert space associated with collective coordinates. With this operator a function $\omega(q, p)$ is associated in the following way:

$$\omega(q,p) = \int \mathrm{d}q' \, \mathrm{d}q'' \, \langle q' | \Omega | q'' \rangle \delta \left\{ q - \frac{q' + q''}{2} \right\} \exp\left\{ -\frac{\mathrm{i}p}{\hbar}(q' - q'') \right\} . \tag{16.100}$$

One can show that transformation (16.100) is the most natural way to associate a classical variable with a quantum operator. We simply point out here that (16.100) is linear, that it transforms I, Q and P into 1, q and p, and that it transforms a Hermitian operator into a real function (Exercise 16.14).[9]

---

[9] The importance of transformation (16.100) is not limited to a discussion of classical properties. Indeed let us apply it to the projector $|\psi(t)\rangle\langle\psi(t)|$, where $|\psi(t)\rangle$ is a normalized state vector. One gets (if q and p are one-dimensional variables)

$$\omega(q,p) = 2 \int ds \, \psi(q - s) \psi^*(q + s) \mathrm{e}^{2\mathrm{i}ps/\hbar} .$$

This is Wigner's *phase-space distribution function*, which has numerous applications [132].

The classical (in general multidimensional) variables q and p define phase space. The classical Hamiltonian is obtained from (16.100) by letting $\Omega = H_c$. It determines the time evolution of q and p through Hamilton's equations.

In classical mechanics a *property* consists in specifying that the values of q and p are in a certain region of phase space, that we can call C. We assume that the shape of C is smooth enough and that the volume of C is much larger than h^N (where N is the number of collective coordinates). Thus the region corresponds, through (16.90), to a very large number of quantum states.

With region C in phase space one can associate a quantum operator acting in the space of collective variables. For this purpose consider the *characteristic function* of C, i.e. the function $f(q,p)$ equal to 1 if (q,p) is in C and to 0 otherwise. This function determines a Hermitian operator through (16.100). That is the operator we are looking for. For all practical purposes it coincides with a projector.[10]

Let C_1 and C_2 be two regular regions in phase space, to which correspond projectors or quasi-projectors F_1 and F_2. One can show that to the region $C_1 \cap C_2$ corresponds essentially the quasi-projector $F_1F_2 \approx F_2F_1$, and that to the region $C_1 \cup C_2$ corresponds essentially the quasi-projector $F_1 + F_2 - F_1F_2$. This entails that classical properties defined in the quantum-mechanical formalism and properties defined in the classical formalism correspond to each other consistently.

Consider, at time t_0, a region C_0 in phase space. Each point in C_0 evolves with time according to Hamilton's equations. The region itself is thus transformed, at time t, into a region C_t. Inasmuch as C_0 and C_t are sufficiently regular, one can associate with them quasi-projectors F_0 and F_t. One can show that in a precise sense F_t is essentially equal to $U(t,t_0)F_0U^\dagger(t,t_0)$. This means that classical properties, defined in the quantum-mechanical formalism, essentially evolve as properties in the classical formalism.

Let $t_1 < t_2 < \cdots < t_n$ be a sequence of instants in the internal (t_0,t). With these instants let us associate quasi-projectors F_1, F_2, ..., F_n. They correspond to the regions C_1, C_2, ..., C_n in phase space into which C_0 transforms. In addition consider the complementary quasi-projectors $I - F_i$ $(i = 1, 2, \ldots, n)$. Let ϱ be a density operator such that $F_0\varrho F_0 = \varrho$. One can show that the complete family of histories based on properties F_i and $I - F_i$ at times t_i is essentially consistent. So is it with the logical structure that can be defined through this family. In that logic propositions associated with F_i and F_j imply each other. This means that classical determinism falls within the scope of the interpretation rule on p. 271.

---

[10] Technically it is preferable to use a smoother function f, i.e. a function that does not change abruptly from 1 to 0. There result a number of operators instead of only one. In a precise sense, all these operators are close to one another. They are *quasi-projectors*, i.e. operators for which almost all eigenvalues are close to 0 or 1.

Exercises

16.1. The general formula for the propagator is given in (16.10).
a) Let

$$H(p,x,\tau) = H(p,x) = \frac{p^2}{2}V(x)\,.$$

Show that the Lagrangian that corresponds to this Hamiltonian is given by

$$\mathcal{L}(x,\dot{x}) = \frac{\dot{x}^2}{2V(x)}\,.$$

b) Show that the propagator can be written as

$$U(x,t;x_0,t_0) = \int \frac{\mathcal{D}x}{\sqrt{V(x)}} \exp\left\{\frac{\mathrm{i}}{\hbar}\int_{t_0}^{t} \mathrm{d}\tau\, \mathcal{L}(x,\dot{x})\right\}.$$

Define the element $\mathcal{D}x/\sqrt{V(x)}$ by the appropriate limiting process.

16.2. Fill in the missing steps in the computation of integral $I(\alpha)$ in Sect. 16.2.

16.3. Prove (16.29) by induction.

16.4. Let $W(\tau)$ be a real and bounded function in the interval (t_0, t). Denote by D the linear operator $\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} + W(\tau)$. Let $\{\lambda_n\}$ be the set of eigenvalues of D, with boundary conditions $x(t_0) = 0 = x(t)$. Let $x_\lambda(\tau)$ be the unique function that satisfies the following differential equation and initial conditions:

$$Dx(\tau) = \lambda x(\tau); \qquad x(t_0) = 0, \qquad \left\{\frac{\mathrm{d}}{\mathrm{d}\tau}x(\tau)\right\}_{t_0} = 1.$$

The function $x_\lambda(t)$ is an analytic function of λ with simple zeros at $\lambda = \lambda_n$. Show that if $|\lambda|$ is large and if $0 \neq \arg(\lambda) \neq \pi$, then

$$x_\lambda(\tau) \to \frac{1}{\sqrt{\lambda}} \sinh\left\{\sqrt{\lambda}(\tau - t_0)\right\}.$$

16.5. Let $\lambda_n^{(1)}$, $\lambda_n^{(2)}$, $x_\lambda^{(1)}(\tau)$ and $x_\lambda^{(2)}(\tau)$ be defined through two functions $W_1(\tau)$ and $W_2(\tau)$ as in Exercise (16.4). Let

$$F(\lambda) = \prod_{n=1}^{\infty} \frac{\lambda_n^{(1)} - \lambda}{\lambda_n^{(2)} - \lambda}, \qquad G(\lambda) = \frac{x_\lambda^{(1)}(t)}{x_\lambda^{(2)}(t)}.$$

F and G are two meromorphic functions of λ with simple zeros at $\lambda_n^{(1)}$ and simple poles at $\lambda_n^{(2)}$.

a) Use the result of Exercise (16.4) to show that if $|\lambda|$ is large and if $0 \neq \arg(\lambda) \neq \pi$, then $G(\lambda) \to 1$.
b) Give a reason why $F(\lambda) \to 1$ if $|\lambda|$ is large and if $0 \neq \arg(\lambda) \neq \pi$.
c) Conclude from (a) and (b) that $F(\lambda)/G(\lambda) = 1$ for any λ, hence that (16.41) is true.

16.6. The trajectory of a classical harmonic oscillator is given by (16.49). Show that the classical action of an oscillator going from x_0 to x in time $t - t_0$ is equal to

$$S_{\mathrm{cl}} = \frac{m\omega}{2\sin\left[\omega(t-t_0)\right]}\left\{(x^2 + x_0^2)\cos\left[\omega(t-t_0)\right] - 2xx_0\right\}.$$

16.7. Derive the propagator associated with a linear potential. (See [115].)

16.8. Let $k^2(x) = Cx^n$. Show that the equation

$$\frac{\mathrm{d}^2\psi}{\mathrm{d}x^2} + k^2(x)\psi = 0$$

has the solutions

$$\psi(x) = a\xi^{1/2}k^{-1/2}J_{\pm m}(\xi).$$

Here

$$\xi(x) = \int_0^x \mathrm{d}x' k(x'), \qquad m = \frac{1}{n+2},$$

and $J_{\pm m}(\xi)$ are Bessel functions, satisfying (7.186).

16.9. A particle of mass m is in a potential $V(x)$ given by

$$V(x) = \begin{cases} -V_0\left(1 - \frac{|x|}{a}\right) & \text{if } |x| < a\,, \\ 0 & \text{if } |x| > a\,. \end{cases}$$

Use the WKB approximation to find all bound-state energies in the case where $mV_0a^2/\hbar^2 = 100$.

16.10. Let $V(x)$ be a potential that is everywhere negative, has a finite lower bound and is proportional to $|x|^{-m}$ if $|x|$ is large. Use the WKB approximation to show that $V(x)$ supports an infinite number of bound states only if $m \leq 2$.

16.11. Show that the WKB approximation, applied to the harmonic oscillator $[V(x) = m\omega^2x^2/2]$, yields the exact value of energy for all levels.

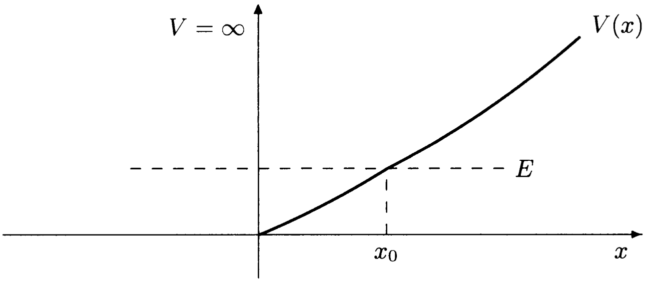

Fig. 16.5. A potential infinite at negative values of x

16.12. Let $V(x)$ be the potential shown in Fig. 16.5:

$$V(x) = \begin{cases} \infty & \text{if } x < 0\,, \\ 0 & \text{if } x = 0\,, \end{cases}$$

and $V(x)$ increases monotonically and slowly if $x > 0$.

a) Let E be a bound-state energy and let $x_0 > 0$ be the corresponding turning point. Show that for $0 < x \ll x_0$ the WKB wave function is correctly given by

$$\psi(x) = \frac{c}{\sqrt{k(x)}} \sin\left\{\int_0^x \mathrm{d}x' k(x')\right\}.$$

b) Show that for such a potential the quantization rule is given by

$$\int_0^{x_0(E)} \mathrm{d}x \,\{2m\,[E - V(x)]\}^{1/2} = \left(n - \frac{1}{4}\right)\pi\hbar, \qquad n = 1, 2, 3, \ldots$$

16.13. Let $V(x)$ be a potential given by

$$V(x) = \begin{cases} \infty & \text{if } x < 0\,, \\ V_0 x/a & \text{if } x > 0\,, \end{cases}$$

where $V_0 > 0$ and $a > 0$. Use the result of Exercise (16.12) to find, in the WKB approximation, the bound-state energies.

16.14. Show that integral (16.100) transforms operators I, Q and P into 1, q and p, respectively, and that it transforms a Hermitian operator into a real function. The proof may be restricted to one collective coordinate.

17 Atomic Orbitals

The central-field model and Hartree's self-consistent equations provide a first approximation of atomic orbitals and corresponding energies. In that context atomic wave functions are taken as products of one-electron wave functions. But this representation is not really adequate. The half-integral spin and the identity of all electrons bring important constraints on atomic wave functions: they must be completely antisymmetric with respect to permutation of the coordinates of any two electrons. This chapter is devoted mainly to the Hartree–Fock method, which determines atomic orbitals by enforcing the antisymmetry of the total wave function. The Hartree–Fock equations will be derived from the variational method, which itself has a large number of applications.[1]

17.1 The Variational Method

Consider a quantum system whose Hamiltonian H does not depend on time. Let $|\phi\rangle$ be a normalized vector in the state space. We will examine the functional $\langle\phi|H|\phi\rangle$. More specifically, we will show that its stationarity under variations is equivalent to the eigenvalue equation of H, and that it provides an upper bound on the lowest eigenvalue of H.

The following theorem makes the first of these statements more precise.

Theorem The functional $\langle\phi|H|\phi\rangle$ is stationary under variations of $|\phi\rangle$, with the constraint $\langle\phi|\phi\rangle = 1$, if and only if $|\phi\rangle$ is an eigenvector of H.

Proof Following Lagrange's method of undetermined multipliers, stationarity of $\langle\phi|H|\phi\rangle$ with the constraint $\langle\phi|\phi\rangle - 1 = 0$ is equivalent to stationarity, without constraint, of the functional

$$F\{\lambda, |\phi\rangle\} = \langle\phi|H|\phi\rangle - \lambda\{\langle\phi|\phi\rangle - 1\}.$$

Here λ is a real variable. Obviously stationarity of F with respect to λ yields the constraint. Stationarity of F with respect to variations of $|\phi\rangle$ entails that

---

[1] The Hartree–Fock method and the determination of atomic orbitals are discussed in a number of books, for instance [25], [59], [86], [214] and [233].

L. Marchildon, *Quantum Mechanics*

$$\begin{aligned} 0 &= F\{\lambda, |\phi\rangle + |\delta\phi\rangle\} - F\{\lambda, |\phi\rangle\} \\ &= \langle\delta\phi|H|\phi\rangle + \langle\phi|H|\delta\phi\rangle - \lambda\langle\delta\phi|\phi\rangle - \lambda\langle\phi|\delta\phi\rangle \\ &= 2\,\mathrm{Re}\left\{\langle\delta\phi|\left[H|\phi\rangle - \lambda|\phi\rangle\right]\right\}, \end{aligned}$$

where only first-order terms in $|\delta\phi\rangle$ have been kept. This equation must hold for any infinitesimal $\langle\delta\phi|$. Since the phase of $\langle\delta\phi|$ is arbitrary one has

$$0 = \langle\delta\phi|\left[H|\phi\rangle - \lambda|\phi\rangle\right].$$

This is true for every $\langle\delta\phi|$ if and only if

$$H|\phi\rangle = \lambda|\phi\rangle. \qquad ♣$$

Note that this search for stationary values can be carried out in the discrete spectrum only, since it involves the normalizability of $|\phi\rangle$. The result just obtained will be used in Sects. 17.4 and 17.5.

Let us now prove the second property of the functional $\langle\phi|H|\phi\rangle$.

Theorem Let $|\phi\rangle$ be a normalized vector and let E_0 be the lowest eigenvalue of H. Then $E_0 \leq \langle\phi|H|\phi\rangle$.

Proof From the spectral decomposition of H one has

$$\begin{aligned} \langle\phi|H|\phi\rangle &= \langle\phi|\left\{\mathsf{S}_E E|E\rangle\langle E|\right\}|\phi\rangle = \mathsf{S}_E E\,|\langle\phi|E\rangle|^2 \\ &\geq E_0\,\mathsf{S}_E\,|\langle\phi|E\rangle|^2 = E_0\langle\phi|\left\{\mathsf{S}_E|E\rangle\langle E|\right\}|\phi\rangle \\ &\geq E_0\langle\phi|\phi\rangle = E_0\,. \end{aligned} \qquad ♣$$

This result leads to a very precise way to determine the ground-state energy of a quantum system, called the *variational method*. It consists in computing the functional $\langle\phi|H|\phi\rangle$ for a number of normalized vectors $|\phi\rangle$ and to approximate E_0 by the lowest of these values. Obviously the choice of vectors $|\phi\rangle$ must be made systematically. As a matter of fact one considers a set of normalized vectors $|\phi(\alpha_1, \ldots, \alpha_n)\rangle$ labeled by real variables α_i. One then writes

$$F(\alpha_1, \ldots, \alpha_n) = \langle\phi(\alpha_1, \ldots, \alpha_n)|H|\phi(\alpha_1, \ldots, \alpha_n)\rangle. \tag{17.1}$$

For all α_i the value of $F(\alpha_1, \ldots, \alpha_n)$ is an upper bound on E_0. This, in particular, is true for values $\alpha_i = \bar{\alpha}_i$ that minimize F, i.e. for which[2]

$$\left\{\frac{\partial}{\partial\alpha_j} F(\alpha_1, \ldots, \alpha_n)\right\}_{\alpha_i = \bar{\alpha}_i} = 0. \tag{17.2}$$

---

[2] For a genuine minimum, all eigenvalues of the matrix of second derivatives of F evaluated at $\alpha_i = \bar{\alpha}_i$ must be positive. We also assume that the $\bar{\alpha}_i$ correspond to the absolute minimum of F.

The expression $F(\bar{\alpha}_1, \ldots, \bar{\alpha}_n)$ is the smallest upper bound on E_0 that can be obtained from such vectors as $|\phi(\alpha_1, \ldots, \alpha_n)\rangle$. We can then make the approximation

$$E_0 \approx F(\bar{\alpha}_1, \ldots, \bar{\alpha}_n). \tag{17.3}$$

This approximation is so much better as $|\phi(\bar{\alpha}_1, \ldots, \bar{\alpha}_n)\rangle$ is closer to $|\phi_0\rangle$, the eigenvector of H corresponding to E_0. The accuracy of the approximation finally depends on the choice of the set of vectors $|\phi(\alpha_1, \ldots, \alpha_n)\rangle$. Physical considerations will most often narrow down this choice.

In some cases the variational method can be used to determine energies of the first few excited states. Suppose for example that $|\phi_0\rangle$ is known exactly. Let $|\phi\rangle$ be a normalized vector orthogonal to $|\phi_0\rangle$. Clearly then $\langle\phi|H|\phi\rangle$ is an upper bound on the first excited state's energy E_1. That energy can thus be estimated by the above method. Note that it is not always necessary to know $|\phi_0\rangle$ exactly in order to find an orthogonal $|\phi\rangle$. Suppose for instance that there is a Hermitian operator K which commutes with H and is such that $K|\phi_0\rangle = \eta_0|\phi_0\rangle$. Then any eigenvector $|\phi\rangle$ of K, with eigenvalue different from η_0, is orthogonal to $|\phi_0\rangle$. Since operators H and K can be diagonalized simultaneously, eigenvectors of H can be picked among eigenvectors of K. Hence $\langle\phi|H|\phi\rangle$ will be an upper bound on the lowest energy for which the eigenvalue of K is different from η_0.

17.2 The Ground-State Energy of the Helium Atom

We now use the variational method to find the ground-state energy of the helium atom.

In Chap. 9 we obtained the Hamiltonian of an atom in terms of relative coordinates. For neutral helium (9.18) can be written as

$$H = -\frac{\hbar^2}{2m}\left(\nabla_1^2 + \nabla_2^2\right) - 2e^2\left(\frac{1}{r_1} + \frac{1}{r_2}\right) + \frac{e^2}{r_{12}}. \tag{17.4}$$

Here m is the electron's reduced mass and e is given in (9.4); r_1 and r_2 stand for the distance between each electron and the nucleus and r_{12} is the distance between the two electrons; ∇_1^2 and ∇_2^2 are Laplacian operators relative to each electron's coordinates.

The helium atom wave function depends on the coordinates of the two electrons. Should the term e^2/r_{12} be absent from the Hamiltonian, the problem of the helium atom would separate in two hydrogen-like ion problems. Clearly then the ground-state wave function would be given by a product of hydrogen-like functions, each one corresponding to the lowest energy. Thus

$$\psi(\boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{1}{\pi}\left(\frac{Z}{a_0}\right)^3 \exp\left\{-\frac{Z}{a_0}(r_1 + r_2)\right\}, \tag{17.5}$$

with $Z = 2$ and $a_0 = \hbar^2/me^2$. Note that function $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2)$ is normalized for any value of Z. Replacing Z by the real variable α we will write

$$\phi(\alpha; \boldsymbol{r}_1, \boldsymbol{r}_2) = \frac{1}{\pi} \left(\frac{\alpha}{a_0}\right)^3 \exp\left\{-\frac{\alpha}{a_0}(r_1 + r_2)\right\}, \tag{17.6}$$

and we will apply the variational method to the set $\{\phi(\alpha; \boldsymbol{r}_1, \boldsymbol{r}_2)\}$.

Following the procedure described in Sect. 17.1 we must first compute the function $F(\alpha) = \langle\phi(\alpha)|H|\phi(\alpha)\rangle$. From (17.4) and (17.6) for H and for $\phi(\alpha; \boldsymbol{r}_1, \boldsymbol{r}_2) \equiv \langle \boldsymbol{r}_1, \boldsymbol{r}_2|\phi(\alpha)\rangle$ one can see that

$$F(\alpha) = -\frac{e^2}{2a_0}\{2F_1(\alpha) + 2F_2(\alpha) + F_3(\alpha)\}, \tag{17.7}$$

where

$$F_1(\alpha) = \frac{1}{\pi}\left(\frac{\alpha}{a_0}\right)^3 \int \mathrm{d}\boldsymbol{r}_1\, \mathrm{e}^{-\alpha r_1/a_0} a_0^2\, \nabla_1^2 \mathrm{e}^{-\alpha r_1/a_0}, \tag{17.8}$$

$$F_2(\alpha) = \frac{4}{\pi}\left(\frac{\alpha}{a_0}\right)^3 \int \mathrm{d}\boldsymbol{r}_1\, \mathrm{e}^{-\alpha r_1/a_0} \frac{a_0}{r_1} \mathrm{e}^{-\alpha r_1/a_0}, \tag{17.9}$$

$$F_3(\alpha) = -\frac{2}{\pi^2}\left(\frac{\alpha}{a_0}\right)^6 \int \mathrm{d}\boldsymbol{r}_1\, \mathrm{d}\boldsymbol{r}_2\, \mathrm{e}^{-\alpha(r_1+r_2)/a_0} \frac{a_0}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} \mathrm{e}^{-\alpha(r_1+r_2)/a_0}. \tag{17.10}$$

The first two integrals are easily calculated and one gets (Exercise 17.5)

$$F_1(\alpha) = -\alpha^2, \tag{17.11}$$

$$F_2(\alpha) = 4\alpha. \tag{17.12}$$

To compute $F_3(\alpha)$ we use the expansion of $|\boldsymbol{r}_1 - \boldsymbol{r}_2|^{-1}$ in a series of spherical harmonics, given in (7.173) and (7.174):

$$\frac{1}{|\boldsymbol{r}_1 - \boldsymbol{r}_2|} = \frac{1}{r_>}\sum_{l=0}^{\infty}\left(\frac{r_<}{r_>}\right)^l \frac{4\pi}{2l+1}\sum_{m=-l}^{l} Y_{lm}^*(\theta_1, \phi_1) Y_{lm}(\theta_2, \phi_2). \tag{17.13}$$

Here $r_<$ and $r_>$ are, respectively, the smaller and the larger of r_1 and r_2, θ_1 and ϕ_1 are the polar and azimuthal angles of $\boldsymbol{r}_1$, and similarly with θ_2 and ϕ_2. Substituting the last expression into that of $F_3(\alpha)$, one sees through the spherical harmonics' orthogonality relations that only the term with $l = 0$ and $m = 0$ survives the angular integration. One thus finds

$$F_3(\alpha) = -32a_0\left(\frac{\alpha}{a_0}\right)^6 \int_0^\infty \mathrm{d}r_1\, r_1^2 \int_0^\infty \mathrm{d}r_2\, r_2^2 \mathrm{e}^{-2\alpha(r_1+r_2)/a_0} \frac{1}{r_>}.$$

The double integral's domain can be separated in two regions where $r_1 > r_2$ and $r_1 < r_2$. The symmetry of the integrand shows that both regions contribute equally to the integral, so that one can write

$$F_3(\alpha) = -32a_0 \left(\frac{\alpha}{a_0}\right)^6 2\int_0^\infty \mathrm{d}r_1\, r_1^2 \int_0^{r_1} \mathrm{d}r_2\, r_2^2 \mathrm{e}^{-2\alpha(r_1+r_2)/a_0} \frac{1}{r_1} = -\frac{5}{4}\alpha. \tag{17.14}$$

Substituting values of $F_1(\alpha)$, $F_2(\alpha)$ and $F_3(\alpha)$ in (17.7), one finds that $F(\alpha)$ is equal to

$$F(\alpha) = -\frac{e^2}{2a_0}\left\{-2\alpha^2 + \frac{27}{4}\alpha\right\}. \tag{17.15}$$

This function has its minimum at $\bar{\alpha} = 27/16$. Therefore the variational approximation for the helium atom's ground-state energy is given by[3]

$$E_0 = F(\bar{\alpha}) = -\frac{e^2}{2a_0} 2\left(\frac{27}{16}\right)^2 = -5.695\,\mathrm{Ry}. \tag{17.16}$$

The result obtained for $\bar{\alpha}$ means that if the helium atom's wave function is represented as a product of hydrogen-like functions, then the best choice consists in letting $Z = 27/16$. Much as in the central-field model, each electron hides part of the nuclear charge from the other electron.

The helium atom's ground-state energy is equal to minus the energy necessary to remove both electrons from the atom in its ground state. The experimental value is given by -5.807 Ry. The relative difference between the experimental value and the variational calculation is about 0.02. Note that the variational calculation has indeed yielded an upper bound on E_0.

The trial function we used depends on one parameter only. It is natural to believe that the approximation for E_0 will improve if we use a more complicated trial function. To choose trial functions properly, it is useful to inquire why (17.6) does not represent the true wave function of the ground state of helium. In helium the electrons' electrostatic repulsion entails that their positions are correlated. Expression (17.6), which is a product of one-electron wave functions, does not take this correlation into account.

A few years after the advent of quantum mechanics, E. A. Hylleraas proposed a set of trial functions that depend on many parameters and take the correlation of electronic positions into account. The method is described in [202], Chap. 16, where a computer program can be found for numerical calculations. Variational calculations of the first few energies of helium, corrected to take terms omitted in Hamiltonian (17.4) into account, are among the most accurate in quantum mechanics [72]. The relative difference between theoretical calculations and experimental values is presently lower than one part in 10^7 [147], [219].

---

[3] With the number of digits shown $e^2/2a_0$ is equal to 1 Ry. A more precise calculation should take into account the difference between a_0 and the Bohr radius, defined in (9.21).

17.3 Antisymmetric Wave Functions

All electrons in an atom, and in the universe, are identical. This statement has little consequences in the framework of classical mechanics. But in quantum mechanics it has major import.

Consider two electrons whose (position and spin) coordinates will be denoted by (1) and (2). Let ϕ_a and ϕ_b be two electronic wave functions. We have already introduced combined wave functions, like $\phi_a(1)\phi_b(2)$, meant to represent a state where the first electron is in state ϕ_a and the second one in ϕ_b. Such functions were used, for instance, in connection with the central-field model and Hartree's equations.

Obviously any result obtained from the function $\phi_a(1)\phi_b(2)$ could have been obtained equally well from the function $\phi_b(1)\phi_a(2)$. If electrons are identical there should be no physical difference between the state represented by $\phi_a(1)\phi_b(2)$ and the one represented by $\phi_b(1)\phi_a(2)$. But then what is the true wave function representing a state wherein one electron is in ϕ_a and the other one in ϕ_b? In fact, it is neither $\phi_a(1)\phi_b(2)$, nor $\phi_b(1)\phi_a(2)$.

The fundamental principle that determines the wave function of a set of electrons is the following: the wave function must be completely antisymmetric with respect to permutation of any two electrons' coordinates. In the particular case where there are only two electrons, one in state ϕ_a and the other one in ϕ_b, this implies that the total wave function is proportional to

$$\phi_a(1)\phi_b(2) - \phi_b(1)\phi_a(2)\,. \tag{17.17}$$

Up to a constant factor, this is the only linear combination of $\phi_a(1)\phi_b(2)$ and $\phi_b(1)\phi_a(2)$ that switches sign under permutation of coordinates (1) et (2).

The origin of the antisymmetry of many-electron wave functions lies in quantum field theory. It is a particular case of a general theorem proved by W. Pauli in 1940, that establishes a relation between a particle's spin and the behavior of wave functions under permutation of two identical particles. Pauli's result is the following: wave functions of half-integral-spin identical particles are antisymmetric under any permutation of the coordinates of two particles, whereas wave functions of integral-spin identical particles are symmetric under such permutations.[4] Note that the so-called Pauli exclusion principle is a consequence of the principle of antisymmetry of electronic wave functions. Indeed the wave function representing a situation where two electrons are in the same state ϕ_a can only be

$$\phi_a(1)\phi_a(2) - \phi_a(1)\phi_a(2)\,, \tag{17.18}$$

and we see that it vanishes identically. Hence two electrons cannot occupy the same quantum state.

---

[4] The theorem is proved in books on quantum field theory, in particular [30], Sect. 16.12, [122], p. 149 and [220], Sect. 4.4.

At this stage it is convenient to write down electronic wave functions more explicitly. They are (tensor) products of spatial and spin wave functions, that is,

$$\phi_a(1) = u_a(\boldsymbol{r}_1)\chi_a(1) . \tag{17.19}$$

A complete set of electronic wave functions can be obtained from a complete set of spatial wave functions and a complete set of spin vectors. For spin 1/2 the latter are two in number. They can be denoted by χ_+ and χ_-, these vectors representing states where the spin component along a given axis is $\hbar/2$ and $-\hbar/2$ respectively.

The two-electron wave function (17.17) can be represented as

$$\phi_a(1)\phi_b(2)-\phi_b(1)\phi_a(2) = \begin{vmatrix} \phi_a(1) & \phi_a(2) \\ \phi_b(1) & \phi_b(2) \end{vmatrix} . \tag{17.20}$$

We will see that this notation in terms of a determinant is especially useful in the case of many electrons.

From two spatial wave functions, four two-electron wave functions can be made that are antisymmetric and linearly independent. They are, for example,

$$\Phi_1 = \begin{vmatrix} u_a(\boldsymbol{r}_1)\chi_+(1) & u_a(\boldsymbol{r}_2)\chi_+(2) \\ u_b(\boldsymbol{r}_1)\chi_+(1) & u_b(\boldsymbol{r}_2)\chi_+(2) \end{vmatrix} , \tag{17.21}$$

$$\Phi_2 = \begin{vmatrix} u_a(\boldsymbol{r}_1)\chi_+(1) & u_a(\boldsymbol{r}_2)\chi_+(2) \\ u_b(\boldsymbol{r}_1)\chi_-(1) & u_b(\boldsymbol{r}_2)\chi_-(2) \end{vmatrix} , \tag{17.22}$$

$$\Phi_3 = \begin{vmatrix} u_a(\boldsymbol{r}_1)\chi_-(1) & u_a(\boldsymbol{r}_2)\chi_-(2) \\ u_b(\boldsymbol{r}_1)\chi_+(1) & u_b(\boldsymbol{r}_2)\chi_+(2) \end{vmatrix} , \tag{17.23}$$

$$\Phi_4 = \begin{vmatrix} u_a(\boldsymbol{r}_1)\chi_-(1) & u_a(\boldsymbol{r}_2)\chi_-(2) \\ u_b(\boldsymbol{r}_1)\chi_-(1) & u_b(\boldsymbol{r}_2)\chi_-(2) \end{vmatrix} . \tag{17.24}$$

Obviously, any linear combination of antisymmetric functions is itself antisymmetric. Take for instance the following four linear combinations:

$$\Phi_1 = \{u_a(\boldsymbol{r}_1)u_b(\boldsymbol{r}_2) - u_b(\boldsymbol{r}_1)u_a(\boldsymbol{r}_2)\}\, \chi_+(1)\chi_+(2) , \tag{17.25}$$

$$\begin{aligned} \frac{1}{\sqrt{2}}(\Phi_2 + \Phi_3) &= \{u_a(\boldsymbol{r}_1)u_b(\boldsymbol{r}_2) - u_b(\boldsymbol{r}_1)u_a(\boldsymbol{r}_2)\} \\ &\quad \times \frac{1}{\sqrt{2}} \{\chi_+(1)\chi_-(2) + \chi_-(1)\chi_+(2)\} , \end{aligned} \tag{17.26}$$

$$\Phi_4 = \{u_a(\boldsymbol{r}_1)u_b(\boldsymbol{r}_2) - u_b(\boldsymbol{r}_1)u_a(\boldsymbol{r}_2)\}\, \chi_-(1)\chi_-(2) , \tag{17.27}$$

$$\begin{aligned} \frac{1}{\sqrt{2}}(\Phi_2 - \Phi_3) &= \{u_a(\boldsymbol{r}_1)u_b(\boldsymbol{r}_2) + u_b(\boldsymbol{r}_1)u_a(\boldsymbol{r}_2)\} \\ &\quad \times \frac{1}{\sqrt{2}} \{\chi_+(1)\chi_-(2) - \chi_-(1)\chi_+(2)\} . \end{aligned} \tag{17.28}$$

Expressions (17.25)–(17.28) share the property of being the product of a two-electron spatial wave function and a spin vector. The first three have the same spatial wave function, which is antisymmetric with respect to permutation of $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. It is not difficult to see that their spin vectors correspond to the $+1$, 0 and -1 components of a total spin $S = 1$. These three states with the same spatial wave function and total spin 1 are collectively called *triplet.* The fourth expression's spatial wave function is symmetric with respect to permutation of $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. Its total spin vanishes and the corresponding state is called *singlet.*

Inasmuch as the helium atom's Hamiltonian is given by (17.4), it does not involve electronic spin. Hence H commutes with the total spin operator. This means that eigenfunctions of H are much better represented by expressions (17.25)–(17.28) than by expressions (17.21)–(17.24).

As for the Hamiltonian's eigenvalues, a qualitative argument shows that the triplet's energy is in general lower than the singlet's. In the singlet state the spatial wave function is symmetric. Thus there is a significant probability that the two electrons are close to one another. Their repulsive force then corresponds to a large positive potential energy. The triplet's wave function, on the other hand, is spatially antisymmetric. Electrons are seldom close to one another and their repulsive force contributes less to the positive potential energy. For a given choice of electronic orbitals, therefore, spin 1 states have less energy than the spin 0 state. Due to the wave functions' antisymmetry, the total energy of an atom depends on the spin even though no spin term is included in the Hamiltonian.

Whenever the spatial wave functions u_a and u_b coincide, expressions (17.25)–(17.27) vanish identically. Only (17.28) remains. To a good approximation, the wave function of the ground state of helium is made of two $1s$ wave functions. Hence the ground state is necessarily a singlet. The spatial part of the wave function is, approximately, a product like $u_a(\boldsymbol{r}_1)u_a(\boldsymbol{r}_2)$, which justifies the starting point of the calculation in Sect. 17.2.

It is not difficult to construct a completely antisymmetric wave function for the general case of a collection of Q electrons. Let (1), (2), …, (Q) be the electrons' coordinates and let ϕ_1, ϕ_2, …, ϕ_Q be the one-electron wave functions. Let

$$\Phi(1,2,\ldots,Q) = \frac{1}{\sqrt{Q!}} \begin{vmatrix} \phi_1(1) & \phi_1(2) & \ldots & \phi_1(Q) \\ \phi_2(1) & \phi_2(2) & \ldots & \phi_2(Q) \\ \ldots & \ldots & \ldots & \ldots \\ \phi_Q(1) & \phi_Q(2) & \ldots & \phi_Q(Q) \end{vmatrix} . \tag{17.29}$$

Properties of determinants entail that $\Phi(1,2,\ldots,Q)$ is antisymmetric with respect to permutation of any two coordinates. Such determinants are called *Slater determinants*, after J. C. Slater who introduced them. The factor $(Q!)^{-1/2}$ ensures normalization, as we will soon see.

We now compute for later use matrix elements of certain operators $O(1,2,\ldots,Q)$ in a basis of Slater determinants. Specifically, we assume that

$O(1, 2, \dots, Q)$ is symmetric with respect to any permutation of two of its coordinates. Moreover, we assume that one-electron wave functions are all built from a complete set of orthonormal spatial wave functions u_i and two spin vectors χ_+ and χ_-. These objects are put together so that the $\phi_i(1) = u_i(\boldsymbol{r}_1)\chi_i(1)$ are orthonormal, that is,

$$\langle\phi_i|\phi_j\rangle = \int \mathrm{d}\boldsymbol{r}_1\, u_i^*(\boldsymbol{r}_1)u_j(\boldsymbol{r}_1){\chi_i}^\dagger(1)\chi_j(1) = \delta_{ij} . \tag{17.30}$$

Due to the symmetry of operator O and the antisymmetry of the wave function Φ one can see that

$$\begin{aligned}
&\langle\Phi|O|\Phi'\rangle \\
&= \frac{1}{Q!}\int \mathrm{d}\boldsymbol{r}_1 \dots \mathrm{d}\boldsymbol{r}_Q \begin{vmatrix} \phi_1^*(1) \dots \phi_1^*(Q) \\ \dots \quad \dots \quad \dots \\ \phi_Q^*(1) \dots \phi_Q^*(Q) \end{vmatrix} O(1, \dots, Q) \begin{vmatrix} \phi_1'(1) \dots \phi_1'(Q) \\ \dots \quad \dots \quad \dots \\ \phi_Q'(1) \dots \phi_Q'(Q) \end{vmatrix} \\
&= \int \mathrm{d}\boldsymbol{r}_1 \dots \mathrm{d}\boldsymbol{r}_Q \left\{\phi_1^*(1) \dots \phi_Q^*(Q)\right\} O(1, \dots, Q) \begin{vmatrix} \phi_1'(1) \dots \phi_1'(Q) \\ \dots \quad \dots \quad \dots \\ \phi_Q'(1) \dots \phi_Q'(Q) \end{vmatrix} .
\end{aligned} \tag{17.31}$$

By $\phi_i^*(j)$ is meant $u_i^*(\boldsymbol{r}_j){\chi_i}^\dagger(j)$. Note that for simplicity, two identical wave functions in the two determinants will always be labeled by the same index, which means we will never have $\phi_i = \phi_j'$ for $i \neq j$.

Let us now examine three specific cases for the operator $O(1, 2, \dots, Q)$.

i) $O(1, 2, \dots, Q) = CI$, where C is a constant and I is the identity. Clearly the matrix element vanishes unless all ϕ_i coincide with all ϕ_i'. Explicitly

$$\langle\Phi|C|\Phi'\rangle = \begin{cases} C & \text{if } \phi_i = \phi_i' \text{ for all } i \,, \\ 0 & \text{otherwise} \,. \end{cases} \tag{17.32}$$

Letting $C = 1$, one concludes that $|\Phi\rangle$ is normalized.

ii) $O(1, 2, \dots, Q) = \sum_i F(i)$.
$F(i)$ is an operator that involves the coordinates of only one electron. Here the matrix element is different from zero not only if all ϕ_i coincide with the ϕ_i', but also if just one ϕ_j differs from the corresponding ϕ_j'. One thus finds[5]

$$\langle\Phi|\sum_{i=1}^{Q} F(i)|\Phi'\rangle = \begin{cases} \sum_{i=1}^{Q}\langle\phi_i|F|\phi_i\rangle & \text{if } \phi_i = \phi_i' \text{ for all } i \,, \\ \langle\phi_j|F|\phi_j'\rangle & \text{if } \phi_j \neq \phi_j' \text{ for one } j \,, \\ 0 & \text{otherwise} \,. \end{cases} \tag{17.33}$$

---

[5] It is important to clearly understand the meaning of the notations used. In $\sum_i F(i)$ the summation is carried out on electronic coordinates. In $\sum_i\langle\phi_i|F|\phi_i\rangle$, on the other hand, it is carried out on electronic wave functions.

iii) $O(1,2,\dots,Q)=\sum_{i<j}G(i,j)$.
$G(i,j)$ is an operator that involves the coordinates of two electrons symmetrically. The matrix element is different from zero if all ϕ_i coincide with the ϕ_i', with at most two exceptions. One has[6]

$$\langle\Phi|\sum_{i<j}^{Q}G(i,j)|\Phi'\rangle$$

$$=\begin{cases}\sum_{i<j}^{Q}\{\langle\phi_i\phi_j|G|\phi_i\phi_j\rangle-\langle\phi_i\phi_j|G|\phi_j\phi_i\rangle\} & \text{if } \phi_i=\phi_i' \text{ for all } i\,,\\ \sum_{i=1,i\neq k}^{Q}\{\langle\phi_i\phi_k|G|\phi_i\phi_k'\rangle-\langle\phi_i\phi_k|G|\phi_k'\phi_i\rangle\} & \text{if } \phi_k\neq\phi_k' \text{ for one } k\,,\\ \langle\phi_k\phi_l|G|\phi_k'\phi_l'\rangle-\langle\phi_k\phi_l|G|\phi_l'\phi_k'\rangle & \text{if } \phi_k\neq\phi_k' \text{ and } \phi_l\neq\phi_l' \\ & \text{for one } k \text{ and one } l\,,\\ 0 & \text{otherwise}\,.\end{cases}\tag{17.34}$$

In closing this section we note that, for Q electrons as well as in the particular case where $Q=2$, the atomic Hamiltonian's eigenfunctions are in general linear combinations of Slater determinants. To a good approximation these linear combinations correspond to a well-defined value of the total spin. In general (here the case where $Q=2$ is an exception), these eigenfunctions cannot be written as (tensor) products of spatial wave functions and spin vectors, but as linear combinations of such products. Nevertheless the spatial part of wave functions is usually more antisymmetric if the total spin is large. This corresponds to lower atomic energies. These remarks justify what is called *Hund's rule*: in the ground state degenerate orbitals tend to be occupied by one electron at a time, with all spins parallel.[7]

17.4 The Hartree–Fock Equations

The Hamiltonian of an atom of Q electrons and Z protons is given in terms of relative coordinates by (9.18), which can be written as

$$H=\sum_{i=1}^{Q}\left\{-\frac{\hbar^2}{2m}\nabla_i^2-\frac{Ze^2}{r_i}\right\}+\sum_{i<j}^{Q}\frac{e^2}{r_{ij}}\,.\tag{17.35}$$

Here m is the reduced mass of each electron,[8] r_i the distance between electron i and the nucleus and r_{ij} the distance between electron i and electron j. With the notations of Sect. 17.3, (17.35) can be written as

---

[6] In $\langle\phi_k\phi_l|G|\phi_k'\phi_l'\rangle$ the functions ϕ_k and ϕ_k' are associated with the first argument of G, while ϕ_l and ϕ_l' are associated with the second one.

[7] For a general formulation and an assessment of Hund's rule see [59], Sect. 4.16.

[8] In many-electron atoms we will no longer distinguish reduced mass from real mass. The symbol m may thus stand for the electron's real mass.

$$H = \sum_{i=1}^{Q} F(i) + \sum_{i<j}^{Q} G(i,j), \tag{17.36}$$

where $F(i)$ and $G(i,j)$ are given by

$$F(i) = -\frac{\hbar^2}{2m}\nabla_i^2 - \frac{Ze^2}{r_i}, \qquad G(i,j) = \frac{e^2}{r_{ij}}. \tag{17.37}$$

$F(i)$ and $G(i,j)$ are operators that involve the coordinates of one and two electrons, respectively.

Let $\Phi(1,2,\ldots,Q)$ be a determinantal function:

$$\Phi(1,2,\ldots,Q) = \frac{1}{\sqrt{Q!}}\begin{vmatrix} \phi_1(1) & \ldots & \phi_1(Q) \\ \ldots & \ldots & \ldots \\ \phi_Q(1) & \ldots & \phi_Q(Q) \end{vmatrix}. \tag{17.38}$$

Here $\phi_i(j)$ is the product of a spatial function and a spin vector, that is,

$$\phi_i(j) = u_i(\boldsymbol{r}_j)\chi_i(j). \tag{17.39}$$

We assume that the u_i are picked among a set of orthonormal functions and that χ_i is equal to χ_+ or χ_-. In this way the ϕ_i are orthonormal and $\Phi(1,2,\ldots,Q)$ is normalized.

According to the antisymmetrization postulate, physically acceptable eigenfunctions of H are antisymmetric with respect to any permutation of the coordinates of two electrons. We will assume that determinantal functions can be a good approximation of the Hamiltonian's eigenfunctions.[9] Following the theorem on p. 391, exact eigenfunctions make the functional $\langle\phi|H|\phi\rangle$ stationary. Hence it is reasonable to believe that the best approximation of an eigenfunction by means of a determinantal function is that which makes $\langle\Phi|H|\Phi\rangle$ stationary with respect to variations of spatial functions u_i.

Let us first compute $\langle\Phi|H|\Phi\rangle$, where H is given by (17.36) and Φ by (17.38). Making use of (17.33) and (17.34) one finds that

$$\begin{aligned}\langle\Phi|H|\Phi\rangle &= \sum_{i=1}^{Q}\langle\phi_i|F(i)|\phi_i\rangle \\ &+ \sum_{i<j}^{Q}\left\{\langle\phi_i\phi_j|G(i,j)|\phi_i\phi_j\rangle - \langle\phi_i\phi_j|G(i,j)|\phi_j\phi_i\rangle\right\}.\end{aligned} \tag{17.40}$$

---

[9] As we pointed out for helium (p. 398), atomic wave functions are usually better represented by linear combinations of Slater determinants than by one determinant only. However there is always, among quantum states associated with a given electronic configuration, at least one state which is well represented by one determinant only. This is all we need for the determination of orbitals.

In terms of spatial functions and spin vectors one gets

$$\langle \Phi | H | \Phi \rangle = \sum_{i=1}^{Q} \int \mathrm{d}\boldsymbol{r}_1 \, u_i^*(\boldsymbol{r}_1) F(1) u_i(\boldsymbol{r}_1) + \sum_{i<j}^{Q} \int \mathrm{d}\boldsymbol{r}_1 \, \mathrm{d}\boldsymbol{r}_2 \\ \times u_i^*(\boldsymbol{r}_1) u_j^*(\boldsymbol{r}_2) G(1,2) \left\{ u_i(\boldsymbol{r}_1) u_j(\boldsymbol{r}_2) - \delta(i,j) u_j(\boldsymbol{r}_1) u_i(\boldsymbol{r}_2) \right\} , \quad (17.41)$$

where $\delta(i,j) = {\chi_i}^\dagger \chi_j$ is equal to 1 or 0 according as χ_i and χ_j are identical or orthogonal.

Let us now analyze the stationarity. Since the u_i are picked among a set of orthonormal functions and the ϕ_i are orthonormal, one has the following constraints:

$$\int \mathrm{d}\boldsymbol{r}_1 \, u_i^*(\boldsymbol{r}_1) u_i(\boldsymbol{r}_1) = 1 , \quad (17.42)$$

$$\int \mathrm{d}\boldsymbol{r}_1 \, u_i^*(\boldsymbol{r}_1) u_j(\boldsymbol{r}_1) = 0 \quad \text{if } i \neq j \text{ and } \chi_i = \chi_j . \quad (17.43)$$

In other words, spatial functions are always normalized and they are orthogonal for identical spins. According to Lagrange's method of undetermined multipliers, the problem consists in finding the stationary value, without constraints, of the functional $F(u_i)$, where

$$F(u_i) = \langle \Phi | H | \Phi \rangle - \sum_{i=1}^{Q} \lambda_{ii} \left\{ \int \mathrm{d}\boldsymbol{r}_1 \, u_i^*(\boldsymbol{r}_1) u_i(\boldsymbol{r}_1) - 1 \right\} \\ - \sum_{i<j}^{Q} \left\{ \lambda_{ij} \int \mathrm{d}\boldsymbol{r}_1 \, u_i^*(\boldsymbol{r}_1) u_j(\boldsymbol{r}_1) + \lambda_{ji} \int \mathrm{d}\boldsymbol{r}_1 \, u_j^*(\boldsymbol{r}_1) u_i(\boldsymbol{r}_1) \right\} . \quad (17.44)$$

Here $\lambda_{ji} = (\lambda_{ij})^*$ and $\lambda_{ij} = 0$ if $\chi_i \neq \chi_j$. One can see that the complex multiplier λ_{ij} corresponds to two distinct constraints on the real and imaginary parts of $\int \mathrm{d}\boldsymbol{r}_1 \, u_i^*(\boldsymbol{r}_1) u_j(\boldsymbol{r}_1)$. The functional $F(u_i)$ is real.

With some algebraic manipulations and using the fact that operators $F(1)$ and $G(1,2)$ are Hermitian, one obtains (Exercise 17.7)

$$F(u_i + \delta u_i) - F(u_i) \\ = 2 \,\mathrm{Re} \sum_{i=1}^{Q} \int \mathrm{d}\boldsymbol{r}_1 \, \delta u_i^*(\boldsymbol{r}_1) \left\{ F(1) u_i(\boldsymbol{r}_1) - \sum_{j=1}^{Q} \lambda_{ij} u_j(\boldsymbol{r}_1) \right. \\ \left. + \sum_{j=1}^{Q} \int \mathrm{d}\boldsymbol{r}_2 \, u_j^*(\boldsymbol{r}_2) G(1,2) \left[u_i(\boldsymbol{r}_1) u_j(\boldsymbol{r}_2) - \delta(i,j) u_j(\boldsymbol{r}_1) u_i(\boldsymbol{r}_2) \right] \right\} .$$

This expression vanishes for any function δu_i^* if and only if

$$\begin{aligned} F(1)u_i(\boldsymbol{r}_1) &- \sum_{j=1}^{Q} \lambda_{ij} u_j(\boldsymbol{r}_1) \\ &+ \sum_{j=1}^{Q} \int \mathrm{d}\boldsymbol{r}_2 \, u_j^*(\boldsymbol{r}_2) G(1,2) \left[u_i(\boldsymbol{r}_1) u_j(\boldsymbol{r}_2) - \delta(i,j) u_j(\boldsymbol{r}_1) u_i(\boldsymbol{r}_2) \right] = 0 \,. \end{aligned} \tag{17.45}$$

These equations are essentially what we were looking for, but they can be simplified further. For this purpose let us introduce a unitary matrix V, of dimension $Q \times Q$, and such that $\det V = 1$ and $V_{ik} = 0$ if $\delta(i,k) = 0$. Let us also write

$$u_i(\boldsymbol{r}) = \sum_{k=1}^{Q} V_{ik} u'_k(\boldsymbol{r}) \,. \tag{17.46}$$

The matrix V relates spatial functions associated with the same spin vector. One can check that Slater determinants as well as constraints assume the same form in terms of functions u'_k as in terms of the u_i. The unitary matrix V can be chosen so as to diagonalize the Hermitian matrix $[\lambda_{ij}]$ (Exercise 17.8). In this case one gets from (17.45) equations for the u'_k which, once the primes are removed and substitution is made for $F(1)$ and $G(1,2)$, can be written as

$$\begin{aligned} &\left\{ -\frac{\hbar^2}{2m} \nabla_1^2 - \frac{Ze^2}{r_1} - \lambda_i \right\} u_i(\boldsymbol{r}_1) \\ &\quad + \sum_{j=1}^{Q} \int \mathrm{d}\boldsymbol{r}_2 \, u_j^*(\boldsymbol{r}_2) \frac{e^2}{r_{12}} \left[u_i(\boldsymbol{r}_1) u_j(\boldsymbol{r}_2) - \delta(i,j) u_j(\boldsymbol{r}_1) u_i(\boldsymbol{r}_2) \right] = 0 \,. \end{aligned} \tag{17.47}$$

The real numbers λ_i are eigenvalues of the matrix $[\lambda_{ij}]$. The system of Q coupled differential equations that we just found is called the *Hartree–Fock equations.*

To understand the physical meaning of the Hartree–Fock equations, it is useful to rewrite them slightly differently as

$$\begin{aligned} &\left\{ -\frac{\hbar^2}{2m} \nabla_1^2 - \frac{Ze^2}{r_1} \right\} u_i(\boldsymbol{r}_1) + \left\{ \int \mathrm{d}\boldsymbol{r}_2 \, \frac{e^2}{r_{12}} \sum_{j=1}^{Q} |u_j(\boldsymbol{r}_2)|^2 \right\} u_i(\boldsymbol{r}_1) \\ &\quad - \left\{ \frac{q_e}{4\pi\varepsilon_0} \int \mathrm{d}\boldsymbol{r}_2 \, \varrho_{\mathrm{ex}}^i(\boldsymbol{r}_2, \boldsymbol{r}_1) \frac{1}{r_{12}} \right\} u_i(\boldsymbol{r}_1) = \lambda_i u_i(\boldsymbol{r}_1) \,. \end{aligned} \tag{17.48}$$

The function $\varrho_{\mathrm{ex}}^i(\boldsymbol{r}_2, \boldsymbol{r}_1)$, called the *exchange charge density*, is given by

$$\varrho_{\mathrm{ex}}^{i}(\boldsymbol{r}_2,\boldsymbol{r}_1) = q_e \sum_{j=1}^{Q} \delta(i,j) \left\{ \frac{u_i^*(\boldsymbol{r}_1)u_j^*(\boldsymbol{r}_2)u_j(\boldsymbol{r}_1)u_i(\boldsymbol{r}_2)}{u_i^*(\boldsymbol{r}_1)u_i(\boldsymbol{r}_1)} \right\}. \tag{17.49}$$

The Hartree–Fock equations (17.48) differ from Hartree's equations (9.49) in that the sum on j includes the term $j = i$, and by the presence of the exchange term.[10]

Equation (17.48) for the function $u_i(\boldsymbol{r}_1)$ looks like the eigenvalue equation of a Hamiltonian whose potential energy is made up of three terms: (i) the energy of an electron at a distance r_1 from the nucleus; (ii) the energy of an electron in the average field of all electrons, including itself; and (iii) minus the energy of an electron in the field produced by the exchange charge density.

We will soon take a closer look at the exchange charge density. But before, let us examine the nature of the solutions of the Hartree–Fock equations. The analogy with the Hamiltonian's eigenvalue equation for a particle in a potential suggests that each λ_i has a discrete spectrum of negative values, and that there can be an infinite number of solutions for the function $u_i(\boldsymbol{r})$.

The Hartree–Fock equations are nonlinear coupled equations for the functions $u_i(\boldsymbol{r})$. A solution of the system of equations thus consists in a set of real numbers λ_j and functions $u_j(\boldsymbol{r})$, where j goes from 1 to Q.[11] As a matter of fact, the solution is always obtained self-consistently. One starts with an estimate $u_j^{(1)}(\boldsymbol{r})$ of the form of occupied spatial functions. Q electrons are then assigned to these orbitals, with well-defined spins and in accordance with Pauli's principle. Next the second and third terms in curly brackets in (17.48) are evaluated for i. Note that in general, this equation is different for each value of i. Its solution, for each i, yields a better estimate of the form of spatial functions, that can be denoted by $u_j^{(2)}(\boldsymbol{r})$. The iterative process is carried out till convergence.

To obtain the ground-state atomic configuration, one assigns electrons to spatial functions with the lowest possible energies, while always complying with Pauli's principle. Excited configurations are obtained by assigning electrons to appropriate spatial functions.

The Hartree–Fock equations were developed on the basis of orthogonality relations (17.43). The following result can be proved directly (Exercise 17.9) from (17.47).

Theorem Two solutions u_k and u_l of (17.47), associated with the same spin and corresponding to different constants λ_k and λ_l, are orthogonal.

Insofar as the atomic wave function is well represented by just one Slater determinant, the atom's total energy is given by (17.41), which represents

---

[10] Hartree's equations can be obtained by finding the stationary value of the functional $\langle\psi|H|\psi\rangle$, where ψ is given by (9.25). The ψ_k are normalized but not necessarily orthogonal.

[11] By extension, each of the functions u_i that belong to the set of Q functions satisfying (17.48) is also called a solution.

the mean value of H in $|\Phi\rangle$. One can rewrite this expression by substituting the Hartree–Fock equations (17.47) in it, thereby obtaining (Exercise 17.10):

$$\langle\Phi|H|\Phi\rangle = \sum_{i=1}^{Q}\lambda_i - \sum_{i<j}^{Q}\int \mathrm{d}\boldsymbol{r}_1\,\mathrm{d}\boldsymbol{r}_2\,u_i^*(\boldsymbol{r}_1)u_j^*(\boldsymbol{r}_2) \times G(1,2)\left\{u_i(\boldsymbol{r}_1)u_j(\boldsymbol{r}_2) - \delta(i,j)u_j(\boldsymbol{r}_1)u_i(\boldsymbol{r}_2)\right\}. \tag{17.50}$$

Clearly the total energy is not simply equal to the sum of the λ_i. These have important physical meaning. Indeed $-\lambda_i$ represents approximately the energy needed to extract electron i from the atom. This statement follows from the following theorem, due to T. C. Koopmans (Exercise 17.11).

Koopmans's Theorem Consider an atom and the ion obtained from it by extracting electron i. Suppose the atom's and the ion's wave functions are each represented by just one determinant, and that these determinants are built from the same spatial functions u_j. Then λ_i is equal to the difference between the atom's energy and the ion's energy.

The following theorem, due to L. Brillouin, is useful in configuration interaction calculations, which will be introduced in Chap. 18. It can be proved (Exercise 17.12) by evaluating the matrix element considered and making use of the Hartree–Fock equations.

Brillouin's Theorem Let Φ and Φ' be two Slater determinants built from orthogonal spatial wave functions that satisfy the Hartree–Fock equations. If Φ and Φ' differ by one spatial orbital only, then $\langle\Phi'|H|\Phi\rangle = 0$.

Let us now come back to the exchange charge density. It is, in the Hartree–Fock equations, the quantity whose interpretation is least obvious. One can check that it has the following properties ([214], Sect. 16.4):

i) The total charge $\int \mathrm{d}\boldsymbol{r}_2\,\varrho_{\mathrm{ex}}^i(\boldsymbol{r}_2,\boldsymbol{r}_1)$ associated with the exchange charge density is equal to q_e if u_i is an occupied orbital, and it vanishes otherwise.
ii) Only electrons whose spin projection is identical to that of electron i contribute to ϱ_{ex}^i.
iii) $\varrho_{\mathrm{ex}}^i(\boldsymbol{r}_1,\boldsymbol{r}_1) = q_e\sum_{j=1}^{Q}\delta(i,j)u_j^*(\boldsymbol{r}_1)u_j(\boldsymbol{r}_1)$. This means that at $\boldsymbol{r}_2 = \boldsymbol{r}_1$, the exchange charge density is equal to the charge density of all electrons that have the same spin projection as electron i.

In the search for the form of orbital $u_i(\boldsymbol{r})$ by means of (17.48), the contribution of electron i must be subtracted from the total electronic charge density. This contribution is represented by the exchange charge density. Because electrons are indistinguishable, ϱ_{ex}^i is represented by a sum on all occupied spatial functions associated with the same spin projection as that of u_i. If u_i is occupied, the total charge associated with ϱ_{ex}^i is precisely equal to the charge of an electron.

As we mentioned, solutions of the Hartree–Fock equations can be used to build determinantal wave functions. It is very hard, however, to obtain solutions of the Hartree–Fock equations in the form given above. This mainly comes from the fact that in (17.48) the exchange charge density and the total potential energy are not spherically symmetric. We will see how this problem can be solved by a relatively mild simplification of the Hartree–Fock equations.

17.5 Using the Hartree–Fock Equations

In our investigation of the central-field model we saw that the total electronic charge density (9.50) is not, in general, independent of angles. It is independent only if all electronic shells are filled. The electronic charge density was thus replaced by its average on all solid angle, so as to obtain a spherically symmetric electrostatic potential.

Similar manipulations can be made with the Hartree–Fock equations. Suppose first that all electronic shells are filled. It is then natural to look for solutions of the form[12]

$$u_i(\boldsymbol{r}) = u_{n_i l_i}(r) Y_{l_i m_i}(\theta, \phi) . \tag{17.51}$$

Let us substitute these relations in (17.47). From the properties of spherical harmonics one can show (Exercise 17.13) that the angular dependence disappears and that there only remains a system of coupled equations for radial functions $u_{n_i l_i}$.

Let us now turn to an electronic configuration with a few electrons outside closed shells. One can reasonably believe that a spherically symmetric approximation will not appreciably alter the form of solutions of (17.47). It is convenient to make the approximation on angular and spin variables at the same time. We will not look at details of the calculation, an example of which can be found in [25], p. 68. One then obtains a system of integro-differential equations for radial functions $u_{n_i l_i}(r)$, that depends neither on angular nor on spin variables. That system is called the *restricted Hartree–Fock equations.*

The restricted Hartree–Fock equations are solved self-consistently. For a given choice of occupied orbitals, the equation that determines the radial function $u_{n_i l_i}(r)$ depends neither on the magnetic quantum number m_i nor on the spin quantum number m_{s_i}. It does, however, depend on n_i and l_i. In other words, the equation for $u_{n_i l_i}$ is different from the one for $u_{n_j l_j}$.[13]

---

[12] We use the same symbol u to denote three-dimensional spatial wave functions $u_i(\boldsymbol{r})$ and radial functions $u_{n_i l_i}(r)$. The magnetic quantum number associated with l_i is simply denoted by m_i rather than by m_{l_i}.

[13] In the self-consistent calculation, the "potential" that determines $u_{n_i l_i}$ is not the same as that which determines $u_{n_j l_j}$. Hence these radial functions are no longer orthogonal if $l_i = l_j$.

The restricted Hartree–Fock equations are extremely useful and they have given rise to numerous explicit calculations. We will come back to them at the end of this section. For now we will try to simplify them further and obtain for the radial function $u_{n_i l_i}$ an equation that does not depend on quantum numbers n_i and l_i. In other words we will simplify the Hartree–Fock equations so as to bring them closer to the central-field model.

For this purpose let us go back to (17.41) which can be rewritten as

$$\begin{aligned}\langle\Phi|H|\Phi\rangle &= \sum_{i=1}^{Q}\int \mathrm{d}\boldsymbol{r}_1\, u_i^*(\boldsymbol{r}_1)F(1)u_i(\boldsymbol{r}_1)\\ &\quad+\frac{1}{2}\sum_{i,j=1}^{Q}\int \mathrm{d}\boldsymbol{r}_1\,\mathrm{d}\boldsymbol{r}_2\, u_i^*(\boldsymbol{r}_1)u_j^*(\boldsymbol{r}_2)G(1,2)u_i(\boldsymbol{r}_1)u_j(\boldsymbol{r}_2) - K\,, \end{aligned}\tag{17.52}$$

where

$$K=\frac{1}{2}\sum_{i,j=1}^{Q}\delta(i,j)\int \mathrm{d}\boldsymbol{r}_1\,\mathrm{d}\boldsymbol{r}_2\, u_i^*(\boldsymbol{r}_1)u_j^*(\boldsymbol{r}_2)\frac{e^2}{r_{12}}u_j(\boldsymbol{r}_1)u_i(\boldsymbol{r}_2). \tag{17.53}$$

The factors 1/2 are introduced as the sum on $i<j$ is transformed into a sum on i and j. Terms with $i=j$ in the two sums cancel each other.

We now evaluate the quantity K approximately, using hypotheses on which the Thomas–Fermi model is based. In this model there are exactly two electrons in each occupied spatial wave function. This allows to replace, in the expression for K, the function $\delta(i,j)$ by 1/2. Spatial functions coincide with plane waves, so that

$$\sum_{i=1}^{Q}u_i^*(\boldsymbol{r}_1)u_i(\boldsymbol{r}_2)=\frac{1}{(2L)^3}\sum_{i=1}^{Q}\exp\left\{i\boldsymbol{k}_i\cdot(\boldsymbol{r}_2-\boldsymbol{r}_1)\right\}. \tag{17.54}$$

To approximately evaluate the sum over i, we replace it by an integral over $\mathrm{d}\boldsymbol{k}$, carried out on a sphere of radius k_F. Here k_F is the maximum value of k, associated with the Fermi energy.[14] One thus finds

$$\begin{aligned}\sum_{i=1}^{Q}u_i^*(\boldsymbol{r}_1)u_i(\boldsymbol{r}_2)&=\frac{1}{(2L)^3}2\left(\frac{L}{\pi}\right)^3\int_0^{k_F}\mathrm{d}\boldsymbol{k}\,\exp\left\{i\boldsymbol{k}\cdot\boldsymbol{r}_{21}\right\}\\ &=\frac{1}{\pi^2(r_{12})^3}\left\{\sin(k_F r_{12})-(k_F r_{12})\cos(k_F r_{12})\right\}. \end{aligned}\tag{17.55}$$

Let us substitute this expression in (17.53). Before integrating it is convenient to make the following change of variables:

---

[14] Specifically, k_F is the wave number of an electron whose kinetic energy is equal to the absolute value of the potential energy at $\boldsymbol{r}$. We assume that k_F has the same value at $\boldsymbol{r}_1$ as at $\boldsymbol{r}_2$. This assumption is valid if $|\boldsymbol{r}_1-\boldsymbol{r}_2|$ is small. If, on the other hand, $|\boldsymbol{r}_1-\boldsymbol{r}_2|$ is not small, then the sum over $\boldsymbol{k}_i$ is negligible.

$$\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2\,, \qquad \boldsymbol{s} = \frac{1}{2}(\boldsymbol{r}_1 + \boldsymbol{r}_2)\,. \tag{17.56}$$

Since the Jacobian of the transformation is equal to 1 one finds that (Exercise 17.14)

$$\begin{aligned} K &= \frac{e^2}{4}\int \mathrm{d}\boldsymbol{s}\,\mathrm{d}\boldsymbol{r}\,\frac{1}{r}\left\{\frac{1}{\pi^2 r^3}\left[\sin(k_F r) - (k_F r)\cos(k_F r)\right]\right\}^2 \\ &= \frac{e^2}{4}\int \mathrm{d}\boldsymbol{s}\,k_F^4 \int \mathrm{d}\boldsymbol{\varrho}\,\frac{1}{\varrho}\left\{\frac{1}{\pi^2\varrho^3}\left[\sin\varrho - \varrho\cos\varrho\right]\right\}^2 \\ &= \frac{e^2}{4}\int \mathrm{d}\boldsymbol{s}\,k_F^4\frac{1}{\pi^3}\,. \end{aligned} \tag{17.57}$$

The quantity k_F was left inside the integral over $\mathrm{d}\boldsymbol{s}$, since the Fermi energy depends on position. In fact k_F is related to the total electronic density through (9.33) which, in the present case, reads as

$$k_F^3 = 3\pi^2 \sum_{i=1}^{Q} u_i^*(\boldsymbol{s})u_i(\boldsymbol{s})\,. \tag{17.58}$$

Thus one finds, replacing the integration variable $\boldsymbol{s}$ by $\boldsymbol{r}_1$

$$K = \frac{3e^2}{4}\left(\frac{3}{\pi}\right)^{1/3}\int \mathrm{d}\boldsymbol{r}_1 \left\{\sum_{i=1}^{Q} u_i^*(\boldsymbol{r}_1)u_i(\boldsymbol{r}_1)\right\}^{4/3}\,. \tag{17.59}$$

The Hartree–Fock equations were obtained by finding the stationary value of functional (17.41). Let us carry out the process once more, using functional (17.52) as a starting point, with K defined as in (17.59). One then finds (Exercise 17.15) the following equations:

$$\begin{aligned} &\left\{-\frac{\hbar^2}{2m}\nabla_1^2 - \frac{Ze^2}{r_1} + e^2\int \mathrm{d}\boldsymbol{r}_2\,\frac{1}{r_{12}}\sum_{j=1}^{Q}|u_j(\boldsymbol{r}_2)|^2\right. \\ &\qquad \left. - e^2\left[\frac{3}{\pi}\sum_{j=1}^{Q} u_j^*(\boldsymbol{r}_1)u_j(\boldsymbol{r}_1)\right]^{1/3}\right\} u_i(\boldsymbol{r}_1) = \lambda_i u_i(\boldsymbol{r}_1)\,. \end{aligned} \tag{17.60}$$

Obviously, the potential that determines orbital $u_i(\boldsymbol{r})$ no longer depends on i. Only the angular dependence remains to be eliminated. For this purpose the total electronic density is replaced by its average on all solid angle, as in (9.51). The potential (9.53) results with, as anticipated, the form (9.54) for the correction term V_{corr}.

The calculation just made is an approximate way to take into account the exchange term in the Hartree–Fock equations. To put things in perspective we should point out that there are other ways to compute V_{corr}, and it is not a priori obvious which one is best. Two of them will be briefly mentioned here. We refer to [59], Chap. 7, for a more elaborate discussion.

At large values of r radial functions associated with atomic orbitals go to zero exponentially. Hence the term V_{corr} just calculated also decreases exponentially. Let $V(r)$ denote the total spherically symmetric potential associated with (17.60) and given explicitly in (9.53) and (9.54). It is easy to see that for large r, $V(r)$ behaves as $-(Z - Q)e^2/r$. Intuitively, one would rather expect a behavior like $-(Z + 1 - Q)e^2/r$, since an electron very far from the nucleus sees a total charge equal to $-[Z - (Q - 1)]q_e$. This suggests changing $V(r)$ by adding a term which, for large r, reproduces the expected asymptotic form for the total potential. Specifically let $\bar{r}$ be such that $V(\bar{r}) = -(Z + 1 - Q)e^2/\bar{r}$. Then one uses $V(r)$ if $r < \bar{r}$ and replaces $V(r)$ by $-(Z + 1 - Q)e^2/r$ if $r > \bar{r}$. Energies displayed in Table 9.1 were computed with this modification.

In contrast with Bethe and Jackiv ([25], p. 71), whom our calculation of V_{corr} follows, Slater ([214], p. 343) applies hypotheses of the Thomas–Fermi model not to the functional (17.41) but to the exchange charge density (17.49). His argument leads to equations similar to (17.60), except that the last term inside the curly brackets is multiplied by a factor 3/2. It seems that the best radial functions are calculated by a factor somewhere between 1 and 3/2. Reference [108] used Slater's correction and the change in $V(r)$ at large values of r to carry out detailed calculations of atomic energies and orbitals.

No matter how hard one tries, a spherically symmetric potential can clearly never perfectly reproduce results obtained from the Hartree–Fock equations. But these equations also appear under different guises, that we now briefly point out.

In every case the Hartree–Fock equations are obtained through the stationary value of a functional. We have used the functional (17.44), in which the wave function Φ is represented by just one Slater determinant. But atomic wave functions are most often represented by linear combinations of Slater determinants, corresponding to well-defined values of angular momentum. The Hartree–Fock equations can be developed from such wave functions ([86], Chap. 2; [87]). With these wave functions $\langle\Phi|H|\Phi\rangle$ represents the energy of a specific quantum state belonging to the electronic configuration picked. One can also consider the mean value of all these energies. This mean value is a very useful functional that will be evaluated explicitly in Sect. 18.3. It is also an appropriate starting point for the derivation of Hartree–Fock equations ([59], Chap. 7).

In all actual atomic calculations the Hartree–Fock equations boil down to equations for radial functions $u_{n_i l_i}$ in (17.51). It is possible to substitute (17.51) directly in the functional we start with [in (17.44) for instance], and to write the latter in terms of radial functions. This is the method used by [86] and [59]. The former reference gives detailed numerical results for all neutral atoms up to radon.

In Chap. 20 we will see that the Hartree–Fock method can be adapted to molecules. The Hamiltonian then involves electrostatic interaction terms between each electron and each nucleus, the latter assumed at rest. Note however that any spherical symmetry is lost, so that spatial orbitals are no longer products of radial functions and spherical harmonics.

Exercises

17.1. Let $|E\rangle$ be a normalized eigenvector of a Hamiltonian H and let E be the corresponding eigenvalue. Suppose that $|\phi\rangle$ differs from $|E\rangle$ by a very small quantity, so that

$$|\phi\rangle = |E\rangle + \eta|\phi_1\rangle\,,$$

where $\eta \ll 1$. Here $|\phi_1\rangle$ is normalized but not necessarily orthogonal to $|E\rangle$. Show that $\langle\phi|\phi\rangle^{-1}\langle\phi|H|\phi\rangle$ differs form E by a term of order η^2, and find this term. Thus a first-order error on an eigenvector entails only a second-order error on the corresponding eigenvalue.

17.2. A particle of mass m is in a potential $V(r) = -V_0\mathrm{e}^{-r/a}$ where $mV_0a^2 = (4/3)\hbar^2$. Use the variational method with a trial function proportional to $\mathrm{e}^{-\alpha r}$ to obtain an upper bound on the ground-state energy. [A numerical calculation yields $E_0 = -0.0324\,V_0$.]

17.3. The harmonic oscillator's Hamiltonian is given by

$$H = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{1}{2}m\omega^2x^2\,.$$

a) Find an upper bound on the ground-state energy by using a trial function proportional to $\exp\{-\alpha|x|\}$, where α is a variational parameter.
b) Redo the calculation with a trial function proportional to $\exp\{-\alpha x^2\}$. How do the results compare with the exact value of the ground-state energy?

17.4. Use a trial function proportional to $\exp\{-\alpha x^2\}$ to obtain an approximate value of the ground-state energy associated with the following Hamiltonian:

$$H = -\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \lambda x^4\,, \quad \text{where} \quad \lambda > 0\,.$$

17.5. Compute functions $F_1(\alpha)$ and $F_2(\alpha)$ defined in (17.8) and (17.9).

17.6. The helium atom's Hamiltonian (17.4) can be written as

$$H = H_0 + W \,,$$

where $W = e^2/r_{12}$ and H_0 is the Hamiltonian of two hydrogen-like atoms. In stationary perturbation theory an approximation of the lowest eigenvalue of H is given by

$$E_0 = \varepsilon_0 + \langle 0|W|0\rangle \,,$$

where ε_0 and $|0\rangle$ are the eigenvalue and the normalized eigenvector associated with the ground state of H_0. Compare this approximation with the one obtained in Sect. 17.2 by a variational calculation.

17.7. Neglecting second-order terms in δu_i, obtain $F(u_i + \delta u_i) - F(u_i)$ from the definition (17.44) of the functional $F(u_i)$.

17.8. Carry out transformation (17.46) on (17.45) and show that matrix V can be picked so as to obtain (17.47).

17.9. Show that two solutions u_k and u_l of (17.47), associated with the same spin and corresponding to different constants λ_k and λ_l, are orthogonal.

17.10. Derive (17.50).

17.11. Prove Koopmans's theorem.

17.12. Prove Brillouin's theorem.

17.13. Substitute (17.51), for all i, in the Hartree–Fock equations (17.47). Show that if all electronic shells are filled, the angular dependence disappears, i.e. there only remains a system of coupled equations for radial functions.

17.14. Using integration by parts check that

$$\int_0^\infty \mathrm{d}\varrho \, \frac{1}{\varrho^5} \left(\sin\varrho - \varrho\cos\varrho\right)^2 = \frac{1}{4} \,.$$

17.15. Derive (17.60) from the stationary value of the functional (17.52).

18 Atomic Terms and Multiplets

The stationary energies and wave functions of an atom are obtained by diagonalizing its Hamiltonian. This diagonalization is carried out here in the state space associated with an electronic configuration. Insofar as the Hamiltonian only involves kinetic and potential energy terms, it commutes with the atom's orbital and spin angular momentum operators. In simple cases this yields a good approximation for atomic wave functions and energies, the latter being expressed with respect to the average energy of the configuration. Eventually, coupling terms between spin and orbital angular momenta are added to the Hamiltonian, as well as interaction terms with an external magnetic field. Degeneracy of atomic energies is then completely removed.

18.1 *LS* Coupling

The aim of this chapter is to develop, in simple cases at least, reasonably accurate methods for the determination of stationary atomic wave functions and energies. To set the problem in context let us recall some of the results already obtained.

We consider a neutral atom or an ion with Z protons and Q electrons. We assume for the moment that the Hamiltonian is given by (17.35), i.e. it only involves kinetic and potential energy terms. We are interested not only in the ground state but, in principle, in all discrete-spectrum states.

Atomic wave functions are the completely antisymmetric eigenfunctions of operator (17.35) and atomic energies are the corresponding eigenvalues.

We know how to build completely antisymmetric functions with respect to exchange of electronic coordinates. We must put together, as in (17.29), electronic wave functions $\phi_i(j)$ in Slater determinants. With the Hartree–Fock equations one can compute functions ϕ_i. We assume they have been obtained, are orthonormal and can be written in the form

$$\phi_i(1) = u_i(\boldsymbol{r}_1)\chi_i(1) = u_{n_i l_i}(r_1)Y_{l_i m_i}(\theta_1, \phi_1)\chi_i(1). \tag{18.1}$$

The representation of atomic wave functions by one Slater determinant, or by a linear combination of determinants associated with a given electronic configuration, is never more than an approximation. The usefulness of the electronic wave functions ϕ_i comes from the fact that this approximation is often rather good, but also and especially from the possibility to improve it systematically.

Recall that an *orbital* is an electronic wave function u_i with given quantum numbers n_i, l_i and m_i. The set of orbitals with the same n_i and l_i makes up an electronic *shell*. An *electronic configuration* consists in assigning the Q electrons of an atom to specific shells (some filled and some partly filled), i.e. in specifying the Q values of n_i and the Q values of l_i.

An electronic configuration in general corresponds to several Slater determinants. In a given configuration (that is, for fixed values of the n_i and l_i), a determinantal wave function is entirely specified by values of the quantum number m_i and m_{s_i}. A determinantal wave function associated with a given configuration can thus be denoted by

$$\Phi\left(m_1^{\mathrm{sign}(m_{s_1})}, m_2^{\mathrm{sign}(m_{s_2})}, \dots, m_Q^{\mathrm{sign}(m_{s_Q})}\right). \tag{18.2}$$

This expression looks complicated but actually it is rather simple. For three electrons, for example, one writes $\Phi(1^+, 0^-, -1^+)$, where integral numbers are values of the m_i and the signs denote values of the m_{s_i}.

Suppose one has a complete set of orthonormal electronic orbitals. In principle there are infinitely many ways to assign an atom's electrons to electronic shells, and therefore infinitely many configurations. Consider the set $\{|\Phi_a\rangle\}$ of all Slater determinants associated with all possible electronic configurations. One can reasonably believe that this set is a basis of the state space associated with the atom. If such is the case, the search for stationary atomic wave functions and energies is equivalent to the diagonalization of Hamiltonian (17.35) in this basis.

Obviously this diagonalization cannot be carried out exactly. It can, however, be carried out approximately if the state space is restricted to a finite number of configurations. The larger the number of configurations used, the better the approximation. In most of this chapter the state space will be restricted to Slater determinants belonging to just one configuration. We will come back to multiple configurations in Sect. 18.7.

Let $\mathcal{V}$ be the subspace of the state space associated with a given electronic configuration and let $\mathcal{N}$ be the dimension of $\mathcal{V}$. When restricted to $\mathcal{V}$, the Hamiltonian of an atom is a $\mathcal{N} \times \mathcal{N}$ matrix.[1] The value of $\mathcal{N}$ can quickly become rather large. Suppose for instance that the atom contains two p electrons outside closed shells. A p electron has three possible values of m_i and, for each of these, two possible values of m_{s_i}. If the electrons are not equivalent (that is, if their quantum numbers n_i are different), this amounts to 36 states. For equivalent electrons the Pauli principle reduces this number to 15 since there are $6 \times 5/2$ ways to pick two electronic states among

---

[1] In the central-field model this matrix is diagonal and has only one $\mathcal{N}$-fold degenerate eigenvalue. Indeed electronic energies then in no way depend on angular variables.

6 different states. If the atom has one p electron and one d electron outside closed shells there are 60 orthogonal states. The reader can check that for the lowest-energy configuration of uranium ($5f^3 6d\, 7s^2$ outside closed shells), $\mathcal{N} = 3640$.

Our problem thus consists in diagonalizing Hamiltonian (17.35) in the vector space $\mathcal{V}$ associated with a given configuration. This operation is considerably simplified by a close look at operators that commute with H.

Let us first consider exchange operators for electronic coordinates. Specifically, let P_{ij} be the operator that permutes the coordinates of electron i with those of electron j. Clearly P_{ij} commutes with H. Any antisymmetric function is an eigenvector of P_{ij} with eigenvalue -1. Therefore the exclusive use of such functions automatically diagonalizes the P_{ij}.[2]

Let us now turn to spin and angular momentum operators. For each electron we introduce an orbital angular momentum operator $\boldsymbol{L}_i$ and a spin operator $\boldsymbol{S}_i$. Let $\boldsymbol{L}$ and $\boldsymbol{S}$ denote the total orbital angular momentum and the total spin, defined as

$$\boldsymbol{L} = \sum_{i=1}^{Q} \boldsymbol{L}_i , \qquad \boldsymbol{S} = \sum_{i=1}^{Q} \boldsymbol{S}_i . \tag{18.3}$$

One can check (Exercise 18.1) that the following commutation relations hold:

$$[H, \boldsymbol{L}_i] \neq 0 , \qquad [H, \boldsymbol{L}] = 0 , \tag{18.4}$$

$$[H, \boldsymbol{S}_i] = 0 , \qquad [H, \boldsymbol{S}] = 0 . \tag{18.5}$$

Hence a priori, eigenfunctions of H could also be eigenfunctions of $\boldsymbol{L}^2$ and L_z, and of $\boldsymbol{S}_i^2$ and S_{iz} for every i.

But here the postulate of antisymmetry brings additional restrictions. Indeed for all i and j,

$$[P_{ij}, \boldsymbol{L}_i] \neq 0 , \qquad [P_{ij}, \boldsymbol{L}] = 0 , \tag{18.6}$$

$$[P_{ij}, \boldsymbol{S}_i] \neq 0 , \qquad [P_{ij}, \boldsymbol{S}] = 0 . \tag{18.7}$$

The two inequalities to the left entail that in general, antisymmetric functions are not eigenfunctions of L_{iz} and S_{iz}. The two equations to the right, by contrast, imply that diagonalization of $\boldsymbol{L}^2$, L_z, $\boldsymbol{S}^2$ and S_z is compatible with the antisymmetry property. To sum up, the completely antisymmetric eigenfunctions of H can always be picked as eigenfunctions of $\boldsymbol{L}^2$, L_z, $\boldsymbol{S}^2$ and S_z. In general, however, they are not eigenfunctions of L_{iz} or S_{iz}.[3]

---

[2] Note that P_{ij} and P_{kl} do not commute if one of the indices ij coincides with one of the indices kl. In general P_{ij} and P_{kl} are therefore not simultaneously diagonalizable. But antisymmetric wave functions, which make up one-dimensional irreducible representation of S_n, are eigenfunctions of all operators P_{ij}.

[3] Eigenfunctions of H are also eigenfunctions of ${\boldsymbol{S}_i}^2$ for every i. This comes from the fact that all electrons have a spin equal to 1/2. By contrast, eigenfunctions of H are not eigenfunctions of $\boldsymbol{L}_i^2$.

Operators $\boldsymbol{L}^2$, L_z, $\boldsymbol{S}^2$ and S_z, together with H [given by (17.35)], make up a (not complete) set of commuting operators. Denote by L, M_L, S and M_S the quantum numbers associated with eigenvalues of $\boldsymbol{L}^2$ L_z, $\boldsymbol{S}^2$ and S_z, respectively. The diagonalization of H can thus be carried out in subspaces of $\mathcal{V}$ associated with fixed values of L, M_L, S and M_S.[4]

In Sect. 18.2 we will carry out this calculation for a few simple configurations. Here let us introduce some notation. A *multiplet* is the set of all states which, for a given configuration, have same L, same S and different values of M_L and M_S, and can all be obtained one from another through application of operators L_+, L_-, S_+ and S_-. Within a configuration, different multiplets are labeled by means of the spectroscopic notation ${}^{2S+1}L$.[5] Capital letters S, P, D, F, ... stand for values 0, 1, 2, 3, ... of L. There are multiplets ${}^1S$, ${}^3P$, ${}^2D$, etc. The energy interval between multiplets that belong to the same configuration is typically of order one eV or less, whereas the energy interval between different configurations is of the order of several eV.

Values of the total orbital angular momentum L are obtained by coupling one after another, with Clebsch–Gordan coefficients, orbital angular momenta of all the electrons. So is it with spin. This procedure is called *LS coupling.* In complex atoms there are in principle many ways to carry out LS coupling. But actually, atoms are most often made mostly of closed shells, with one or two partly filled shells. It is then extremely convenient first to couple angular momenta of electrons that belong to the same shell.

Let $\bar{L}$ and $\bar{S}$ be the quantum numbers associated with the total orbital and spin angular momentum of a given electronic shell. Owing to the antisymmetry of wave functions, the coupling of equivalent electrons does not produce all values of $\bar{L}$ and $\bar{S}$ one would expect on the basis of Clebsch–Gordan coefficients. For a closed shell, in particular, the following result holds.

Lemma Let $\bar{L}$ and $\bar{S}$ be the quantum numbers associated with the total orbital and spin angular momentum of a completely filled electronic shell. Then $\bar{L} = 0 = \bar{S}$.

Proof Define

$$\bar{L}_+ = \sum_{\text{shell}} L_{i+} , \qquad \bar{S}_+ = \sum_{\text{shell}} S_{i+} , \tag{18.8}$$

---

[4] Note also that the space inversion operator P commutes with the Hamiltonian. From (7.169) one can see that ϕ_i given in (18.1) is an eigenfunction of P with eigenvalue $(-1)^{l_i}$. Hence a Slater determinant built with functions ϕ_i is also an eigenfunction of P, with eigenvalue $\prod_i (-1)^{l_i}$. Thus all wave functions associated with a given configuration have the same parity. Parity becomes relevant when one considers different configurations.

[5] We point out that within a configuration there can be several multiplets with the same L and S. This, however, only happens if there are more than two electrons outside closed shells.

and similarly with $\bar{L}_-$ and $\bar{S}_-$. The action of $\bar{L}_+$ on a Slater determinant is the sum of the actions of all L_{i+}. But these vanish if $m_i = l_i$ and produce a determinant with two identical columns if $m_i \neq l_i$. Hence the action of $\bar{L}_+$ on a determinant vanishes. So is it with the action of $\bar{S}_+$, $\bar{L}_-$ and $\bar{S}_-$. ♣

In this way the electrons' angular momenta are first coupled within each shell, so as to obtain the total orbital and the total spin angular momentum of each shell. Next the angular momenta of each shell are coupled to yield the total L and S. Only angular momenta of partly filled shells contribute to the total angular momenta, since angular momenta of closed shells vanish.

As an example, consider a configuration with a p electron and a d electron outside closed shells. The total orbital angular momentum comes from the sum of two angular momenta $l_1 = 1$ and $l_2 = 2$. Hence $L = 1, 2$ or 3. Similarly, $S = 0$ or 1. This configuration therefore gives rise to multiplets ${}^1P$, ${}^1D$, ${}^1F$, ${}^3P$, ${}^3D$ and ${}^3F$.

This schematic discussion of the coupling of atomic angular momenta should not lead one to believe that the determination of atomic angular momenta and wave functions is a simple matter. As soon as several equivalent electrons occupy partly filled shells, the problem becomes extremely complex. A few simple cases will be investigated in the following section. More complicated cases require the use of Racah's formalism [188], [189], [190], developed for instance in [59].

At any rate suppose that for a given configuration, allowed values of L and S have been obtained. Suppose furthermore that to each pair (L, S) there corresponds only one multiplet. In the approximation where the state space is restricted to the subspace $\mathcal{V}$ associated with the configuration given, the eigenfunctions of H coincide with vectors $|L, S; M_L, M_S\rangle$, uniquely determined. If, by contrast, several multiplets correspond to a pair (L, S), then the determination of eigenfunctions of H requires diagonalizing this operator in the subspace of vectors with the same L, S, M_L and M_S.

18.2 Atomic Wave Functions

We now derive the wave functions of multiplets associated with a few simple electronic configurations. We know that closed shells, with zero angular momenta, do not bear on how the configuration's various states are distributed into multiplets. To avoid proliferation of symbols, only electrons outside closed shells will be considered explicitly. Thus the determinant $\Phi(0^+, 0^-, 0^+, 0^-, 1^+, 1^-, 0^+, 0^-, -1^+, -1^-, 0^+)$, associated with the $1s^2 2s^2 2p^6 3s$ configuration of sodium, will simply be denoted by $\Phi(0^+)$.

i) No Electrons

Even though this case is trivial it should be mentioned. A configuration where all shells are filled only gives rise to a ${}^1S$ multiplet.

ii) One Electron

A single electron outside closed shells yields a single ^2L multiplet, with $L = l_1$ and $S = s_1 = 1/2$. One has ^2S, ^2P, ^2D, ... if $L = 0, 1, 2, \ldots$ The multiplet contains $2(2L+1)$ orthogonal states with wave functions denoted by $\Phi(m_1{}^{\pm})$.

iii) Two *s* Electrons

Here $l_1 = 0 = l_2$ so that $L = 0$. Since $s_1 = 1/2 = s_2$ one gets $S = 0$ or $S = 1$. Actually, wave functions were obtained in (17.25)–(17.28). They are given in Table 18.1 in our present notation.

Table 18.1. Wave functions of multiplets of two nonequivalent *s* electrons

^{2S+1}L	M_L	M_S	Wave functions
^3S	0	1	$\Phi(0^+,0^+)$
	0	0	$\frac{1}{\sqrt{2}}\{\Phi(0^+,0^-)+\Phi(0^-,0^+)\}$
	0	-1	$\Phi(0^-,0^-)$
^1S	0	0	$\frac{1}{\sqrt{2}}\{\Phi(0^+,0^-)-\Phi(0^-,0^+)\}$

Thus a configuration of two *s* electrons normally gives rise to two multiplets, a triplet and a singlet. Both multiplets are always there when the two *s* orbitals are different. If, on the other hand, they are equivalent (two 1*s* orbitals, for example), only the singlet survives. This was pointed out in Sect. 17.3, where we saw that the triplet's spatial wave function is antisymmetric with respect to the exchange of $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$ [(17.25)–(17.27)]. Thus wave functions given in Table 18.1 are appropriate to the case of two different *s* orbitals, and then they are normalized. For two equivalent *s* orbitals, the singlet's normalized wave function is simply given by $\Phi(0^+, 0^-)$.

iv) One *s* Electron and One *p* Electron

One sees easily that $L = 1$ and $S = 0$ or 1. The configuration gives rise to two multiplets, ^1P and ^3P. There are twelve linearly independent wave functions. We will not write them explicitly, but we point out that they can all be obtained form Table 18.1 by replacing the second zero in determinantal wave functions successively by $+1$, 0 and -1. Obviously the case of equivalent orbitals is not relevant here.

v) Two *p* Electrons

Here is a more complicated case. One has $l_1 = 1 = l_2$, so that $L = 0$, 1 or 2. Of course $S = 0$ or 1. The configuration gives rise to six multiplets, i.e. three triplets (^3S, ^3P, ^3D) and three singlets (^1S, ^1P, ^1D). There are 36 determinantal wave functions. Eigenfunctions of $\boldsymbol{L}^2$, L_z, $\boldsymbol{S}^2$ and S_z are linear combinations of these determinantal functions. Let us see how to derive them.

The only determinantal wave function for which $M_L = 2$ and $M_S = 1$ is $\Phi(1^+, 1^+)$. It necessarily belongs to the ^3D multiplet. Other wave functions in the ^3D are obtained from this one by applying the operators L_- and S_-. For example,[6]

$$\begin{aligned} L_-\Phi\left(m_1^+, m_2^+\right) &= L_{1-}\Phi\left(m_1^+, m_2^+\right) + L_{2-}\Phi\left(m_1^+, m_2^+\right) \\ &= \hbar\left\{(l_1 + m_1)(l_1 + 1 - m_1)\right\}^{1/2}\Phi\left((m_1 - 1)^+, m_2^+\right) \\ &\quad + \hbar\left\{(l_2 + m_2)(l_2 + 1 - m_2)\right\}^{1/2}\Phi\left(m_1^+, (m_2 - 1)^+\right). \end{aligned} \tag{18.9}$$

Using (18.9) and the corresponding relation for S_-, and taking normalization into account, one finds (Exercise 18.3) the 15 wave functions of the ^3D multiplet. A few of them are given in Table 18.2.

Let us now turn to the ^3P multiplet. Its wave functions must be orthogonal to those of the ^3D. This (up to a phase) determines the wave function associated with values $M_L = 1$ and $M_S = 1$. The other ones are obtained by applying operators L_- and S_- (Exercise 18.4). In all there are 9 different wave functions in the ^3P, a few of them shown in Table 18.2.

The wave functions of other multiplets are similarly derived, by using orthogonality conditions and applying operators L_- and S_- (Exercise 18.5). Table 18.2 displays a few of them.

The wave functions we have derived are valid if the two p orbitals are different. When they are equivalent one can see that some multiplets disappear due to the antisymmetry of determinantal wave functions. Let us examine, for instance, the $M_L = 2$, $M_S = 1$ wave function of the ^3D triplet. It is given by $\Phi(1^+, 1^+)$. One easily sees that this determinant vanishes identically if the two spatial functions coincide. Since all wave functions in the multiplet are obtained from the first one by applying operators L_- and S_-, the whole multiplet disappears. Such is also the case with the ^3S triplet and the ^1P singlet. Thus only the ^3P, ^1D and ^1S survive. Wave function determination is straightforward. Take for instance the $M_L = 1$, $M_S = 1$ state in the ^3P triplet. Since orbitals are equivalent, one has $\Phi(1^+, 0^+) = -\Phi(0^+, 1^+)$. The required wave function is therefore simply equal to $\Phi(1^+, 0^+)$. Wave functions of all the other states in the ^3P, ^1D and ^1S multiplets, for equivalent orbitals, are similarly obtained by regrouping equal terms in Table 18.2 and suitably correcting the normalization factor.

---

[6] Strictly speaking the notation in (18.9) is different from the one used elsewhere. Operator L_{1-}, for instance, does not denote here the angular momentum associated with electron (1), but rather the angular momentum associated with wave function (1). The substitution is allowed, however, because L_- involves the coordinates of all electrons symmetrically. A similar remark applies to (18.8).

Table 18.2. Wave functions of multiplets of two nonequivalent p electrons

^{2S+1}L	M_L	M_S	Wave functions
^3D	2	1	$\Phi(1^+,1^+)$
	1	1	$\frac{1}{\sqrt{2}}\left\{\Phi(1^+,0^+)+\Phi(0^+,1^+)\right\}$
	0	1	$\frac{1}{\sqrt{6}}\left\{\Phi(1^+,-1^+)+2\Phi(0^+,0^+)+\Phi(-1^+,1^+)\right\}$
	2	0	$\frac{1}{\sqrt{2}}\left\{\Phi(1^+,1^-)+\Phi(1^-,1^+)\right\}$
	1	0	$\frac{1}{2}\left\{\Phi(1^+,0^-)+\Phi(0^+,1^-)+\Phi(1^-,0^+)+\Phi(0^-,1^+)\right\}$
	0	0	$\frac{1}{2\sqrt{3}}\{\Phi(1^+,-1^-)+2\Phi(0^+,0^-)+\Phi(-1^+,1^-)$
			$+\Phi(1^-,-1^+)+2\Phi(0^-,0^+)+\Phi(-1^-,1^+)\}$
^3P	1	1	$\frac{1}{\sqrt{2}}\left\{\Phi(1^+,0^+)-\Phi(0^+,1^+)\right\}$
	0	1	$\frac{1}{\sqrt{2}}\left\{\Phi(1^+,-1^+)-\Phi(-1^+,1^+)\right\}$
	1	0	$\frac{1}{2}\left\{\Phi(1^+,0^-)+\Phi(1^-,0^+)-\Phi(0^+,1^-)-\Phi(0^-,1^+)\right\}$
	0	0	$\frac{1}{2}\left\{\Phi(1^+,-1^-)+\Phi(1^-,-1^+)-\Phi(-1^+,1^-)-\Phi(-1^-,1^+)\right\}$
^3S	0	1	$\frac{1}{\sqrt{3}}\left\{\Phi(1^+,-1^+)-\Phi(0^+,0^+)+\Phi(-1^+,1^+)\right\}$
	0	0	$\frac{1}{\sqrt{6}}\{\Phi(1^+,-1^-)+\Phi(1^-,-1^+)-\Phi(0^+,0^-)$
			$-\Phi(0^-,0^+)+\Phi(-1^+,1^-)+\Phi(-1^-,1^+\}$
^1D	2	0	$\frac{1}{\sqrt{2}}\left\{\Phi(1^+,1^-)-\Phi(1^-,1^+)\right\}$
	1	0	$\frac{1}{2}\left\{\Phi(1^+,0^-)+\Phi(0^+,1^-)-\Phi(1^-,0^+)-\Phi(0^-,1^+)\right\}$
	0	0	$\frac{1}{2\sqrt{3}}\{\Phi(1^+,-1^-)+2\Phi(0^+,0^-)+\Phi(-1^+,1^-)$
			$-\Phi(1^-,-1^+)-2\Phi(0^-,0^+)-\Phi(-1^-,1^+\}$
^1P	1	0	$\frac{1}{2}\left\{\Phi(1^+,0^-)-\Phi(0^+,1^-)-\Phi(1^-,0^+)+\Phi(0^-,1^+)\right\}$
	0	0	$\frac{1}{2}\{\Phi(1^+,-1^-)-\Phi(-1^+,1^-)-\Phi(1^-,-1^+)+\Phi(-1^-,1^+\}$
^1S	0	0	$\frac{1}{\sqrt{6}}\{\Phi(1^+,-1^-)-\Phi(0^+,0^-)+\Phi(-1^+,1^-)$
			$-\Phi(1^-,-1^+)+\Phi(0^-,0^+)-\Phi(-1^-,1^+)\}$

vi) Three s Electrons

A configuration with three s electrons gives rise to one ^4S and two ^2S multiplets. Normalized wave functions of the ^4S are obtained by applying operator S_- to the determinant $\Phi(0^+,0^+,0^+)$. They are given in Table 18.3.

By contrast, the wave functions of the ^2S multiplets cannot be derived uniquely by the method illustrated in the previous example. Indeed there are three linearly independent wave functions that correspond to values $M_L = 0$ and $M_S = 1/2$. The orthogonality condition to the wave function associated with the ^4S determines a two-dimensional vector space, and nothing more. To properly derive the two ^2S multiplets, it is necessary to diagonalize Hamiltonian (17.35) in this two-dimensional space.

Table 18.3. Wave functions of the ^4S multiplet of three nonequivalent s electrons

M_L	M_S	Wave functions
0	3/2	$\Phi(0^+,0^+,0^+)$
0	1/2	$\frac{1}{\sqrt{3}}\{\Phi(0^+,0^+,0^-)+\Phi(0^+,0^-,0^+)+\Phi(0^-,0^+,0^+)\}$
0	−1/2	$\frac{1}{\sqrt{3}}\{\Phi(0^+,0^-,0^-)+\Phi(0^-,0^+,0^-)+\Phi(0^-,0^-,0^+)\}$
0	−3/2	$\Phi(0^-,0^-,0^-)$

Note, however, that if the first two s orbitals are equivalent, the configuration gives rise to only one multiplet, a ^2S doublet. Associated wave functions are easily obtained.

vii) More Complicated Cases

In every case with only two electrons outside closed shells, there is only one multiplet for each value of the pair (L, S). Wave functions can all be obtained by using the orthogonality condition and applying operators L_- and S_-. For equivalent orbitals one can show ([233], Sect. 20.1) that a ^{2S+1}L multiplet survives if and only if $L + S$ is even. Specifically ^1S, ^3P, ^1D, ^3F, ... survive while ^3S, ^1P, ^3D, ^1F, ... disappear.

When more than two electrons are found outside closed shells, a given value of (L, S) corresponds in general to more than one multiplet. In the approximation where the state space is restricted to just one configuration, the atomic wave functions can be obtained as follows. In the space $\mathcal{V}$ corresponding to states in the configuration, one considers the subspace made up of all vectors like $|L, S; M_L, M_S\rangle$, with fixed values for these four quantum numbers. Next the Hamiltonian is diagonalized in that subspace. This manipulation must be repeated for every value of the pair (L, S), but nor for different values of M_L and M_S. Indeed eigenvectors of H with fixed (L, S) and variable (M_L, M_S) are all obtained from one another through application of operators L_+, L_-, S_+ and S_-.

In the process of diagonalizing the Hamiltonian the evaluation of matrix elements soon becomes very complicated, especially with many equivalent electrons. Examples of such calculation can be found in [56], Sect. 8.7, [211], Chaps. 21 and 22, [233], Chap. 21 and, with much generality, in [59], Chaps. 10 and 12.

Two electronic configurations are called *conjugate* if they involve the same closed shells, and if to each partly filled nl shell with q electrons there corresponds, in the other configuration, an nl shell with $\{2(2l + 1) - q\}$ electrons. Thus $1s^22s^22p^2$ and $1s^22s^22p^4$ are conjugate, and similarly with $1s^22s^22p3s$ and $1s^22s^22p^53s$. One can show that conjugate configurations give rise to the same multiplets. Wave functions associated with conjugate configurations are very similar. The quantum numbers associated in one case with electrons outside closed shells correspond to quantum numbers associated in the other case with *holes*, i.e. with electrons extracted from closed shells. By way of example, recall that the wave function of the $M_L = 1$, $M_S = 1$ state in the ${}^3P$ triplet of two equivalent p electrons in equal to $\Phi(1^+, 0^+)$. The wave function of the $M_L = -1$, $M_S = -1$ state in the ${}^3P$ triplet of four equivalent p electrons is then given by $\Phi(-1^-, 0^-, 1^-, -1^+)$.

18.3 Average Energy of a Configuration

Insofar as the Hamiltonian of an atom is given by (17.35), all states in a multiplet have the same energy. We will compute this energy, in simple cases, in Sect. 18.4. Before that, however, it is useful to inquire into the value of energy that can roughly be associated with a given configuration. Indeed we know that the energy difference between multiplets in the same configuration is, in general, much smaller than the energy difference between multiplets belonging to different configurations.

Consider an electronic configuration involving $\mathcal{N}$ quantum states. Let $\{|\Phi_a\rangle,\ a = 1,\ 2,\ \ldots,\ \mathcal{N}\}$ be a basis of orthonormal vectors associated with this configuration. The *average energy of the configuration* is defined as

$$E_{\mathrm{av}} = \frac{1}{\mathcal{N}} \sum_{a=1}^{\mathcal{N}} \langle \Phi_a | H | \Phi_a \rangle \,. \tag{18.10}$$

Since E_{av} is proportional to the trace of the Hamiltonian in the subspace associated with the configuration, E_{av} does not depend on the choice of orthonormal vectors $|\Phi_a\rangle$. We will use determinantal functions, i.e. functions that do not necessarily belong to a specific multiplet.

The atom's Hamiltonian can be written as in (17.36) and (17.37). Wave functions Φ_a are given by (17.38), with orbitals specified as in (18.1). Following results of Sect. 17.3, the computation of $\langle \Phi_a | H | \Phi_a \rangle$ involves terms like $\langle \phi_i | F | \phi_i \rangle$, $\langle \phi_i \phi_j | G | \phi_i \phi_j \rangle$ and $\langle \phi_i \phi_j | G | \phi_j \phi_i \rangle$. We will first evaluate these terms, in the general form $\langle \phi_i | F | \phi_j \rangle$ and $\langle \phi_i \phi_j | G | \phi_s \phi_t \rangle$.

Expressions $\langle \phi_i | F | \phi_j \rangle$ are easily evaluated with the help of orthogonality properties of spherical harmonics. Since radial functions are real one finds that

$$\begin{aligned} \langle \phi_i | F | \phi_j \rangle &= \delta(i,j) \delta_{l_i l_j} \delta_{m_i m_j} \int_0^\infty \mathrm{d}r\, u_{n_i l_i}(r) \\ &\times \left\{ -\frac{\hbar^2}{2m} \left[\frac{\mathrm{d}}{\mathrm{d}r} \left(r^2 \frac{\mathrm{d}}{\mathrm{d}r} \right) - l_j(l_j+1) \right] - Ze^2 r \right\} u_{n_j l_j}(r) \,. \end{aligned} \tag{18.11}$$

Recall that $\delta(i,j)$ is equal to one or zero, according as m_{s_i} does or does not coincide with m_{s_j}. In the case where $i = j$ it is convenient to introduce the notation $I(n_i l_i) \equiv \langle \phi_i | F | \phi_i \rangle$. One shows that

$$I(n_il_i) \equiv \langle \phi_i | F | \phi_i \rangle$$
$$= \int_0^\infty \mathrm{d}r \left\{ \frac{\hbar^2}{2m} r^{2l_i+2} \left[\frac{\mathrm{d}}{\mathrm{d}r} \left(\frac{u_{n_il_i}}{r^{l_i}} \right) \right]^2 - Ze^2 r \left[u_{n_il_i}(r) \right]^2 \right\}. \quad (18.12)$$

The evaluation of $\langle \phi_i \phi_j | G | \phi_s \phi_t \rangle$ is somewhat more complicated. To carry it out one uses the expression of $1/r_{12}$ in a series of spherical harmonics given in (17.13). There result integrals involving products of three spherical harmonics. Define the coefficient $c^l(l_s m_s; l_i m_i)$ as

$$\sqrt{\frac{4\pi}{2l+1}} \int \mathrm{d}\Omega_1 \, Y^*_{l_i m_i}(\theta_1, \phi_1) Y^*_{lm}(\theta_1, \phi_1) Y_{l_s m_s}(\theta_1, \phi_1)$$
$$= \delta_{m_i + m, m_s} c^l(l_s m_s; l_i m_i) \,. \quad (18.13)$$

Properties of spherical harmonics entail that coefficients $c^l(l_s m_s; l_i m_i)$ are real and satisfy the equations

$$c^l(l_i m_i; l_s m_s) = (-1)^{m_s - m_i} c^l(l_s m_s; l_i m_i) \,, \quad (18.14)$$
$$c^0(l_s m_s; l_i m_i) = \delta_{l_i l_s} \delta_{m_i m_s} \,. \quad (18.15)$$

Actually, the coefficients $c^l(l_s m_s; l_i m_i)$ can be written in terms of Clebsch–Gordan coefficients. To see this one writes $Y^*_{l_i m_i} = (-1)^{m_i} Y_{l_i, -m_i}$ and uses formula (14.82) for the product $Y_{l_s m_s} Y_{l_i, -m_i}$ that appears in (18.13). Orthogonality of spherical harmonics then implies that

$$c^l(l_s m_s; l_i m_i) = (-1)^{m_i} \frac{\sqrt{(2l_i+1)(2l_s+1)}}{2l+1}$$
$$\times \langle l_s, l_i; m_s, -m_i | l, m_s - m_i \rangle \langle l_s, l_i; 0, 0 | l, 0 \rangle \,. \quad (18.16)$$

Clearly $c^l(l_s m_s; l_i m_i)$ vanishes unless $|l_i - l_s| \leq l \leq l_i + l_s$. One finds in Table 18.4 values of the coefficients c^l for s and p shells.[7]

With the help of coefficients c^l the matrix element $\langle \phi_i \phi_j | G | \phi_s \phi_t \rangle$ can be written as (Exercise 18.6)

$$\langle \phi_i \phi_j | G | \phi_s \phi_t \rangle = \delta(i, s) \delta(j, t) \delta_{m_i + m_j, m_s + m_t}$$
$$\times \sum_{l=0}^{\infty} c^l(l_s m_s; l_i m_i) c^l(l_j m_j; l_t m_t) R^l(ij; st) \,, \quad (18.17)$$

where

$$R^l(ij; st) = e^2 \int_0^\infty \mathrm{d}r_1 \, r_1^2 \int_0^\infty \mathrm{d}r_2 \, r_2^2 \frac{r_<^l}{r_>^{l+1}} u_{n_i l_i}(r_1)$$
$$\times u_{n_j l_j}(r_2) u_{n_s l_s}(r_1) u_{n_t l_t}(r_2) \,. \quad (18.18)$$

---

[7] More complete tables can be found in [56], Sect. 6.9 and in [211], App. 20.

Table 18.4. Values of $c^l(l_i m_i; l_j m_j)$

			l		
$l_i l_j$	m_i	m_j	0	1	2
ss	0	0	1	0	0
sp	0	± 1	0	$-\sqrt{1/3}$	0
	0	0	0	$\sqrt{1/3}$	0
pp	0	± 1	0	0	$-\sqrt{3/25}$
	0	0	1	0	$\sqrt{4/25}$
	1	1	1	0	$-\sqrt{1/25}$
	1	-1	0	0	$-\sqrt{6/25}$
	-1	-1	1	0	$-\sqrt{1/25}$

Note that the sum over l does not really extend to infinity, because it is restricted by properties of coefficients c^l. Integrals on $\mathrm{d}r_1$ and $\mathrm{d}r_2$ can be carried out if radial functions are known.

It is convenient to introduce notations specific to the case where $(ij) = (st)$ or (ts). Letting

$$F^l(n_i l_i; n_j l_j) = R^l(ij; ij) , \tag{18.19}$$

$$G^l(n_i l_i; n_j l_j) = R^l(ij; ji) , \tag{18.20}$$

one gets

$$\langle \phi_i \phi_j | G | \phi_i \phi_j \rangle = \sum_l c^l(l_i m_i; l_i m_i) c^l(l_j m_j; l_j m_j) F^l(n_i l_i; n_j l_j) , \tag{18.21}$$

$$\langle \phi_i \phi_j | G | \phi_j \phi_i \rangle = \delta(i,j) \sum_l \left[c^l(l_j m_j; l_i m_i) \right]^2 G^l(n_i l_i; n_j l_j) . \tag{18.22}$$

It is easy to see that for every l, $F^l(n_i l_i; n_j l_j) > 0$. One can also show [189] that $G^l(n_i l_i; n_j l_j) > 0$.

We now have everything needed to evaluate the matrix element $\langle \Phi_a | H | \Phi_a \rangle$. Making use of formulas (17.33), (17.34), (18.12), (18.21) and (18.22) we obtain

$$\langle \Phi_a | H | \Phi_a \rangle = \sum_{i=1}^{Q} E_i^a + \sum_{i<j}^{Q} E_{ij}^a , \tag{18.23}$$

where

$$E_i^a = I(n_i l_i) , \tag{18.24}$$

$$E_{ij}^a = \sum_l c^l(l_i m_i; l_i m_i) c^l(l_j m_j; l_j m_j) F^l(n_i l_i; n_j l_j)$$
$$- \delta(i,j) \sum_l \left[c^l(l_j m_j; l_i m_i)\right]^2 G^l(n_i l_i; n_j l_j) . \tag{18.25}$$

One should note that the index a, which labels a determinantal function, corresponds to values of m_i and m_{s_i}.

To obtain the average energy of the configuration, one must use (18.10). All wave functions in the configuration have the same values of the quantum numbers n_i and l_i, but they are distinguished by values of m_i and m_{s_i}. Thus different terms in the sum (18.10) correspond to different possible values of these numbers.[8] As a first step, it is useful to perform the summation on m_j and m_{s_j} of the two different combinations of coefficients c^l that appear in (18.25). Equations (18.13) and (18.15), together with properties of spherical harmonics, imply that

$$\begin{aligned}
\sum_{m_j, m_{s_j}} c^l(l_i m_i; l_i m_i) c^l(l_j m_j; l_j m_j) &= 2c^l(l_i m_i; l_i m_i) \sum_{m_j} c^l(l_j m_j; l_j m_j) \\
&= 2c^l(l_i m_i; l_i m_i) \sum_{m_j} \sqrt{\frac{4\pi}{2l+1}} \int \mathrm{d}\Omega_1 \, |Y_{l_j m_j}(\theta_1, \phi_1)|^2 Y_{l0}^*(\theta_1, \phi_1) \\
&= 2c^l(l_i m_i; l_i m_i) \sqrt{\frac{4\pi}{2l+1}} \frac{(2l_j+1)}{4\pi} \int \mathrm{d}\Omega_1 \, Y_{l0}^*(\theta_1, \phi_1) \\
&= 2(2l_j+1)\delta_{0l} \, .
\end{aligned} \tag{18.26}$$

Moreover, using properties of the Clebsch–Gordan coefficients one finds that

$$\begin{aligned}
\sum_{m_j, m_{s_j}} \delta(i,j) \left[c^l(l_j m_j; l_i m_i)\right]^2 &= \sum_{m_j} \left[c^l(l_j m_j; l_i m_i)\right]^2 \\
&= \frac{(2l_i+1)(2l_j+1)}{(2l+1)^2} \langle l_j, l_i; 0, 0 | l, 0\rangle^2 \sum_{m_j} \langle l_j, l_i; m_j, -m_i | l, m_j - m_i\rangle^2 \\
&= \frac{2l_j+1}{2l+1} \langle l_j, l_i; 0, 0 | l, 0\rangle^2 \sum_{m_j} \langle l, l_j; m_i - m_j, m_j | l_i, m_i\rangle^2 \\
&= \frac{2l_j+1}{2l+1} \langle l_j, l_i; 0, 0 | l, 0\rangle^2 \, .
\end{aligned} \tag{18.27}$$

We can now evaluate the configuration average of each term on the right-hand side of (18.23). Note first that all terms E_i^a are independent of m_i and m_{s_i}. Thus

---

[8] The correspondence is not one-to-one. Nevertheless, with some thought one can see that an average computed as in (18.10) coincides with an average on all values of m_i and m_{s_i}, provided the Pauli exclusion principle is taken into account.

$$E_i^{\mathrm{av}} = \frac{1}{\mathcal{N}}\sum_{a=1}^{\mathcal{N}} E_i^a = I(n_i l_i)\,. \tag{18.28}$$

To evaluate the average of E_{ij}^a we must distinguish cases where electrons i and j are or are not equivalent. Suppose first that they are not. There are then $4(2l_i+1)(2l_j+1)$ possible values of the quadruplet $(m_i, m_{s_i}, m_j, m_{s_j})$. Making use of (18.26) and (18.27) and suppressing, for simplicity, the arguments $(n_i l_i; n_j l_j)$ of F^l and G^l, one gets

$$\begin{aligned}
E_{ij}^{\mathrm{av}} &= \frac{1}{\mathcal{N}}\sum_{a=1}^{\mathcal{N}} E_{ij}^a = \frac{1}{4(2l_i+1)(2l_j+1)}\sum_{m_i,m_{s_i}}\sum_{m_j,m_{s_j}} E_{ij}^a \\
&= \frac{1}{4(2l_i+1)(2l_j+1)} \\
&\quad \times \sum_{m_i,m_{s_i}}\sum_{l}\left\{2(2l_j+1)\delta_{0l}F^l - \frac{2l_j+1}{2l+1}\langle l_j, l_i; 0, 0|l, 0\rangle^2 G^l\right\} \\
&= F^0 - \frac{1}{2}\sum_l \frac{\langle l_j, l_i; 0, 0|l, 0\rangle^2}{2l+1} G^l\,.
\end{aligned} \tag{18.29}$$

The last equality comes from the fact that the expression between curly brackets depends neither on m_i nor on m_{s_i}.

If the two electrons are equivalent, then $n_i = n_j$ and $l_i = l_j$. From (18.19) and (18.20) one easily sees that $F^l(n_i l_i; n_i l_i) = G^l(n_i l_i; n_i l_i)$. For each value of m_i and m_{s_i} there are only $4l_i+1$ allowed values of m_j and m_{s_j}. Nevertheless the sum may include the term where $m_j = m_i$ and $m_{s_j} = m_{s_i}$, since it vanishes. In evaluating the average one must divide by $4l_i+1$. Therefore one finds [suppressing arguments $(n_i l_i; n_i l_i)$]

$$\begin{aligned}
E_{ij}^{\mathrm{av}} &= \frac{1}{\mathcal{N}}\sum_{a=1}^{\mathcal{N}} E_{ij}^a = \frac{1}{2(2l_i+1)(4l_i+1)}\sum_{m_i,m_{s_i}}\sum_{m_j,m_{s_j}} E_{ij}^a \\
&= \frac{1}{2(2l_i+1)(4l_i+1)} \\
&\quad \times \sum_{m_i,m_{s_i}}\sum_{l}\left\{2(2l_i+1)\delta_{0l}F^l - \frac{2l_i+1}{2l+1}\langle l_i, l_i; 0, 0|l, 0\rangle^2 F^l\right\} \\
&= \frac{1}{4l_i+1}\left\{2(2l_i+1)F^0 - (2l_i+1)\langle l_i, l_i; 0, 0|0, 0\rangle^2 F^0\right. \\
&\quad \left. - \sum_{l>0}\frac{2l_i+1}{2l+1}\langle l_i, l_i; 0, 0|l, 0\rangle^2 F^l\right\} \\
&= F^0 - \frac{2l_i+1}{4l_i+1}\sum_{l>0}\frac{\langle l_i, l_i; 0, 0|l, 0\rangle^2}{2l+1}F^l\,.
\end{aligned} \tag{18.30}$$

Here the result of Exercise (14.15) was used.

To sum up, the average energy of the configuration is equal to

$$E_{\text{av}} = \sum_{i=1}^{Q} E_i^{\text{av}} + \sum_{i<j}^{Q} E_{ij}^{\text{av}} , \tag{18.31}$$

where E_i^{av} is given by (18.28) and E_{ij}^{av} is given by (18.29) or (18.30), according as electrons i and j are or are not equivalent. Values of E_{ij}^{av} for s, p and d shells are given in Table 18.5.

Table 18.5. Values of E_{ij}^{av}

$l_i l_j$	Equivalent electrons	Nonequivalent electrons
ss	F^0	$F^0 - \frac{1}{2}G^0$
sp		$F^0 - \frac{1}{6}G^1$
sd		$F^0 - \frac{1}{10}G^2$
pp	$F^0 - \frac{2}{25}F^2$	$F^0 - \frac{1}{6}G^0 - \frac{1}{15}G^2$
pd		$F^0 - \frac{1}{15}G^1 - \frac{3}{70}G^3$
dd	$F^0 - \frac{2}{63}F^2 - \frac{2}{63}F^4$	$F^0 - \frac{1}{10}G^0 - \frac{1}{35}G^2 - \frac{1}{35}G^4$

18.4 Energy of Atomic Multiplets

With the formulas just derived we can compute the average energy of any electronic configuration, no matter what its complexity. But this is not enough. We want to find the energy of each multiplet. Investigated with unrestricted generality, this problem is extremely complicated. After a few general remarks we will take a closer look at some simple cases.

The vector space associated with an electronic configuration has a basis of Slater determinants. The evaluation of multiplet energies therefore boils down to the computation of matrix elements of the Hamiltonian in that basis. This is not the most efficient way to carry out the calculation, particularly in complex situations. But this shows that multiplet energies can be expressed in terms of integrals $I(n, l)$, F^l and G^l of Sect. 18.3, since the matrix elements are linear combinations of these integrals.

Let us denote by $E(^{2S+1}L)$ the energies of the various multiplets of a given configuration.[9] One can write

---

[9] If more than one multiplet correspond to a pair (L, S), an additional symbol is needed to distinguish energies.

$$E(^{2S+1}L) = E_{\text{av}} + \Delta E(^{2S+1}L). \tag{18.32}$$

One can prove ([59], Chap. 12) that $\Delta E(^{2S+1}L)$ has the following properties:

a) $\Delta E(^{2S+1}L)$ vanishes if the configuration involves only one multiplet.
b) Electrons in closed shells do not contribute to $\Delta E(^{2S+1}L)$.
c) Let a configuration have just one shell that is not completely filled. The expression of $\Delta E(^{2S+1}L)$ in terms of integrals $I(n,l)$, F^l and G^l is the same for this configuration and its conjugate. Numerical values of integrals, however, are different.

Property (a) is obvious. Property (b) is related to the fact that closed shells produce a spherically symmetric field. Hence their interaction with an electron does not depend on its orientation. Property (c) expresses just how the interaction between holes is analogous to the interaction between electrons.

Property (a) entails that $\Delta E(^{2S+1}L)$ vanishes if all electronic shells are filled, if there is just one electron outside closed shells or if there is just one hole in otherwise closed shells. Let us now evaluate $\Delta E(^{2S+1}L)$ in less trivial cases.

i) Two *s* Electrons

Consider two orbitals n_1s and n_2s. Suppose first that $n_1 \neq n_2$. The three quantum states associated with the triplet have the same energy, presumably different from the singlet's. It is always simpler to compute the energy of a wave function made up of just one Slater determinant. Consider the $M_S = 1$ component of the triplet, given in Table 18.1 (p. 418). According to property (b) electrons in closed shells can be neglected. Actually, this just adds a numerical constant to all energies. Energies without that constant will be denoted with a prime. Using formulas (18.23), (18.24) and (18.25) as well as values in Table 18.4 (p. 424), one gets

$$E'(^3S) = I(n_1s) + I(n_2s) + F^0(n_1s;n_2s) - G^0(n_1s;n_2s). \tag{18.33}$$

Let us now turn to the singlet's energy. One way to evaluate it is to compute the mean value of the Hamiltonian in the singlet's wave function. This requires calculating four matrix elements of H in the basis of determinantal wave functions. But there is a simpler method. Consider the two-dimensional vector space generated by the orthonormal vectors $|\Phi(0^-,0^+)\rangle$ and $|\Phi(0^+,0^-)\rangle$. Clearly, atomic states $|^3S, M_S = 0\rangle$ and $|^1S, M_S = 0\rangle$ also make up an orthonormal basis of that space, obtained from the former by a unitary transformation. But we know that the trace of an operator is invariant under such a change of basis. Applying this property to the Hamiltonian operator restricted to the two-dimensional space, one has

$$E'(^3S) + E'(^1S) \\ = \langle \Phi(0^-,0^+)|H|\Phi(0^-,0^+)\rangle + \langle \Phi(0^+,0^-)|H|\Phi(0^+,0^-)\rangle\,. \tag{18.34}$$

We already know the triplet's energy. Thus we will get the singlet's energy by calculating two matrix elements. One easily finds that

$$E'(^3S)+E'(^1S) = 2\left\{I(n_1s) + I(n_2s) + F^0(n_1s;n_2s)\right\},$$

from which one gets

$$E'(^1S) = I(n_1s)+I(n_2s)+F^0(n_1s;n_2s)+G^0(n_1s;n_2s). \tag{18.35}$$

The average of $E'(^3S)$ and $E'(^1S)$, weighed by the number of states in each multiplet, is equal to $\{3E'(^3S) + E'(^1S)\}/4$. Using (18.33) and (18.35) we conclude that

$$\Delta E(^3S) = -\frac{1}{2}G^0(n_1s;n_2s)\,, \tag{18.36}$$

$$\Delta E(^1S) = \frac{3}{2}G^0(n_1s;n_2s)\,. \tag{18.37}$$

Thus the singlet's and the triplet's energies differ by $2G^0(n_1s;n_2s)$. As pointed out in Sect. 18.3 this quantity is always positive, so that the singlet's energy is always higher than the triplet's.

There remains the case where the two s orbitals are equivalent, i.e. $n_1 = n_2$. Only the singlet survives, so that $\Delta E(^1S) = 0$.

ii) One *s* Electron and One *p* Electron

This case is very similar to the one just investigated, except that the orbitals cannot be equivalent. One finds that (Exercise 18.8)

$$\Delta E(^3P) = -\frac{1}{6}G^1(n_1s;n_2p)\,, \tag{18.38}$$

$$\Delta E(^1P) = \frac{1}{2}G^1(n_1s;n_2p)\,. \tag{18.39}$$

iii) Two *p* Electrons

Let us first examine the case where the two p orbitals are different. The energy of the ^3D multiplet is obtained by calculating the mean value of H in wave function $\Phi(1^+,1^+)$. The energy of the ^3P then follows from application of the trace invariance rule to the vector space generated by $\Phi(1^+,0^+)$ and $\Phi(0^+,1^+)$. In a similar way, one finds the energies of the six multiplets involved. The result is (Exercise 18.9)

$$\Delta E(^3D) = \frac{1}{25}F^2(n_1p;n_2p) - \frac{5}{6}G^0(n_1p;n_2p) + \frac{2}{75}G^2(n_1p;n_2p)\,, \quad (18.40)$$

$$\Delta E(^3P) = -\frac{1}{5}F^2(n_1p;n_2p) + \frac{7}{6}G^0(n_1p;n_2p) - \frac{2}{15}G^2(n_1p;n_2p)\,, \quad (18.41)$$

$$\Delta E(^3S) = \frac{2}{5}F^2(n_1p;n_2p) - \frac{5}{6}G^0(n_1p;n_2p) - \frac{1}{3}G^2(n_1p;n_2p)\,, \quad (18.42)$$

$$\Delta E(^1D) = \frac{1}{25}F^2(n_1p,n_2p) + \frac{7}{6}G^0(n_1p;n_2p) + \frac{8}{75}G^2(n_1p;n_2p)\,, \quad (18.43)$$

$$\Delta E(^1P) = -\frac{1}{5}F^2(n_1p;n_2p) - \frac{5}{6}G^0(n_1p;n_2p) + \frac{4}{15}G^2(n_1p;n_2p)\,, \quad (18.44)$$

$$\Delta E(^1S) = \frac{2}{5}F^2(n_1p;n_2p) + \frac{7}{6}G^0(n_1p;n_2p) + \frac{7}{15}G^2(n_1p;n_2p)\,. \quad (18.45)$$

When the two p orbitals are equivalent, one finds the energy of the ^3P and the ^1D multiplet by calculating the mean value of H in $\varPhi(1^+,0^+)$ and $\varPhi(1^+,1^-)$, respectively. The energy of the ^1S is next obtained by applying the trace invariance rule. One gets (Exercise 18.10)

$$\Delta E(^3P) = -\frac{3}{25}F^2(n_1p;n_1p)\,, \quad (18.46)$$

$$\Delta E(^1D) = \frac{3}{25}F^2(n_1p;n_1p)\,, \quad (18.47)$$

$$\Delta E(^1S) = \frac{12}{25}F^2(n_1p;n_1p)\,. \quad (18.48)$$

From this we find that

$$\Delta E(^1S) - \Delta E(^3P) = \frac{3}{5}F^2(n_1p;n_1p)\,, \quad (18.49)$$

$$\frac{\Delta E(^1S) - \Delta E(^1D)}{\Delta E(^1D) - \Delta E(^3P)} = \frac{3}{2}\,. \quad (18.50)$$

The total width of the multiplet, given by (18.49), depends on the function F^2. The ratio of intervals between adjacent levels, however, does not depend on it. The value of 3/2 follows solely from the assumption that wave functions belong to only one configuration.

Table 18.6 compares predictions (18.49) and (18.50) with experimental values [165], [166], for the ground-state configuration of four neutral atoms. Function F^2 is obtained through the central-field model, by means of radial functions computed in Sect. 9.6. Results tend to overestimate the multiplet's total width. The assumption of only one configuration works better with silicium and sulfur than with carbon and oxygen. Better predictions hinge on configuration interaction calculations, which will be mentioned in Sect. 18.7.

Table 18.6. Separation between multiplets in np^2 or np^4 configurations

	$\Delta E(^1S) - \Delta E(^3P)$ [Ry]		$\frac{\Delta E(^1S)-\Delta E(^1D)}{\Delta E(^1D)-\Delta E(^3P)}$	
	Theory	Experiment	Theory	Experiment
C	0.356	0.197	1.50	1.13
O	0.484	0.307	1.50	1.14
Si	0.247	0.139	1.50	1.48
S	0.317	0.200	1.50	1.43

iv) Three *s* Electrons

The energy of the ^4S multiplet is easily obtained by noting that $\Phi(0^+, 0^+, 0^+)$ is the only wave function associated with the value $M_S = 3/2$. The mean value of H in this state can be computed with formulas (18.23), (18.24) and (18.25), and one finds

$$\Delta E(^4S) = -\frac{1}{2}G^0(n_1s; n_2s) - \frac{1}{2}G^0(n_1s; n_3s) - \frac{1}{2}G^0(n_2s; n_3s). \qquad (18.51)$$

Energies of the two ^2S multiplets are somewhat more complicated to find. Since three linearly independent wave functions correspond to $M_L = 0$ and $M_S = 1/2$, the trace invariance rule is not sufficient to determine the two unknown energies. These are obtained by diagonalizing the Hamiltonian in the appropriate two-dimensional space.

The foregoing discussion applies to the case where the three *s* orbitals are different. When two orbitals are equivalent there is only one ^2S multiplet, and then $\Delta E(^2S) = 0$.

v) More Complicated Cases

Whenever there are only two electrons outside closed shells, multiplet energies can be obtained by computing the mean value of H in determinantal wave functions and using the trace invariance rule.

The problem gets more complicated when more than two electrons are found outside closed shells. It is then necessary to diagonalize the Hamiltonian in finite-dimensional subspaces, as in the determination of wave functions.

This completes our discussion of atomic multiplet energies. We should point out, however, that the energy of a multiplet is not directly observable. In fact, experiments measure excitation and ionization energies. But they can easily be obtained from multiplet energies.[10]

---

[10] Strictly speaking, multiplet energies should be corrected by spin–orbit interaction and the Zeeman effect, as we will do in Sects. 18.5 and 18.6. The present discussion applies unchanged to corrected energies.

Let $|\Psi_a\rangle$ and $|\Psi_b\rangle$ be two stationary states of an atom and let $E(^{2S_a+1}L_a)$ and $E(^{2S_b+1}L_b) > E(^{2S_a+1}L_a)$ be the energies of the multiplets to which these states belong. The *excitation energy* from $|\Psi_a\rangle$ to $|\Psi_b\rangle$ is defined as $E(^{2S_b+1}L_b) - E(^{2S_a+1}L_a)$. The most interesting case is when the multiplets belong to different configurations. Rigorously, one should then compute two distinct sets of orbitals, associated with configurations (a) and (b). The energy $E(^{2S_a+1}L_a)$ is obtained by means of formula (18.32), where E_{av} and ΔE are computed from orbitals associated with configuration (a). The energy $E(^{2S_b+1}L_b)$ is similarly obtained from orbitals associated with configuration (b).

Obviously this kind of calculation is much more accurate than the one which identifies the excitation energy to the difference of orbital energies (the λ_i in Chaps. 9 and 17). Indeed it takes into account the rearrangement of unexcited orbitals and the fact that different multiplets in the same configuration have different energies.

Similar remarks apply to the *ionization energy*. Here $|\Psi_a\rangle$ represents an atomic state and $|\Psi_b\rangle$ represents a state of the ion obtained from the atom by removing an electron. Usually $|\Psi_a\rangle$ and $|\Psi_b\rangle$ coincide with the atomic and ionic ground states. At any rate one must compute orbitals associated with atomic configuration (a) and orbitals associated with ionic configuration (b). The ionization energy is equal to $E(^{2S_b+1}L_b) - E(^{2S_a+1}L_a)$, where $E(^{2S_a+1}L_a)$ is the atom's energy and $E(^{2S_b+1}L_b)$ is the ion's energy.

Table 18.7. Total ionization energies of light atoms (Ry)

	Theory	Experiment		Theory	Experiment
He	5.59	5.81			
Li	14.67	14.96	Na	322.6	324.9
Be	28.88	29.34	Mg	398.0	400.6
B	48.71	49.32	Al	482.5	485.5
C	74.93	75.71	Si	576.4	579.7
N	108.25	109.23	P	680.1	683.9
O	148.90	150.22	S	793.5	798.1
F	197.92	199.62	Cl	917.4	922.8
Ne	256.02	258.11	Ar	1 051.9	1 058.2

Table 18.7 displays the absolute value of the ground-state total energy for the 17 lightest multi-electron atoms. This coincides with the energy required to completely ionize each atom. Experimental values are drawn from [59], p. 12. Theoretical values are computed from results of Sects. 18.3 and 18.4

and from radial functions obtained in Sect. 9.6.[11] The relative difference between theoretical and experimental values is lower than 0.04 for helium, lower than 0.02 from lithium to carbon and lower than 0.01 from nitrogen on.

18.5 Spin–Orbit Interaction

We have hitherto considered in the atomic Hamiltonian only kinetic and electrostatic potential energy terms. But we know that any charged particle with nonvanishing angular momentum carries a magnetic moment proportional to its angular momentum. Magnetic moments of various electrons interact and give rise to additional terms in the Hamiltonian. We will view these terms as a perturbation of Hamiltonian (17.35).[12] This approximation is appropriate insofar as their matrix elements are small compared with differences between multiplet energies.

Let us first examine the effect of terms like $\xi_{ij}\boldsymbol{L}_i\cdot\boldsymbol{L}_j$. Here $\boldsymbol{L}_i$ and $\boldsymbol{L}_j$ are orbital angular momenta of two electrons and ξ_{ij} is a function of r_i, r_j and r_{ij}. Obviously $\xi_{ij}\boldsymbol{L}_i\cdot\boldsymbol{L}_j$ commutes with $\boldsymbol{L}$, the total orbital angular momentum. The only effect of this term is therefore to add the same constant to the energy of all states in an atomic multiplet. It does not remove degeneracy within a multiplet. So is it with $\xi_{ij}\boldsymbol{S}_i\cdot\boldsymbol{S}_j$. Such terms will be neglected.

Much more interesting is the interaction energy between the spin and orbital magnetic moments of the same electron. We saw in (15.84) that it is given by Dirac's relativistic equation as

$$\xi(r_i)\boldsymbol{L}_i\cdot\boldsymbol{S}_i=\frac{1}{2m^2c^2}\frac{1}{r_i}\frac{\mathrm{d}V}{\mathrm{d}r_i}\boldsymbol{L}_i\cdot\boldsymbol{S}_i\,. \tag{18.52}$$

Here $V=V(r)$ is the central potential in the atom. The term (18.52) is called the *spin–orbit interaction* of electron i. The total spin–orbit interaction is given by[13]

$$H_{LS}=\sum_{i=1}^{Q}\xi(r_i)\boldsymbol{L}_i\cdot\boldsymbol{S}_i\,. \tag{18.53}$$

Note that adding H_{LS} to the Hamiltonian does not change the average energy of a configuration. Indeed let $\{|\Phi_a\rangle\}$ be a basis of Slater determinants in the vector space associated with the configuration. From (17.33) one has

---

[11] The separation between multiplets of an np^3 configuration (nitrogen and phosphorous) is obtained in [233], Sect. 21.3 and in [59], Sect. 12.8.

[12] In this section and the next one, the Hamiltonian (17.35) will be denoted by H_0.

[13] Terms $\boldsymbol{L}_i\cdot\boldsymbol{S}_j$ $(i\neq j)$, less important than the ones just mentioned, are neglected.

$$\begin{aligned}\sum_a \langle \Phi_a | H_{LS} | \Phi_a \rangle &= \sum_a \sum_i \langle \phi_{ai} | \xi(r_i) \boldsymbol{L}_i \cdot \boldsymbol{S}_i | \phi_{ai} \rangle \\ &= \sum_a \sum_i \langle u_{ai} | \xi(r_i) L_{iz} | u_{ai} \rangle \langle \chi_{ai} | S_{iz} | \chi_{ai} \rangle \,. \qquad (18.54)\end{aligned}$$

If $|\Phi_a\rangle$ belongs to the configuration, then the determinant $|\Phi'_a\rangle$ obtained from $|\Phi_a\rangle$ by keeping the same spatial functions and inverting all spins also belongs to the configuration. Hence terms in the sum pairwise cancel.

We now introduce operators $\boldsymbol{J}_i$ and $\boldsymbol{J}$, the total angular momentum of electron i and the total electronic angular momentum. One has

$$\boldsymbol{J}_i = \boldsymbol{L}_i + \boldsymbol{S}_i \,, \qquad \boldsymbol{J} = \sum_{i=1}^{Q} \boldsymbol{J}_i \,. \qquad (18.55)$$

The following commutation relations are easily checked:

$$\begin{aligned}&[\xi(r_i)\boldsymbol{L}_i \cdot \boldsymbol{S}_i, \boldsymbol{L}_i] \neq 0 \,, \qquad [H_{LS}, \boldsymbol{L}] \neq 0 \,, \qquad (18.56)\\ &[\xi(r_i)\boldsymbol{L}_i \cdot \boldsymbol{S}_i, \boldsymbol{S}_i] \neq 0 \,, \qquad [H_{LS}, \boldsymbol{S}] \neq 0 \,, \qquad (18.57)\\ &[\xi(r_i)\boldsymbol{L}_i \cdot \boldsymbol{S}_i, \boldsymbol{J}_i] = 0 \,, \qquad [H_{LS}, \boldsymbol{J}] = 0 \,. \qquad (18.58)\end{aligned}$$

Thus in the state space associated with an atomic multiplet, the spin–orbit interaction is not diagonal. It does, however, commute with $\boldsymbol{J}$. This means that the Hamiltonian, with the spin–orbit interaction, can be diagonalized simultaneously with $\boldsymbol{J}^2$ and J_z.

For various multiplets we have built eigenvectors of $\boldsymbol{L}^2$, $\boldsymbol{S}^2$, L_z and S_z. In general these vectors can be denoted by $|\gamma; L, S; M_L, M_S\rangle$. Here γ stands for all other indices necessary to the full specification of a state. Within a multiplet, γ, L and S are fixed while M_L and M_S vary. We now prove the following lemma.

Lemma

$$\begin{aligned}&\langle \gamma; L, S; M_L, M_S | \sum_{i=1}^{Q} \xi(r_i) \boldsymbol{L}_i \cdot \boldsymbol{S}_i | \gamma; L, S; M'_L, M'_S \rangle \\ &\quad = \Gamma(\gamma; L, S) \langle \gamma; L, S; M_L, M_S | \boldsymbol{L} \cdot \boldsymbol{S} | \gamma; L, S; M'_L, M'_S \rangle \,, \qquad (18.59)\end{aligned}$$

where Γ only depends on the arguments shown.

Proof As in Exercise (14.18) one can write

$$\boldsymbol{L}_i \cdot \boldsymbol{S}_i = \sum_{q=-1}^{1} (-1)^q L_{iq} S_{i,-q} \,, \qquad (18.60)$$

where L_{iq} denotes components (14.89) of orbital angular momentum. Consider the expression

$$\langle\gamma; L, S; M_L, M_S|\xi(r_i)L_{iq}S_{ip}|\gamma; L, S; M'_L, M'_S\rangle. \tag{18.61}$$

Clearly $\xi(r_i)L_{iq}$ is an irreducible tensor operator of rank 1 with respect to $\boldsymbol{L}$. Furthermore, $S_{ip}|\gamma; L, S; M'_L, M'_S\rangle$ is an eigenvector of $\boldsymbol{L}^2$ and L_z. The Wigner–Eckart theorem therefore implies that the M_L, M'_L and q dependence of (18.61) is entirely contained in the Clebsch–Gordan coefficient $\langle 1, L; q, M'_L|L, M_L\rangle$. Likewise the M_S, M'_S and p dependence of (18.61) is entirely contained in the Clebsch–Gordan coefficient $\langle 1, S; p, M'_S|S; M_S\rangle$. A completely similar argument shows that the matrix element

$$\langle\gamma; L, S; M_L, M_S|L_q S_p|\gamma; L, S; M'_L, M'_S\rangle$$

has the same dependence. Thus we can conclude that

$$\begin{aligned}&\langle\gamma; L, S; M_L, M_S|\xi(r_i)L_{iq}S_{ip}|\gamma; L, S; M'_L, M'_S\rangle\\&\quad= \alpha_i(\gamma; L, S)\langle\gamma; L, S; M_L, M_S|L_q S_p|\gamma; L, S; M'_L, M'_S\rangle\,.\end{aligned}$$

Using (18.60) one finds that

$$\begin{aligned}&\langle\gamma; L, S; M_L, M_S|\xi(r_i)\boldsymbol{L}_i\cdot\boldsymbol{S}_i|\gamma; L, S; M'_L, M'_S\rangle\\&\quad= \alpha_i(\gamma; L, S)\langle\gamma; L, S; M_L, M_S|\boldsymbol{L}\cdot\boldsymbol{S}|\gamma; L, S; M'_L, M'_S\rangle\,.\end{aligned}$$

Summing over the index i, we get the result. ♣

At this stage it is convenient to switch to the basis $\{|\gamma; L, S; J, M\rangle\}$ of eigenvectors of $\boldsymbol{L}^2$, $\boldsymbol{S}^2$, $\boldsymbol{J}^2$ and J_z. These vectors are, of course, related to the $|\gamma; L, S; M_L, M_S\rangle$ by Clebsch–Gordan coefficients. Operator H_{LS} commutes with $\boldsymbol{J}^2$ and J_z. Following the lemma on p. 25, H_{LS}, restricted to the state space associated with the multiplet $(\gamma; L, S)$, is therefore diagonal in the basis $\{|\gamma; L, S; J, M\rangle\}$. Let us apply stationary perturbation theory for the case of degenerate energy. Vectors $|\gamma; L, S; J, M\rangle$ are, in zeroth order, eigenvectors of the perturbed Hamiltonian. In first order, the energy correction is simply given by

$$\varepsilon_{1,J} = \langle\gamma; L, S; J, M|H_{LS}|\gamma; L, S; J, M\rangle. \tag{18.62}$$

Since H_{LS} commutes with $\boldsymbol{J}$, we already know that the energy correction does not depend on M. According to (18.59), operators H_{LS} and $\Gamma\boldsymbol{L}\cdot\boldsymbol{S}$ coincide in the state space associated with the multiplet (γ, L, S). Hence

$$\begin{aligned}&\langle\gamma; L, S; J, M|H_{LS}|\gamma; L, S; J, M\rangle\\&\quad= \Gamma(\gamma; L, S)\langle\gamma; L, S; J, M|\boldsymbol{L}\cdot\boldsymbol{S}|\gamma; L, S; J, M\rangle\,. \end{aligned}\tag{18.63}$$

From the identity

$$\boldsymbol{L}\cdot\boldsymbol{S} = \frac{1}{2}\{\boldsymbol{J}\cdot\boldsymbol{J} - \boldsymbol{L}\cdot\boldsymbol{L} - \boldsymbol{S}\cdot\boldsymbol{S}\}\,, \tag{18.64}$$

one finds for the energy correction

$$\begin{aligned}\varepsilon_{1,J} &= \langle\gamma; L, S; J, M|H_{LS}|\gamma; L, S; J, M\rangle\\ &= \Gamma(\gamma; L, S)\langle\gamma; L, S; J, M|\frac{1}{2}\left\{\boldsymbol{J}\cdot\boldsymbol{J} - \boldsymbol{L}\cdot\boldsymbol{L} - \boldsymbol{S}\cdot\boldsymbol{S}\right\}|\gamma; L, S; J, M\rangle\\ &= \frac{\hbar^2}{2}\Gamma(\gamma; L, S)\left\{J(J+1) - L(L+1) - S(S+1)\right\}. \end{aligned} \tag{18.65}$$

Thus the spin–orbit interaction partly removes the degeneracy within a multiplet. States associated with different values of J have different energies. By a *term* is meant the set of all states which, within a multiplet, have the same value of J.[14] The notation $^{2S+1}L_J$ is used to represent different terms in a multiplet ^{2S+1}L. Recall that within a term, degeneracy with respect to M remains.

To find the numerical value of corrections $\varepsilon_{1,J}$, it is necessary to evaluate the constant $\Gamma(\gamma; L, S)$. Vectors $|\gamma; L, S; J, M\rangle$ are linear combinations of Slater determinants. The mean value of H_{LS} in state $|\gamma; L, S; J, M\rangle$ can thus be evaluated by means of (17.33). From this, constant $\Gamma(\gamma; L, S)$ follows. References [25], Chap. 8 and [233], Sect. 21.4 develop different methods to carry out the computation.

When there is only one electron outside closed shells, all computations are easily performed explicitly. Vectors $|\gamma; L, S; M_L, M_S\rangle$ are then each given by just one Slater determinant. Moreover, two states in the same multiplet differ only by the values of m_l and m_s associated with the outer orbital. Call this orbital Q. One shows in Exercise (18.11) that in every case

$$\begin{aligned}&\langle\gamma; L, S; M_L, M_S|H_{LS}|\gamma; L, S; M'_L, M'_S\rangle\\ &\quad= \langle n_Q, l_Q, m_Q, m_{s_Q}|\xi(r_Q)\boldsymbol{L}_Q\cdot\boldsymbol{S}_Q|n_Q, l_Q, m'_Q, m'_{s_Q}\rangle. \end{aligned} \tag{18.66}$$

From this one gets

$$\begin{aligned}\varepsilon_{1,J} &= \langle\gamma; L, S; J, M|H_{LS}|\gamma; L, S; J, M\rangle\\ &= \sum_{M_L, M_S, M'_L, M'_S} \langle L, S; M_L, M_S|J, M\rangle\langle L, S; M'_L, M'_S|J, M\rangle\\ &\qquad\times\langle\gamma; L, S; M_L, M_S|H_{LS}|\gamma; L, S; M'_L, M'_S\rangle\\ &= \sum_{m_l, m_s, m'_l, m'_s} \langle l, s; m_l, m_s|j, m\rangle\langle l, s; m'_l, m'_s|j, m\rangle\\ &\qquad\times\langle n, l, m_l, m_s|\xi(r_Q)\boldsymbol{L}_Q\cdot\boldsymbol{S}_Q|n, l, m'_l, m'_s\rangle\end{aligned}$$

---

[14] The terminology is not uniform. Weissbluth [233] and Cowan [59], for instance, call "term" what we call "multiplet" and "level" what we call "term". Of course this sense of the word is completely different from the one that refers to various elements of a sum, as when one talks about "terms in the Hamiltonian."

$$= \langle n, l, j, m | \frac{1}{2} \xi(r_Q) \{ \boldsymbol{J}_Q \cdot \boldsymbol{J}_Q - \boldsymbol{L}_Q \cdot \boldsymbol{L}_Q - \boldsymbol{S}_Q \cdot \boldsymbol{S}_Q \} | n, l, j, m \rangle$$

$$= \left(\frac{\hbar}{2mc} \right)^2 \left\{ j(j+1) - l(l+1) - \frac{3}{4} \right\} \int_0^\infty \mathrm{d}r \, r \frac{\mathrm{d}V}{\mathrm{d}r} \left[u_{nl}(r) \right]^2 . \quad (18.67)$$

In this equation quantum numbers written with lower case letters refer to orbital Q.

We emphasize that all results found in this section are valid when the spin–orbit interaction energy is small compared with the difference between multiplet energies. This assumption is correct for small atoms but it starts deteriorating as atoms get bigger.

What happens when the spin–orbit interaction is comparable to the difference between multiplet energies in the same configuration? The search for the Hamiltonian's eigenvalues and eigenfunctions then requires diagonalizing $H = H_0 + H_{LS}$ in the space of all states of a given configuration. This calculation is carried out in detail in [59], Chaps. 10 and 12.

The case where the spin–orbit interaction is much larger than the difference between multiplet energies gives rise to *jj coupling*. It is useful then to build basis functions by first coupling each angular momentum $\boldsymbol{L}_i$ with the corresponding spin $\boldsymbol{S}_i$, thereby obtaining an angular momentum $\boldsymbol{J}_i$. Operators $\boldsymbol{J}_i$ are then used to obtain a total angular momentum $\boldsymbol{J}$. Each operator $\boldsymbol{L}_i \cdot \boldsymbol{S}_i$ is diagonal in these basis functions. The Hamiltonian's electrostatic terms are afterwards diagonalized. Such calculations are described, for example, in [56], Chap. 10.

18.6 The Zeeman Effect

Even when the spin–orbit interaction is taken into account, an atom's energy remains degenerate with respect to M. We will see that the degeneracy is removed by the presence of a magnetic field.

Let $\boldsymbol{B}$ be a constant magnetic field. In (7.136) we saw that interaction between the field and the orbital magnetic moment of an electron (with charge q_e and angular momentum $\boldsymbol{L}_i$) results in the Hamiltonian acquiring an additional term given by $(-q_e/2m)\boldsymbol{B} \cdot \boldsymbol{L}_i$.[15] As far as interaction between field and spin magnetic moment is concerned, the corresponding term is given by Dirac's equation as $(-q_e/m)\boldsymbol{B} \cdot \boldsymbol{S}_i$ (Sect. 15.5). Hence the Hamiltonian of an atom with Q electrons acquires an additional term H_{mag} given by

$$H_{\text{mag}} = \sum_{i=1}^{Q} \left\{ -\frac{q_e}{2m} \boldsymbol{B} \cdot \boldsymbol{L}_i - \frac{q_e}{m} \boldsymbol{B} \cdot \boldsymbol{S}_i \right\}$$

---

[15] There is also a term proportional to the square of the magnetic field, which is neglected.

$$= -\frac{q_e}{2m} \boldsymbol{B} \cdot (\boldsymbol{L} + 2\boldsymbol{S}) = -\frac{q_e}{2m} B \, (J_z + S_z) \, . \tag{18.68}$$

Here the z axis is taken as the field's direction. One easily checks (Exercise 18.12) that H_{mag} does not change the average energy of a configuration.

In this section we assume that LS coupling conditions apply, i.e. the spin–orbit interaction is much smaller than the energy difference between multiplets. Furthermore, we assume initially that magnetic terms are much smaller than the spin–orbit interaction.

With these assumptions, the magnetic terms can be considered as a perturbation of the Hamiltonian $H_0 + H_{LS}$, where H_0 is Hamiltonian (17.35). Let us pick the $|\gamma; L, S; J, M\rangle$ as unperturbed vectors. Since S_z and J_z commute, the lemma on p. 25 implies that matrix elements of S_z vanish in states with different M. In first-order perturbation theory the correction to energies is therefore given by

$$\varepsilon_{1,M} = \langle \gamma; L, S; J, M | \frac{-q_e}{2m} B \, (J_z + S_z) \, |\gamma; L, S; L, M\rangle \, . \tag{18.69}$$

Obviously $\boldsymbol{J} + \boldsymbol{S}$ is an irreducible tensor operator of rank 1 with respect to $\boldsymbol{J}$. From the Wigner–Eckart theorem one then gets

$$\begin{aligned} \varepsilon_{1,M} &= -\frac{q_e B}{2m} \langle \gamma; L, S; J, M | \, (J_z + S_z) \, |\gamma; L, S; J, M\rangle \\ &= -\frac{q_e B}{2m} g \langle \gamma; L, S; J, M | J_z | \gamma; L, S; J, M\rangle \\ &= -\frac{q_e B}{2m} g M \hbar \, . \end{aligned} \tag{18.70}$$

Here g is essentially a reduced matrix element, that we will presently evaluate.

For this purpose let us compute in two different ways the mean value of operator $(\boldsymbol{J} + \boldsymbol{S}) \cdot \boldsymbol{J}$ in state $|\gamma; L, S; J, M\rangle$. On the one hand,

$$\begin{aligned} &\langle \gamma; L, S; J, M | \, (\boldsymbol{J} + \boldsymbol{S}) \cdot \boldsymbol{J} | \gamma; L, S; J, M\rangle \\ &\quad = \langle \gamma; L, S; J, M | \frac{1}{2} \left\{ 3\boldsymbol{J} \cdot \boldsymbol{J} - \boldsymbol{L} \cdot \boldsymbol{L} + \boldsymbol{S} \cdot \boldsymbol{S} \right\} |\gamma; L, S; J, M\rangle \\ &\quad = \frac{\hbar^2}{2} \left\{ 3J(J+1) - L(L+1) + S(S+1) \right\} . \end{aligned}$$

On the other hand, writing the scalar product as in (18.60) and introducing a complete set of vectors $|\gamma'; L', S', J', M'\rangle$, one finds that

$$\begin{aligned} &\langle \gamma; L, S; J, M | \sum_{q=-1}^{1} (-1)^q \, (J_q + S_q) \, J_{-q} | \gamma; L, S; J, M\rangle \\ &\quad = \sum_{q,M'} (-1)^q \langle \gamma; L, S; J, M | \, (J_q + S_q) \, |\gamma; L, S; J, M'\rangle \\ &\qquad \times \langle \gamma; L, S; J, M' | J_{-q} | \gamma; L, S; J, M\rangle \end{aligned}$$

$$
\begin{aligned}
&= \sum_{q,M'} (-1)^q g \langle \gamma; L, S; J, M | J_q | \gamma; L, S; J, M' \rangle \\
&\qquad \times \langle \gamma; L, S; J, M' | J_{-q} | \gamma; L, S; J, M \rangle \\
&= g \langle \gamma; L, S; J, M | \sum_{q=-1}^{1} (-1)^q J_q J_{-q} | \gamma; L, S; J, M \rangle \\
&= g\hbar^2 J(J+1) \,.
\end{aligned}
$$

Comparing both results we get

$$
g(L, S, J) = \frac{3J(J+1) - L(L+1) + S(S+1)}{2J(J+1)} \,. \tag{18.71}
$$

The constant g, whose dependence on L, S and J is displayed here, is called the *Landé factor.*

To sum up, the correction $\varepsilon_{1,M}$ of energies of atomic terms, due to a magnetic field, can be written as

$$
\varepsilon_{1,M} = -\frac{q_e \hbar}{2m} g(L, S, J) B M = \mu_B g B M \,. \tag{18.72}
$$

Within an atomic term the energy of an M state acquires a correction proportional to M and proportional to the magnetic field's intensity. This is the so-called *Zeeman effect*, for weak fields.

What happens when the Hamiltonian's magnetic terms are of the same order of magnitude as the spin–orbit interaction? Then H_{mag} can no longer be viewed as a perturbation of $H_0 + H_{LS}$. In pure LS coupling, however, $H_{LS} + H_{\mathrm{mag}}$ can be taken as a perturbation of H_0. The problem then consists in diagonalizing the operator $H_{LS} + H_{\mathrm{mag}}$ in the state space associated with a multiplet. Exercise (18.13) gives an example of such a calculation, for an alkali atom.

In very strong magnetic fields $H_0 + H_{\mathrm{mag}}$ is identified with the unperturbed Hamiltonian and H_{LS} with the perturbation. Then $H_0 + H_{\mathrm{mag}}$ has to be diagonalized in the space of all states of a given configuration. Some of these situations are analyzed in [56], Chap. 16, among others.

To close this section we display in Fig. 18.1 the 15 quantum states associated with a configuration of two nonequivalent p electrons. In an ideal case the three multiplets' energies are shifted from the average energy according to (18.46)–(18.48). The spin–orbit interaction splits the ${}^3P$ multiplet into three terms, following (18.65). One has $E({}^3P_2) - E({}^3P_1) = 2\{E({}^3P_1) - E({}^3P_0)\}$. Finally, terms are splitted into individual states by the Zeeman effect, following (18.71) and (18.72). The separation between adjacent Zeeman levels is the same in ${}^3P_2$ and in ${}^3P_1$ terms, and reduced by one third in ${}^1D_2$.

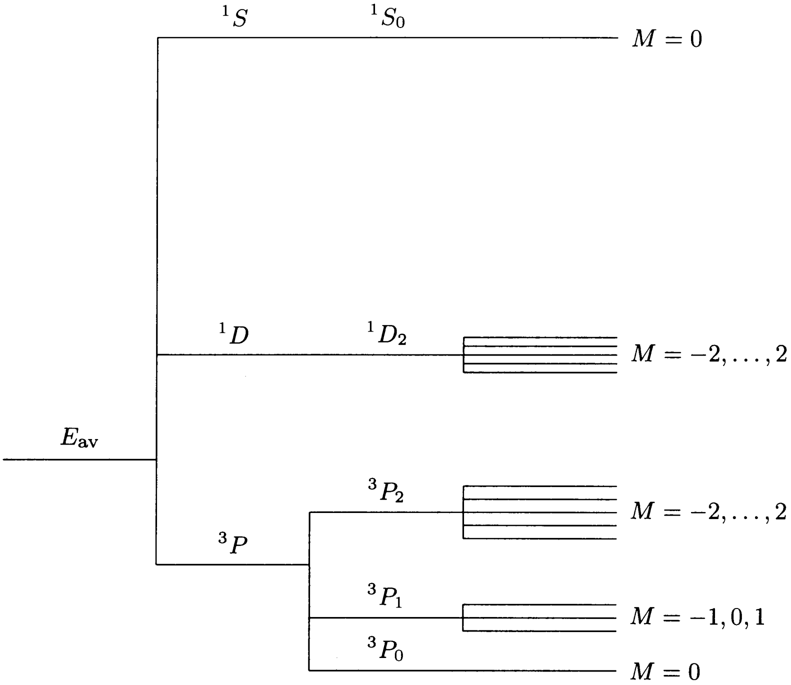

Fig. 18.1. The 15 states associated with a configuration of two nonequivalent p electrons

18.7 Configuration Interaction

In this last section we briefly come back to the computation of atomic energies and wave functions and indicate how it can be improved.

The Hamiltonian of a neutral atom or atomic ion is given by (17.35), to which the term H_{LS} given in (18.53) must be added. With an external magnetic field the term H_{mag} in (18.68) must also be added. We have considered $H_{LS} + H_{\mathrm{mag}}$ as a perturbation of operator (17.35), an assumption which can be dropped if unsuitable.

The state space of an atom has infinite dimension. Nevertheless, the Hamiltonian has hitherto been diagonalized in a finite-dimensional subspace corresponding to a specific configuration. Slater determinants associated with this configuration are built from electronic wave functions (18.1). The latter are computed by one form or another of the Hartree–Fock equations, for instance (9.58) and (9.59).

The diagonalization of the Hamiltonian in the space of a specific configuration yields, in principle, an approximation to the energy and wave function of any atomic state. As a matter of fact, computations quickly become very complicated and the use of Racah's formalism is inevitable.

Although it was not mentioned in this or the previous chapter, it is clear that accurate atomic calculations must take relativistic effects into account [95]. As Z increases they become more important. A relativistic treatment of the hydrogen atom requires replacing Schrödinger's equation by Dirac's equation. Similarly, a relativistic treatment of an atom with atomic number Z requires replacing the Hartree–Fock equations by the Hartree–Fock–Dirac equations, the latter being a four-component extension of the former. This makes the problem significantly more complicated. But there exists a simpler approach that is appropriate to most cases. To the Hartree–Fock equations, one adds terms like (15.83) obtained from Dirac's equation when energy is not too high. Reference [59], Sects. 7.14 and 7.17, explains how to deal with the singularity of these terms at the origin.

At any rate, even if the Hartree–Fock equations could by solved exactly and relativistic effects were taken into account, the wave functions and energies so obtained would not precisely coincide with the Hamiltonian's eigenfunctions and eigenvalues. The *correlation energy* is the difference between the experimentally measured energy of an atom and the energy predicted by the Hartree–Fock equations (including relativistic corrections). The correlation energy is usually smaller than one percent of the total energy, but it is not zero.

The origin of the correlation energy can be seen in the particular case of the helium atom's ground state. The solution of the (restricted) Hartree–Fock equations has the form

$$\Phi(1,2) = u_{1s}(r_1)u_{1s}(r_2)\frac{1}{\sqrt{2}}\left\{\chi_+(1)\chi_-(2) - \chi_+(2)\chi_-(1)\right\}.$$

Functions u_{1s} do not depend on the polar and azimuthal angles of $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$. Thus in the Hartree–Fock approximation the spatial part of $\Phi(1,2)$ separates into a product of functions of r_1 and of r_2. But in fact nothing prevents $\Phi(1,2)$ to depend on distance r_{12}. Indeed one can build eigenfunctions of $\boldsymbol{L}^2$ that are symmetric with respect to an exchange of spatial coordinates and that depend on r_{12}. Such a dependence cannot be obtained from the Hartree–Fock equations, but it is present in the real world. It precisely introduces a correlation of electronic positions.

The accurate computation of atomic wave functions and energies, hence of correlation energies, requires what is called *configuration interaction*. The Hamiltonian is diagonalized not in a vector space associated with only one configuration, but rather in a larger space associated with several configurations. Since, however, the space inversion operator commutes with the Hamiltonian and since all wave functions in a configuration have the same parity, it is enough to consider configurations with the same parity.

The computation of matrix elements of the Hamiltonian in a basis of wave functions belonging to different configurations can become very complicated. It is explained in detail in [59], Chap. 13. Note that matrix elements vanish if configurations differ by more than two electronic wave functions. By Brillouin's theorem, they also vanish if wave functions are each made up of just one Slater determinant, and differ by only one electronic wave function.

We should point out that the Hartree–Fock equations can be generalized to wave functions belonging to several configurations. Such computations are described in detail in [86] and [87], where one can find a number of hints on the proper way to pick configurations.

Exercises

18.1. Let $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ be the position operators of electrons (1) and (2), and $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$ the orbital angular momentum operators.
a) Show that

$$\left[(\boldsymbol{R}_1 - \boldsymbol{R}_2)^2, \boldsymbol{L}_1\right] \neq 0, \qquad \left[(\boldsymbol{R}_1 - \boldsymbol{R}_2)^2, \boldsymbol{L}_1 + \boldsymbol{L}_2\right] = 0.$$

b) From this derive (18.4) and (18.5).

18.2.
a) Find all multiplets coming from the coupling of two nonequivalent d electrons.
b) Find all multiplets coming from the coupling of an s electron, a d electron and an f electron.

18.3. Find the wave functions of a ${}^3D$ multiplet of two p electrons, given in Table 18.2.

18.4. Find the wave functions of a ${}^3P$ multiplet of two p electrons, given in Table 18.2.

18.5. Find the wave functions of the ${}^3S$, ${}^1D$, ${}^1P$ and ${}^1S$ multiplets of two p electrons, given in Table 18.2.

18.6. Derive (18.17) and (18.18).

18.7. Suppose that in (18.10) the Φ_a are products of electronic wave functions [as in (9.25)] rather than Slater determinants. Suppose furthermore that the Pauli exclusion principle remains valid. Show that the average energy of a configuration is still given by (18.31), with E_i^{av} unchanged but with E_{ij}^{av} given in every case by F^0.

18.8. Derive (18.38) and (18.39).

18.9. Derive (18.40)–(18.45).

18.10. Derive (18.46)–(18.48).

18.11. Derive (18.66) in the case where wave functions ϕ_Q and ϕ'_Q are identical. In the case where $\phi_Q \neq \phi'_Q$, (18.66) is obvious.

18.12. Show that adding H_{mag} to an atom's Hamiltonian does not change the average energy of a configuration.

18.13. Consider a ^2P multiplet of an alkali atom. Suppose that states $|\gamma; L, S; M_L, M_S\rangle$ are represented by only one Slater determinant. If the magnetic field vanishes and the spin–orbit interaction is negligible, there are in this multiplet six states of energy E_0. Suppose now that the spin–orbit interaction is taken into account and a magnetic field is applied, so that the following terms are added to the Hamiltonian:

$$H_{LS} + H_{\mathrm{mag}} = \Gamma \boldsymbol{L} \cdot \boldsymbol{S} + \mu_B B \hbar^{-1} (J_z + S_z) .$$

Use stationary perturbation theory to find the Hamiltonian's eigenvalues in first order and the eigenvectors in zeroth order. Here one is not allowed to assume that one of the two added terms is small with respect to the other.

19 Semiclassical Radiation Theory

In previous chapters we have shown how to obtain atomic energies and wave functions. Here we examine the interaction of a quantum system with an electromagnetic wave. Indeed the experimental investigation of atoms is carried out largely by spectroscopy, i.e. by recording the properties of radiation that they emit or absorb. In most cases electromagnetic waves can be considered as small perturbations of the atomic system. We shall see how to compute the probability that a photon is emitted or absorbed by an atom during a transition between different states. In particular, we will obtain selection rules associated with the most important transitions.

19.1 Harmonic Perturbation

Let H_0 be a time-independent Hermitian operator. Denote by E_k and $|E_k\rangle$ the eigenvalues and orthonormal eigenvectors of this operator. In general the spectrum of H_0 can be partly discrete and partly continuous. We assume the index k has been chosen appropriately.

In this chapter we will investigate quantum systems whose Hamiltonian is given by

$$H(t) = H_0 + \bar{H}(t) . \tag{19.1}$$

Here $\bar{H}(t)$ explicitly depends on time. Of course this decomposition is significant only if the operator $\bar{H}(t)$ is specified further. So we make the following three assumptions:

i) The state space associated with operator $H(t)$ does not depend on time and coincides with the state space associated with H_0.
ii) Matrix elements of $\bar{H}(t)$ in the basis of eigenvectors of H_0 are small compared with differences between eigenvalues E_k. In this sense we will say that $\bar{H}(t)$ is small.
iii) When $\bar{H}(t)$ does not vanish, its time dependence is harmonic, i.e. it is given by a trigonometric function or a superposition of trigonometric functions.

L. Marchildon, *Quantum Mechanics*

Note that assumption (ii) looks like the one made on operator W in Chap. 10. Once again we call H_0 the *unperturbed Hamiltonian*, $H(t)$ the *perturbed Hamiltonian* and $\bar{H}(t)$ the *perturbation.*

These assumptions, as one may suspect, are made with the purpose of studying the interaction of atoms or molecules with an electromagnetic wave. We will soon come to this. But before, it is instructive to examine a more general case where $\bar{H}(t)$ is essentially arbitrary, except that assumption (i) is satisfied.

Let $|\psi(t)\rangle$ be the state vector of a quantum system governed by Hamiltonian $H(t)$. From assumption (i) one concludes that at any time, $|\psi(t)\rangle$ is a linear combination of vectors $|E_k\rangle$. Generally, this linear combination can be written as

$$|\psi(t)\rangle = \mathsf{S}_k C_k(t) \mathrm{e}^{-\mathrm{i}E_k t/\hbar} |E_k\rangle . \tag{19.2}$$

Coefficients $C_k(t)$ become constants if $\bar{H}$ vanishes.

We now derive a system of differential equations for the $C_k(t)$. The Schrödinger equation is

$$\mathrm{i}\hbar \frac{\mathrm{d}}{\mathrm{d}t} |\psi(t)\rangle = \left\{ H_0 + \bar{H}(t) \right\} |\psi(t)\rangle . \tag{19.3}$$

Substituting (19.2) into (19.3) and using the fact that $H_0|E_k\rangle = E_k|E_k\rangle$, one easily obtains that

$$\mathrm{i}\hbar \mathsf{S}_k \left(\frac{\mathrm{d}C_k}{\mathrm{d}t} \right) \mathrm{e}^{-\mathrm{i}E_k t/\hbar} |E_k\rangle = \mathsf{S}_k C_k(t) \mathrm{e}^{-\mathrm{i}E_k t/\hbar} \bar{H}(t) |E_k\rangle . \tag{19.4}$$

Let us project both sides of (19.4) on the bra $\langle E_l|$. Making use of orthonormality relations one finds that

$$\frac{\mathrm{d}}{\mathrm{d}t} C_l(t) = \frac{1}{\mathrm{i}\hbar} \mathsf{S}_k C_k(t) \mathrm{e}^{\mathrm{i}\omega_{lk} t} \langle E_l | \bar{H}(t) | E_k \rangle , \tag{19.5}$$

where

$$\omega_{lk} = \frac{E_l - E_k}{\hbar} . \tag{19.6}$$

Note that, inasmuch as assumption (i) is satisfied, the system of differential equations (19.5) for the coefficients $C_l(t)$ is equivalent to the Schrödinger equation.

Before introducing assumption (ii), it is useful to express in a more general way the result just obtained. Suppose the Hamiltonian H depends on a real parameter λ, according to the relation

$$H(\lambda; t) = H_0 + \lambda \bar{H}(t) . \tag{19.7}$$

Clearly then, coefficients C_l are functions of λ. They satisfy the equation

$$\frac{\mathrm{d}}{\mathrm{d}t} C_l(\lambda; t) = \frac{\lambda}{\mathrm{i}\hbar} \mathsf{S}_k C_k(\lambda; t) \mathrm{e}^{\mathrm{i}\omega_{lk} t} \langle E_l | \bar{H}(t) | E_k \rangle . \tag{19.8}$$

Suppose that in the neighborhood of $\lambda = 0$, the C_k can be expanded in series of powers of λ, that is,

$$C_k(\lambda; t) = \sum_{s=0}^{\infty} C_k^{(s)}(t) \lambda^s . \tag{19.9}$$

Substituting (19.9) into (19.8) and equating on both sides coefficients of the same powers of λ, one finds that

$$\frac{\mathrm{d}}{\mathrm{d}t} C_l^{(0)}(t) = 0 , \tag{19.10}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} C_l^{(s+1)}(t) = \frac{1}{\mathrm{i}\hbar} \mathsf{S}_k C_k^{(s)}(t) \mathrm{e}^{\mathrm{i}\omega_{lk} t} \langle E_l | \bar{H}(t) | E_k \rangle . \tag{19.11}$$

Here again, the system of differential and recurrence equations for coefficients $C_l^{(s)}(t)$ is equivalent to Schrödinger's equation. One sees at once that the $C_l^{(0)}(t)$ are constant, while (19.11) can be formally integrated as

$$C_l^{(s+1)}(t) = \frac{1}{\mathrm{i}\hbar} \mathsf{S}_k \int_{-\infty}^{t} \mathrm{d}t' \, C_k^{(s)}(t') \mathrm{e}^{\mathrm{i}\omega_{lk} t'} \langle E_l | \bar{H}(t') | E_k \rangle . \tag{19.12}$$

Let us now introduce assumption (ii). We assume that matrix elements of $\bar{H}$ are small compared with differences between eigenvalues of H_0. Let us denote the matrix element in (19.12) by $\bar{E}$. Due to the oscillating factor, the sign of the integrand changes on a time scale of order $(\omega_{lk})^{-1}$. If we assume that terms corresponding to different values of k have more or less random phases, we find that the right-hand side of (19.12) is of order

$$\frac{1}{\hbar} C_k^{(s)}(t) \frac{1}{\omega_{lk}} \bar{E} ,$$

where k is the index associated with a typical coefficient. Hence if $E_l \neq E_k$,

$$\frac{C_l^{(s+1)}(t)}{C_k^{(s)}(t)} \approx \frac{\bar{E}}{E_l - E_k} \ll 1 . \tag{19.13}$$

Thus assumption (ii) entails that coefficients of order $s+1$ are much smaller than coefficients of order s.

This observation leads to the following approximation. In the expression of $C_k(\lambda; t)$ we will keep coefficients $C_k^{(0)}$ and $C_k^{(1)}(t)$ only. Setting $\lambda = 1$ one finds that the state vector $|\psi(t)\rangle$ of a quantum system governed by Hamiltonian $H_0 + \bar{H}(t)$ is approximately given by

$$|\psi(t)\rangle = \mathsf{S}_k \left\{ C_k^{(0)} + C_k^{(1)}(t) \right\} \mathrm{e}^{-\mathrm{i}E_k t/\hbar} |E_k\rangle . \tag{19.14}$$

In the case where, at $t = -\infty$, $|\psi(t)\rangle$ coincides with an eigenvector $|E_m\rangle$ of H_0, one has $C_k^{(0)} = \delta_{km}$, so that

$$|\psi(t)\rangle = \mathrm{e}^{-\mathrm{i}E_m t/\hbar}|E_m\rangle + \mathrm{S}_k C_k^{(1)}(t)\mathrm{e}^{-\mathrm{i}E_k t/\hbar}|E_k\rangle . \tag{19.15}$$

Coefficients $C_k^{(1)}(t)$ are then given by

$$C_k^{(1)}(t) = \frac{1}{\mathrm{i}\hbar}\int_{-\infty}^{t} \mathrm{d}t'\, \mathrm{e}^{\mathrm{i}\omega_{km}t'}\langle E_k|\bar{H}(t')|E_m\rangle . \tag{19.16}$$

This equation is valid no matter what the perturbation's time dependence.

We now introduce assumption (iii), according to which the time dependence of $\bar{H}(t)$ is harmonic. It is convenient to set the origin of time so that $\bar{H}(t)$ starts acting at $t = 0$. In the simplest case there is only one frequency, so that we can let

$$\bar{H}(t) = \begin{cases} 0 & \text{if } t < 0 , \\ W\mathrm{e}^{-\mathrm{i}\omega t} + W^{\dagger}\mathrm{e}^{\mathrm{i}\omega t} & \text{if } t > 0 . \end{cases} \tag{19.17}$$

Here ω is a positive constant and W an operator. Clearly, for any $t < 0$ the coefficient $C_k^{(1)}(t)$ vanishes. For $t > 0$, $C_k^{(1)}(t)$ is given by

$$\begin{aligned} C_k^{(1)}(t) &= \frac{1}{\mathrm{i}\hbar}\int_0^t \mathrm{d}t'\, \mathrm{e}^{\mathrm{i}\omega_{km}t'} \left\{\langle E_k|W|E_m\rangle \mathrm{e}^{-\mathrm{i}\omega t'} + \langle E_k|W^{\dagger}|E_m\rangle \mathrm{e}^{\mathrm{i}\omega t'}\rangle\right\} \\ &= -\frac{1}{\hbar}\left\{\langle E_k|W|E_m\rangle \frac{\mathrm{e}^{\mathrm{i}(\omega_{km}-\omega)t} - 1}{\omega_{km} - \omega} + \langle E_k|W^{\dagger}|E_m\rangle \frac{\mathrm{e}^{\mathrm{i}(\omega_{km}+\omega)t} - 1}{\omega_{km} + \omega}\right\} . \end{aligned} \tag{19.18}$$

In general, the coefficient $C_k^{(1)}(t)$ is of the order of the ratio of matrix elements of W (i.e. of $\bar{H}$) to eigenvalues of H_0. There are, however, two exceptions to this statement. Indeed the term inside curly brackets can become large if one denominator is close to zero, that is, if

$$E_k - E_m \approx \pm\hbar\omega . \tag{19.19}$$

Therefore one can have, under the effect of a harmonic perturbation, a nonnegligible probability of transition from state $|E_m\rangle$ to state $|E_k\rangle$ if the absolute value of the energy difference is close to the perturbation's angular frequency (multiplied by $\hbar$). When the perturbation is electromagnetic, this transition process corresponds to the absorption or emission of a photon of frequency $\omega/2\pi$.

In Sect. 19.2 we will examine the case where E_m belongs to the discrete and E_k to the continuous spectrum. Later sections will be devoted to the case where E_m and E_k both belong to the discrete spectrum.

19.2 Transition to the Continuous Spectrum

Whenever E_m belongs to the discrete and E_k to the continuous spectrum, $E_k - E_m$ is positive. The second term in the curly brackets of (19.18) is then negligible and one can write

$$C_k^{(1)}(t) = -\frac{1}{\hbar}\langle E_k|W|E_m\rangle \left\{ \frac{\mathrm{e}^{\mathrm{i}(\omega_{km}-\omega)t} - 1}{\omega_{km} - \omega} \right\}. \tag{19.20}$$

This expression is appreciable if k is such that $E_k - E_m \approx \hbar\omega$.

Clearly if E_k belonged to the discrete spectrum, $|C_k^{(1)}(t)|^2$ would represent the probability that at time t the system is in state $|E_k\rangle$.[1] But E_k belongs to the continuous spectrum. Hence quantity $|C_k^{(1)}(t)|^2$ represents a probability density. At any rate

$$\left|C_k^{(1)}(t)\right|^2 = \frac{4|\langle E_k|W|E_m\rangle|^2}{\hbar^2(\omega_{km} - \omega)^2} \sin^2\left\{\frac{1}{2}(\omega_{km} - \omega)t\right\}. \tag{19.21}$$

At this stage it is useful to specify states $|E_k\rangle$ in the continuous spectrum better. Energies in the continuous spectrum are in general highly degenerate. So let us denote states in the continuous spectrum by $|E_i^\alpha\rangle$, where i refers to energy and α labels different states with the same energy. Often the range of the index α is the same for close but different values of energy. In this case it is convenient to write

$$\mathsf{S}_k \to \int \mathrm{d}E_i\, \mathrm{d}\alpha\, \varrho_\alpha(E_i)\,, \tag{19.22}$$

where $\varrho_\alpha(E_i)$ is the Jacobian of the transformation $k \to (E_i, \alpha)$ and represents the density of states of energy E_i and index α.

Define w_α as the probability density, per unit time, that the quantum system makes a transition to a state of index α and arbitrary energy. Thus

$$\begin{aligned} w_\alpha &= \frac{1}{T}\int \mathrm{d}E_i\, \varrho_\alpha(E_i)\left|C_k^{(1)}(T)\right|^2 \\ &= \frac{1}{T}\int \mathrm{d}E_i\, \varrho_\alpha(E_i)\frac{4|\langle E_i^\alpha|W|E_m\rangle|^2}{\hbar^2(\omega_{im} - \omega)^2}\sin^2\left\{\frac{1}{2}(\omega_{im} - \omega)T\right\}. \end{aligned} \tag{19.23}$$

For T large enough the factor

$$f = \frac{1}{(\omega_{im} - \omega)^2}\sin^2\left\{\frac{1}{2}(\omega_{im} - \omega)T\right\}, \tag{19.24}$$

as a function of E_i, has an acute maximum around $E_i = E_m + \hbar\omega$ (Fig. 19.1). Suppose that the density of states $\varrho_\alpha(E_i)$ and the matrix element $\langle E_i^\alpha|W|E_m\rangle$

---

[1] This assumes the probability that the system is in state $|E_m\rangle$ remains close to 1, i.e. $|C_m^{(1)}(t)|^2 \ll 1$. Note that $C_m^{(1)}(t)$ is purely imaginary.

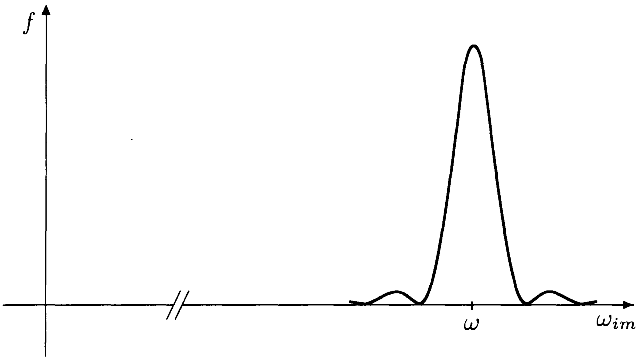

Fig. 19.1. The factor f in (19.24)

do not change much in the interval where this factor significantly differs from zero. One can then write

$$w_\alpha = \frac{4}{\hbar^2 T}\varrho_\alpha(\bar{E}_i)\left|\langle \bar{E}_i^\alpha|W|E_m\rangle\right|^2 \int \mathrm{d}E_i \, \frac{\sin^2\left\{\frac{1}{2}(\omega_{im}-\omega)T\right\}}{(\omega_{im}-\omega)^2}, \tag{19.25}$$

where $\bar{E}_i = E_m + \hbar\omega$. The integral can be evaluated from $-\infty$ to ∞. Using the fact that

$$\int_{-\infty}^{\infty} \mathrm{d}x \, x^{-2}\sin^2 x = \pi, \tag{19.26}$$

one easily finds that

$$w_\alpha = \frac{2\pi}{\hbar}\varrho_\alpha(\bar{E}_i)\left|\langle \bar{E}_i^\alpha|W|E_m\rangle\right|^2. \tag{19.27}$$

The result just obtained is called the *Fermi golden rule.* The probability density, per unit time, that the system makes a transition to a state with index α and arbitrary energy, does not depend on time. Clearly the probability, per unit time, that the system makes a transition to an arbitrary state in the continuous spectrum is given by

$$w = \int \mathrm{d}\alpha \, w_\alpha = \frac{2\pi}{\hbar}\int \mathrm{d}\alpha \, \varrho_\alpha(\bar{E}_i)\left|\langle \bar{E}_i^\alpha|W|E_m\rangle\right|^2. \tag{19.28}$$

Recall that this result holds for weak and harmonic perturbations only.

To illustrate the formalism just developed, we will apply the Fermi golden rule to the ionization of a hydrogen atom.

Consider a hydrogen atom in its ground state which, starting at $t = 0$, is subject to electromagnetic radiation of angular frequency ω. In Sect. 7.7 we saw that the wave's electric and magnetic fields can be obtained from the vector potential

$$\boldsymbol{A}(\boldsymbol{r},t) = 2\boldsymbol{A}_0 \cos(\boldsymbol{q}\cdot\boldsymbol{r} - \omega t + \eta), \tag{19.29}$$

where $\boldsymbol{q} \cdot \boldsymbol{A}_0 = 0$ and $\boldsymbol{q} \cdot \boldsymbol{q} = (\omega/c)^2$. Neglecting the term quadratic in $\boldsymbol{A}$ in (7.134) and checking that $\sum_i [P_i, A_i] = 0$, one sees that the atom's Hamiltonian is modified by the addition of a term

$$\bar{H}(t) = -\frac{q_e}{m}\boldsymbol{A}(\boldsymbol{R},t)\cdot\boldsymbol{P}. \tag{19.30}$$

Here q_e and m are the electron's charge and reduced mass, and $\boldsymbol{P}$ is the momentum operator associated with the relative coordinate.[2]

We will assume that a photon's energy $\hbar\omega$ is larger than the 13.6 eV required to ionize an atom, but that it is not very much larger. Let's say the radiation wavelength is somewhere between 10^2 and 10^3 Å. In this case the scale of spatial variation of the function $\boldsymbol{A}(\boldsymbol{R},t)$ is much larger than atomic dimensions, so that for all practical purposes, $\boldsymbol{A}(\boldsymbol{R},t)$ does not depend on $\boldsymbol{R}$. Thus

$$\bar{H}(t) = -\frac{q_e}{m}\boldsymbol{A}_0 \left\{ \mathrm{e}^{\mathrm{i}\eta - \mathrm{i}\omega t} + \mathrm{e}^{-\mathrm{i}\eta + \mathrm{i}\omega t} \right\}\cdot\boldsymbol{P}, \tag{19.31}$$

from which we get, by comparing with (19.17), that

$$W = -\frac{q_e}{m}\mathrm{e}^{\mathrm{i}\eta}\boldsymbol{A}_0\cdot\boldsymbol{P}. \tag{19.32}$$

One can check (Exercise 19.3) that for an incoming wave of 10^4 W m^{-2}, matrix elements of W are of order 10^{-7} eV, hence much smaller than eigenvalues of H_0.[3]

To apply Fermi's golden rule one must compute the density of states $\varrho_\alpha(E_i)$ and matrix elements $\langle E_i^\alpha | W | E_m \rangle$. For this purpose one should use in principle state vectors in the continuous spectrum of hydrogen. Even though they are known analytically, we will rather use an approximation that will greatly simplify the calculation. The approximation consists in taking the $|E_i^\alpha\rangle$ as free-particle state vectors.[4] These will first be normalized in a cubic box of length $2L$, and the limit $L \to \infty$ will eventually be taken.

---

[2] Strictly speaking, one should write $\bar{H}(t)$ as a sum of terms $\boldsymbol{A}(\boldsymbol{R}_i,t)\cdot\boldsymbol{P}_i$ on the electron and proton coordinates, and then switch to relative coordinates. Up to small corrections, we would recover the result just written. We will come back to this in later sections. The effect of an electromagnetic field on atoms is analyzed in detail in [129].

[3] Neither here nor in later sections will we prove that state spaces associated with H_0 and $H_0 + \lambda\bar{H}$ coincide.

[4] This approximation is valid if the energy of the ejected electron is substantially larger than the absolute value of the ground-state energy of hydrogen. In that case the ejected electron does not feel, so to speak, the Coulomb potential. See [25], Chap. 12 and [26], Chap. 4a, for a more complete discussion of the ionization of hydrogen.

Free-particle normalized wave functions in a box of length $2L$ (with periodic boundary conditions) are given in (9.28) as

$$\psi_{\boldsymbol{k}}(\boldsymbol{r}) \equiv \langle \boldsymbol{r}|E_{\boldsymbol{k}}\rangle = (2L)^{-3/2} \exp\{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}\} .$$

Here

$$\begin{aligned} E_{\boldsymbol{k}} &= \frac{\hbar^2}{2m}\boldsymbol{k}\cdot\boldsymbol{k} , \\ \boldsymbol{k} &= (k_x, k_y, k_z) = \frac{\pi}{L}(n_x, n_y, n_z) , \end{aligned}$$

and n_x, n_y and n_z are integers. Clearly the index i can be identified with $|\boldsymbol{k}|$ while the index α corresponds to the unit vector $\hat{k}$. The differential $\mathrm{d}\alpha$ corresponds to the element of solid angle $\mathrm{d}\Omega_k$, while $\mathrm{d}E_i = (\hbar^2 k/m)\mathrm{d}k$. One shows that the density of states $\varrho_\alpha(E_i)$ is given by[5]

$$\varrho_\alpha(E_i) \equiv \varrho_{\hat{k}}(E_i) = \frac{L^3}{\pi^3}\frac{mk}{\hbar^2} . \tag{19.34}$$

Let us now compute the matrix element $\langle E_i^\alpha|W|E_m\rangle$. Since $\boldsymbol{A}_0$ is constant and since one can write

$$\boldsymbol{P} = \frac{\mathrm{i}m}{\hbar}\,[H_0, \boldsymbol{R}] , \tag{19.35}$$

one easily finds that

$$\begin{aligned} \langle E_i^\alpha|W|E_m\rangle &= -\frac{q_e}{m}\mathrm{e}^{\mathrm{i}\eta}\boldsymbol{A}_0\cdot\langle E_{\boldsymbol{k}}|\frac{\mathrm{i}m}{\hbar}\,[H_0, \boldsymbol{R}]\,|E_m\rangle \\ &= -\mathrm{i}q_e\omega_{km}\mathrm{e}^{\mathrm{i}\eta}\boldsymbol{A}_0\cdot\langle E_{\boldsymbol{k}}|\boldsymbol{R}|E_m\rangle . \end{aligned} \tag{19.36}$$

At this stage it is convenient to switch to the coordinate representation. Call z the direction of vector $\boldsymbol{k}$. Furthermore denote by (θ, ϕ) the polar and azimuthal angles of $\boldsymbol{A}_0$, and by (θ_r, ϕ_r) the polar and azimuthal angles of $\boldsymbol{r}$. Let $\bar{\theta}$ be the angle between $\boldsymbol{A}_0$ and $\boldsymbol{r}$ (Fig. 19.2). Recalling that $|E_m\rangle$ stands for the ground state of hydrogen one has

$$\begin{aligned} \langle E_{\boldsymbol{k}}|\boldsymbol{A}_0\cdot\boldsymbol{R}|E_m\rangle &= \int \mathrm{d}\boldsymbol{r}\,\frac{1}{(2L)^{3/2}}\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}}\left\{A_0\, r\cos\bar{\theta}\right\}\frac{1}{(\pi a_0^3)^{1/2}}\mathrm{e}^{-r/a_0} \\ &= \frac{A_0}{\left(8\pi a_0^3 L^3\right)^{1/2}}\int_0^\infty r^2\mathrm{d}r\int_0^\pi \sin\theta_r\,\mathrm{d}\theta_r\int_0^{2\pi}\mathrm{d}\phi_r\, r\mathrm{e}^{-r/a_0}\mathrm{e}^{-\mathrm{i}kr\cos\theta_r} \\ &\quad\times\left\{\cos\theta\cos\theta_r + \sin\theta\sin\theta_r\cos(\phi-\phi_r)\right\} . \end{aligned}$$

---

[5] In Sect. 9.4 we found the number of states $V(E)$ with energy lower than E. Clearly $\varrho_{\hat{k}}(E_i)$ is related to $V(E)$ as

$$\varrho_{\hat{k}}(E_i) = \frac{1}{4\pi}\left(\frac{\mathrm{d}V}{\mathrm{d}E}\right)_{E_i} . \tag{19.33}$$

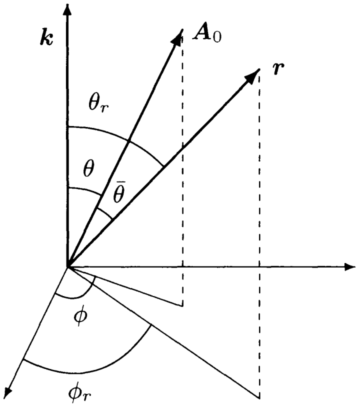

Fig. 19.2. Vectors $\boldsymbol{k}$, $\boldsymbol{r}$ and $\boldsymbol{A}_0$

We have made use of the expansion of $\cos\bar{\theta}$ in terms of the polar and azimuthal angles of $\boldsymbol{A}_0$ and $\boldsymbol{r}$. The integral on $\mathrm{d}\phi_r$ of the term involving $\cos(\phi-\phi_r)$ vanishes. Integrals on $\mathrm{d}\phi_r$ and $\mathrm{d}\theta_r$ are easily carried out and one finds (Exercise 19.4)

$$\langle E_{\boldsymbol{k}}|\boldsymbol{A}_0\cdot\boldsymbol{R}|E_m\rangle = -\frac{32\pi \mathrm{i} a_0^5 A_0}{\left(8\pi a_0^3 L^3\right)^{1/2}}\frac{k\cos\theta}{\left(1+a_0^2k^2\right)^3}. \tag{19.37}$$

We can now apply the Fermi golden rule (19.27), with the following result:

$$w_{\hat{k}} = \frac{256\, m\bar{k}^3\omega^2 q_e^2 A_0^2 a_0^7}{\pi\hbar^3\left(1+a_0^2\bar{k}^2\right)^6}\cos^2\theta. \tag{19.38}$$

This is the probability density, per unit time, that the atom makes a transition to a state in the continuous spectrum indexed by $\hat{k}$. Here $\bar{k}$ is defined as

$$\bar{E}_i = \frac{\hbar^2\bar{k}^2}{2m} = E_m + \hbar\omega. \tag{19.39}$$

Recall that angle θ stands for the direction of $\hat{k}$ relative to $\boldsymbol{A}_0$, the polarization of the incoming wave.

The total probability, per unit time, of a transition to an arbitrary state in the continuous spectrum is given by

$$w = \int \mathrm{d}\hat{k}\, w_{\hat{k}} = \frac{1024\, m\bar{k}^3\omega^2 q_e^2 A_0^2 a_0^7}{3\hbar^3\left(1+a_0^2\bar{k}^2\right)^6}. \tag{19.40}$$

19.3 Transition to the Discrete Spectrum

Let us now turn to a situation where the initial state $|E_m\rangle$ and the final state $|E_k\rangle$ both belong to the discrete spectrum. It may then be that $E_k < E_m$.

Suppose the quantum system, in the initial state $|E_m\rangle$, is subject to an electromagnetic wave of angular frequency ω. Following Sect. 19.1, we can expect transitions from $|E_m\rangle$ to $|E_k\rangle$ if $E_k - E_m \approx \pm\hbar\omega$. This means the transition can be produced by photons of different frequencies, provided they are very close to $\hbar^{-1}|E_k - E_m|$.

Actually, an electromagnetic wave is seldom exactly monochromatic. Provided they have very close frequencies, different components in a wave can contribute to the same transition. Accordingly we will assume that the wave's vector potential is given by a superposition like

$$\boldsymbol{A}(\boldsymbol{r},t) = \sum_{\boldsymbol{q}} 2\boldsymbol{A}_{\boldsymbol{q}} \cos\{\boldsymbol{q}\cdot\boldsymbol{r} - \omega(q)t + \eta_q\}\,. \tag{19.41}$$

Here $\boldsymbol{q}\cdot\boldsymbol{A}_{\boldsymbol{q}} = 0$ and $\omega(q) = qc$.[6]

In general $|E_m\rangle$ and $|E_k\rangle$ represent atomic or molecular states, i.e. states that involve several particles. Let us denote by q_i and m_i the charge and the mass of particle i, which can be an electron or a nucleus. Let $\boldsymbol{R}_i$ and $\boldsymbol{P}_i$ be the position and momentum operators. Under the influence of an electromagnetic wave the atom's Hamiltonian is, for all practical purposes, changed by the addition of a term[7]

$$\begin{aligned}\bar{H}(t) &= -\sum_i \frac{q_i}{m_i}\boldsymbol{A}(\boldsymbol{R}_i,t)\cdot\boldsymbol{P}_i\\ &= -\sum_i \frac{q_i}{m_i}\sum_{\boldsymbol{q}}\{\exp\left[\mathrm{i}\left(\boldsymbol{q}\cdot\boldsymbol{R}_i - \omega(q)t + \eta_q\right)\right]\\ &\qquad + \exp\left[-\mathrm{i}\left(\boldsymbol{q}\cdot\boldsymbol{R}_i - \omega(q)t + \eta_q\right)\right]\}\,\boldsymbol{A}_{\boldsymbol{q}}\cdot\boldsymbol{P}_i\,.\end{aligned} \tag{19.42}$$

To evaluate the coefficient $C_k^{(1)}(t)$ we have to substitute relation (19.42) into (19.16). The quantum system is subject to the electromagnetic wave beginning at $t = 0$, i.e. $\bar{H}(t) = 0$ if $t < 0$. Evaluating the time integral one gets as in Sect. 19.1[8]

---

[6] We use the convention that the sum on $\boldsymbol{q}$ not only covers different values of vector $\boldsymbol{q}$, but also the two linearly independent polarizations associated with a wave of a given $\boldsymbol{q}$.

[7] Terms proportional to the square of $\boldsymbol{A}$ are neglected. Note that coupling terms between spin and magnetic field have not been included either. The effect of these terms will be evaluated in Sect. 19.6. Note finally that at this stage $\boldsymbol{R}_i$ and $\boldsymbol{P}_i$ stand for the coordinates and momenta of each particle, and not for relative coordinates.

[8] Note that $|E_m\rangle$ and $|E_k\rangle$ not only represent the atom's internal configuration, but also the center-of-mass motion. For all practical purposes the center-of-mass motion does not change when a photon is absorbed or emitted. The matrix element therefore involves $\delta(\boldsymbol{k}_k - \boldsymbol{k}_m)$, where $\boldsymbol{k}_k$ and $\boldsymbol{k}_m$ are wave numbers associated with the translational motion of the final and initial states, respectively. The delta function can be ignored if one is concerned only with the probability to obtain a state whose internal configuration coincides with that of $|E_k\rangle$.

$$C_k^{(1)}(t) = \frac{1}{\hbar}\sum_i \frac{q_i}{m_i}\sum_{\boldsymbol{q}}\left\{ \mathrm{e}^{-\mathrm{i}\eta_q}\langle E_k|\mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_i}\boldsymbol{A}_{\boldsymbol{q}}\cdot\boldsymbol{P}_i|E_m\rangle \frac{\mathrm{e}^{\mathrm{i}[\omega_{km}+\omega(q)]t}-1}{\omega_{km}+\omega(q)} \right.$$
$$\left. + \mathrm{e}^{\mathrm{i}\eta_q}\langle E_k|\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_i}\boldsymbol{A}_{\boldsymbol{q}}\cdot\boldsymbol{P}_i|E_m\rangle \frac{\mathrm{e}^{\mathrm{i}[\omega_{km}-\omega(q)]t}-1}{\omega_{km}-\omega(q)} \right\}. \tag{19.43}$$

Let us first examine the case where the final state's energy is higher than the initial state's, i.e. $E_k > E_m$. Only the second term inside curly brackets has appreciable value, so that

$$C_k^{(1)}(t) = \frac{1}{\hbar}\sum_{\boldsymbol{q}} \mathrm{e}^{\mathrm{i}\eta_q}\langle E_k|\sum_i \frac{q_i}{m_i}\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_i}\boldsymbol{A}_{\boldsymbol{q}}\cdot\boldsymbol{P}_i|E_m\rangle \frac{\mathrm{e}^{\mathrm{i}[\omega_{km}-\omega(q)]t}-1}{\omega_{km}-\omega(q)}. \tag{19.44}$$

To evaluate the probability that the system goes from $|E_m\rangle$ to $|E_k\rangle$, we must compute the absolute square of $C_k^{(1)}(t)$. This involves a double sum on $\boldsymbol{q}$ and $\boldsymbol{q}'$. We assume that the superposition of the wave's different components is incoherent, i.e. that phases η_q are random. In this case diagonal terms in the double sum are all positive whereas nondiagonal terms, which are multiplied by factors like $\exp\{\mathrm{i}(\eta_q - \eta_{q'})\}$, essentially cancel each other.[9] This means that

$$\left|C_k^{(1)}(t)\right|^2 = \frac{4}{\hbar^2}\sum_{\boldsymbol{q}}\left|\langle E_k|\sum_i \frac{q_i}{m_i}\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_i}\boldsymbol{A}_{\boldsymbol{q}}\cdot\boldsymbol{P}_i|E_m\rangle\right|^2$$
$$\times \frac{1}{[\omega_{km}-\omega(q)]^2}\sin^2\left\{\frac{1}{2}[\omega_{km}-\omega(q)]\,t\right\}. \tag{19.45}$$

One can show (Exercise 19.5) that the time-averaged total intensity of radiation associated with the electromagnetic wave (19.41) is given by

$$I = 2\varepsilon_0 c\sum_{\boldsymbol{q}} \omega^2(q)\boldsymbol{A}_{\boldsymbol{q}}\cdot\boldsymbol{A}_{\boldsymbol{q}}. \tag{19.46}$$

Define $I_{\hat{q}}(\omega)\mathrm{d}\omega$ as the intensity of radiation directed along $\hat{q}$ and with frequency between ω and $\omega + \mathrm{d}\omega$. One then has

$$I = \sum_{\hat{q}}\int \mathrm{d}\omega\, I_{\hat{q}}(\omega). \tag{19.47}$$

---

[9] There are $N(N-1)$ nondiagonal terms. At any time the sum of these terms is of the order of N times the square of a matrix element, i.e. of the same order as the sum of diagonal terms. But the latter are always positive, whereas the sum of nondiagonal terms is an oscillating function of time. On the average, nondiagonal terms therefore vanish. Note that the assumption of random phases is certainly not appropriate to coherent waves like laser radiation. See [170], Sect. 8.7 and [199].

Thus expression (19.45) can be rewritten as

$$\left|C_k^{(1)}(t)\right|^2 = \frac{2}{\hbar^2\varepsilon_0 c}\sum_{\hat{q}}\int \mathrm{d}\omega\,\frac{1}{\omega^2}I_{\hat{q}}(\omega)\left|\langle E_k|\sum_i \frac{q_i}{m_i}\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_i}\hat{A}_{\hat{q}}\cdot\boldsymbol{P}_i|E_m\rangle\right|^2$$

$$\times\frac{1}{(\omega_{km}-\omega)^2}\sin^2\left\{\frac{1}{2}(\omega_{km}-\omega)t\right\}, \qquad (19.48)$$

where $\hat{A}_{\hat{q}}$ is a unit vector in the direction of $\boldsymbol{A}_{\boldsymbol{q}}$.

The integral on $\mathrm{d}\omega$ once more involves the factor (19.24), with an acute maximum around $\omega = \omega_{km}$. Insofar as $I_{\hat{q}}(\omega)$ and the matrix element are smooth enough functions of ω, they can be taken outside the integral. The probability per unit time that there is a transition from $|E_m\rangle$ to $|E_k\rangle$ is thus given by

$$w^a_{m\to k} = \frac{1}{T}\left|C_k^{(1)}(T)\right|^2$$

$$= \frac{\pi}{\varepsilon_0 c\hbar^2\omega_{km}^2}\sum_{\hat{q}} I_{\hat{q}}(\omega_{km})\left|\langle E_k|\sum_i \frac{q_i}{m_i}\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_i}\hat{A}_{\hat{q}}\cdot\boldsymbol{P}_i|E_m\rangle\right|^2 . \qquad (19.49)$$

The exponent a stands for absorption. It is understood that in the right-hand side, $q = \omega_{km}/c$.

Let us now turn to the case where the final state's energy is lower than the initial state's. By a similar argument one shows that the probability of emission per unit time is given by

$$w^e_{m\to k} = \frac{\pi}{\varepsilon_0 c\hbar^2\omega_{km}^2}\sum_{\hat{q}} I_{\hat{q}}\left(|\omega_{km}|\right)\left|\langle E_k|\sum_i \frac{q_i}{m_i}\mathrm{e}^{-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_i}\hat{A}_{\hat{q}}\cdot\boldsymbol{P}_i|E_m\rangle\right|^2 . \qquad (19.50)$$

It is important to point out that $w^e_{m\to k}$ represents the transition probability from $|E_m\rangle$ to $|E_k\rangle$ produced by an external field, i.e. the probability of *stimulated* emission. The quantity $w^e_{m\to k}$ does not take into account emission that would occur in the absence of external field, i.e. a *spontaneous* emission. We will come back to this in Sect. 19.4.

It is instructive to replace, in (19.50), the matrix element by its complex conjugate. Of course this does not change the value of the absolute square. Because $\boldsymbol{q}\cdot\hat{A}_{\hat{q}} = 0$, operators $\hat{q}\cdot\boldsymbol{R}_i$ and $\hat{A}_{\hat{q}}\cdot\boldsymbol{P}_i$ commute. Hence

$$w^e_{m\to k} = \frac{\pi}{\varepsilon_0 c\hbar^2\omega_{mk}^2}\sum_{\hat{q}} I_{\hat{q}}(\omega_{mk})\left|\langle E_m|\sum_i \frac{q_i}{m_i}\mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_i}\hat{A}_{\hat{q}}\cdot\boldsymbol{P}_i|E_k\rangle\right|^2$$

$$= w^a_{k\to m} . \qquad (19.51)$$

This means that the probability of stimulated emission from $|E_m\rangle$ to $|E_k\rangle$ is equal to the probability of absorption from $|E_k\rangle$ to $|E_m\rangle$.

It may happen that the matrix element in the previous expressions vanishes for every vector $\hat{A}_{\hat{q}}$. This occurs, for example, if $|E_m\rangle$ and $|E_k\rangle$ are both invariant under rotations (Exercise 19.6). The transition from $|E_m\rangle$ to $|E_k\rangle$ is then called *strictly forbidden.* It is important to point out that this does not mean that the transition is impossible. Indeed we have only considered first-order corrections due to the time-dependent perturbation, and we have neglected in the Hamiltonian the term proportional to $\boldsymbol{A}^2$. Strictly forbidden transitions only occur in very strong external fields. They give rise to the simultaneous absorption or emission of two photons.

19.4 Spontaneous Emission

Results of Sect. 19.3 show that in an electromagnetic field of angular frequency ω, a quantum system can go from an initial state with energy E_m to a final state with energy $E_k = E_m \pm \hbar\omega$. Transition probabilities $w^a_{m\to k}$ and $w^e_{m\to k}$, however, vanish if there is no external field.

This statement seems consonant with the general formalism of quantum mechanics. In the absence of external field, the system's Hamiltonian coincides with H_0. Vectors $|E_m\rangle$ and $|E_k\rangle$ are eigenvectors of H_0 which represent stationary states. If the quantum system is initially in state $|E_m\rangle$ it should, in the absence of external field, remain there indefinitely.

What we just described is not, however, what actually happens. Indeed suppose an atom is initially in an excited state. Most often then one observes that even without an external field the atom eventually comes back to its ground state, with the emission of one or several photons. This is called *spontaneous emission.*

How is it that the quantum-mechanical formalism could not predict the phenomenon of spontaneous emission? The answer is that the electromagnetic field has been considered as a classical object. From a fundamental point of view, components of the electromagnetic field at each point of space are dynamical variables which must be associated with operators rather than numbers. Such quantization of the electromagnetic field correctly predicts the probability of spontaneous emission.

We will not here develop the formalism of electromagnetic field quantization. The probability of spontaneous emission will be obtained in a different way, as suggested by Einstein in 1917.

Let us first point out that in the presence of isotropic radiation with intensity per unit frequency $I(\omega)$, the absorption probability $w^a_{m\to k}$ obtained in (19.49) is equal to $BI(\omega_{km})$, where B is a constant given by

$$B = \frac{\pi}{\varepsilon_0 c \hbar^2 \omega_{km}^2} \left\langle \left| \langle E_k | \sum_i \frac{q_i}{m_i} \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_i} \hat{A}_{\hat{q}} \cdot \boldsymbol{P}_i | E_m \rangle \right|^2 \right\rangle . \tag{19.52}$$

Large brackets here mean that an average is carried out on all directions of propagation and all polarizations.

Consider a cavity held at constant temperature T. Electromagnetic radiation inside has an intensity $I(\omega)$ that satisfies Planck's formula. Suppose that the cavity contains a certain number of identical atoms. Let N_m and N_k be the number of atoms in state $|E_m\rangle$ and state $|E_k\rangle$ where, to be specific, $E_k > E_m$. According to Gibbs's distribution

$$\frac{N_k}{N_m} = \exp\left(-\frac{E_k - E_m}{kT}\right) = \exp\left(-\frac{\hbar\omega_{km}}{kT}\right). \tag{19.53}$$

The number of transitions from $|E_m\rangle$ to $|E_k\rangle$, per unit time, is equal to $N_m w^a_{m\to k}$. On the other hand the number of transitions from $|E_k\rangle$ to $|E_m\rangle$, per unit time, is equal to $N_k(w^e_{k\to m} + w^s_{k\to m})$, where $w^s_{k\to m}$ stands for the probability of spontaneous emission per unit time. At equilibrium the number of transitions in both directions must be the same, so that

$$N_m w^a_{m\to k} = N_k(w^e_{k\to m} + w^s_{k\to m}). \tag{19.54}$$

Using the expression of N_k/N_m and the fact that $w^e_{k\to m} = w^a_{m\to k}$, one easily finds that

$$w^s_{k\to m} = \left[\frac{N_m}{N_k} - 1\right] w^a_{m\to k} = \left[\exp\left(\frac{\hbar\omega_{km}}{kT}\right) - 1\right] BI(\omega_{km}). \tag{19.55}$$

The function $I(\omega)$ is given in (1.4). Substituting it in the previous equation one finally obtains

$$w^s_{k\to m} = \frac{\hbar\omega^3_{km}}{\pi^2 c^2} B. \tag{19.56}$$

This is the total probability of spontaneous emission from $|E_k\rangle$ to $|E_m\rangle$, per unit time.

One can show ([200], Sect. 57) that the probability of spontaneous emission of a photon of wave number $\boldsymbol{q}$ and polarization $\hat{A}$, per unit solid angle and time in a transition from $|E_k\rangle$ to $|E_m\rangle$, is given by

$$\frac{\omega_{km}}{8\pi^2\varepsilon_0\hbar c^3}\left|\langle E_k|\sum_i \frac{q_i}{m_i} e^{i\boldsymbol{q}\cdot\boldsymbol{R}_i}\hat{A}\cdot\boldsymbol{P}_i|E_m\rangle\right|^2. \tag{19.57}$$

Thus not only is total spontaneous emission related to average stimulated emission, but the spontaneous emission of a photon with specific $\boldsymbol{q}$ and $\hat{A}$ is governed by the same matrix element as stimulated emission by such a photon.

19.5 Electric Dipole Transitions

Whether it is absorption, stimulated emission or spontaneous emission, the transition is always governed by the same matrix element conveniently denoted by M_{km}:

$$M_{km} = \langle E_k | \sum_i \frac{q_i}{m_i} \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_i} \hat{A}_{\hat{q}} \cdot \boldsymbol{P}_i | E_m \rangle . \tag{19.58}$$

In (19.58), $|E_k\rangle$ and $|E_m\rangle$ represent atomic or molecular states.[10] The spatial extension of wave functions is a few Å. By contrast the magnitude of vector $\boldsymbol{q}$ is equal to $2\pi/\lambda$, where λ is the radiation wavelength. In the case of infrared, visible or near ultraviolet radiation, one has $\lambda > 10^3$ Å. This means that the function $\exp(\mathrm{i}\boldsymbol{q} \cdot \boldsymbol{r}_i)$ changes very little if $\boldsymbol{r}_i$ varies over atomic or molecular dimensions. In this section we will make the approximation $\exp(\mathrm{i}\boldsymbol{q} \cdot \boldsymbol{R}_i) = I$. The matrix element is then given by

$$M_{km}^{\mathrm{E1}} = \langle E_k | \sum_i \frac{q_i}{m_i} \hat{A}_{\hat{q}} \cdot \boldsymbol{P}_i | E_m \rangle . \tag{19.59}$$

But as in (19.35),

$$\boldsymbol{P}_i = \frac{\mathrm{i}m_i}{\hbar} [H_0, \boldsymbol{R}_i] . \tag{19.60}$$

Hence by the same argument as in Sect. 19.2,

$$M_{km}^{\mathrm{E1}} = \mathrm{i}\omega_{km} \hat{A}_{\hat{q}} \cdot \langle E_k | \sum_i q_i \boldsymbol{R}_i | E_m \rangle . \tag{19.61}$$

The operator $\sum_i q_i \boldsymbol{R}_i$, which can be denoted by $\boldsymbol{D}$, coincides with the total dipole moment of the atom or molecule. For this reason the approximation that consists in identifying M_{km} to M_{km}^{E1} is called the *electric dipole approximation*. A transition is called *forbidden* if it cannot occur in the electric dipole approximation, i.e. if $M_{km}^{\mathrm{E1}} = 0$.

We have already pointed out that operators $\boldsymbol{R}_i$ stand for the coordinates of each electron or nucleus. Vectors $|E_k\rangle$ and $|E_m\rangle$, however, are most conveniently expressed in terms of center-of-mass and relative coordinates. Hence it is appropriate to express the total dipole moment in terms of these coordinates (Exercise 19.8). For a neutral atom or molecule made of N particles, one finds that the dipole moment does not depend on center-of-mass coordinates and can be written in terms of relative coordinates $\boldsymbol{R}_{iN} = \boldsymbol{R}_i - \boldsymbol{R}_N$ $(i = 1, 2, \ldots, N-1)$ as

$$\boldsymbol{D} = \sum_{i=1}^{N-1} q_i \boldsymbol{R}_{iN} . \tag{19.62}$$

---

[10] They can also represent nuclear states, which we will not explicitly consider.

For charged atoms or molecules this expression must be modified. The change is particularly slight in the case of an atom. At any rate we will restrict ourselves, for simplicity, to neutral atoms and molecules. But to keep notation simple we will still write the dipole moment as $\sum_i q_i \boldsymbol{R}_i$, keeping in mind that $\boldsymbol{R}_i$ stands for relative coordinates and that the sum goes from 1 to $N-1$.

In principle, matrix elements M_{km} can be calculated if vectors $|E_k\rangle$ and $|E_m\rangle$ are known. It happens however that many M_{km} vanish, at least in the electric dipole approximation. In Sect. 13.5 we have defined *selection rules* as necessary conditions for M_{km} to be different from zero. We now derive selection rules for electric dipole transitions in atoms. Molecules will be examined in Chap. 20.

Let us first recall that for every atom, the Hamiltonian commutes with total angular momentum components and space inversion. The latter operator also commutes with $\boldsymbol{J}$. Hence vectors $|E_k\rangle$ and $|E_m\rangle$ can always be picked so as to diagonalize H_0, $\boldsymbol{J}^2$, J_z and P.

Let ε_k and ε_m be eigenvalues of the space inversion operator associated with states $|E_k\rangle$ and $|E_m\rangle$. One has

$$\langle E_k|\boldsymbol{D}|E_m\rangle = \langle E_k|P\{P\boldsymbol{D}P\}P|E_m\rangle = \varepsilon_k\varepsilon_m\langle E_k|-\boldsymbol{D}|E_m\rangle, \tag{19.63}$$

where we used the fact that $P\boldsymbol{D}P = -\boldsymbol{D}$. Clearly M_{km}^{E1} vanishes if $\varepsilon_k = \varepsilon_m$. Therefore in every electric dipole transition, parity must change.

The dipole moment is a linear combination of position operators. Thus it is an irreducible tensor operator of rank 1. Denote by D_q ($q = 1, 0, -1$) its components like (14.89). Following the Wigner–Eckart theorem

$$\langle E_k|D_q|E_m\rangle \sim \langle 1, J_m; q, M_m|J_k, M_k\rangle . \tag{19.64}$$

Here J_m and M_m are quantum numbers associated with the total angular momentum and its projection on $\hat{z}$, for state $|E_m\rangle$; J_k and M_k are corresponding quantum numbers for $|E_k\rangle$. The Clebsch–Gordan coefficient vanishes unless $|J_m - 1| \le J_k \le J_m + 1$ and $q + M_m = M_k$. Hence we get the following selection rules:

$$\begin{aligned} J_k - J_m &= 0\,, \pm 1 \qquad (\text{except } J_k = 0 = J_m)\,; \\ M_k - M_m &= 0\,, \pm 1\,. \end{aligned} \tag{19.65}$$

Note however that due to symmetry properties of the Clebsch–Gordan coefficients, the transition $M_k = 0 = M_m$, $J_k = J_m$ is forbidden. Selection rules (19.65), like that of parity, are valid for electric dipole transitions in any atom.

Let us now focus on the more specific case of an atom whose wave functions are correctly obtained by LS coupling. Operators $\boldsymbol{L}$ and $\boldsymbol{S}$ then separately commute with the Hamiltonian. Obviously, the dipole moment is an irreducible tensor operator not only with respect to $\boldsymbol{J}$ but also with respect to $\boldsymbol{L}$. By the above argument one thus gets

$$L_k - L_m = 0\,, \pm 1 \qquad (\text{except } L_k = 0 = L_m)\,. \tag{19.66}$$

The spin, for its part, commutes with dipole moment. Hence

$$S_k - S_m = 0\,. \tag{19.67}$$

Recall that these rules hold in LS coupling. In real atoms they are often more or less violated, since the spin–orbit interaction prevents $\boldsymbol{L}$ and $\boldsymbol{S}$ from commuting with the Hamiltonian. Nevertheless atomic wave functions are often represented as linear combinations of functions with well-defined L and S. The LS coupling selection rules are then useful to compute matrix elements of $\boldsymbol{D}$ in these basis functions.[11]

Independently of LS coupling, let us examine the situation where each atomic wave function is adequately represented by just one configuration. Each wave function then is built from Q radial functions only. Following Sect. 17.3, the matrix element M_{km}^{E1} can be expressed entirely in terms of functions $\langle \phi_{ki} | \boldsymbol{R}_i | \phi_{mi} \rangle$, where $|\phi_{ki}\rangle$ and $|\phi_{mi}\rangle$ represent the i-th electronic wave function in Slater determinants making up states $|E_k\rangle$ and $|E_m\rangle$. One can see that M_{km}^{E1} vanishes if radial functions of more than one vector $|\phi_{ki}\rangle$ differ from those of the corresponding $|\phi_{mi}\rangle$. Furthermore $\langle \phi_{ki} | \boldsymbol{R}_i | \phi_{mi} \rangle$ clearly vanishes unless the parities of $|\phi_{ki}\rangle$ and $|\phi_{mi}\rangle$ are different. From this we conclude that M_{km}^{E1} vanishes except if, for one and only one i (say $i = \bar{\imath}$), radial functions associated with $|\phi_{k\bar{\imath}}\rangle$ and $|\phi_{m\bar{\imath}}\rangle$ are different. From the Wigner–Eckart theorem one easily sees that $l_{k\bar{\imath}} - l_{m\bar{\imath}} = 0, \pm 1$. This observation, together with the requirement of parity change, entails the following selection rule:

$$l_{k\bar{\imath}} - l_{m\bar{\imath}} = \pm 1\,. \tag{19.68}$$

This implies, among other things, that there can be no electric dipole transition between terms and multiplets of the same configuration.

Consider an atom with only one electron outside closed shells. Suppose the wave function is correctly described by just one configuration. If $|E_k\rangle$ and $|E_m\rangle$ represent states where only the valence electron is excited, then wave functions can be expressed by just one Slater determinant. From (17.33) and (19.61) one gets

$$M_{km}^{\mathrm{E1}} = \mathrm{i} q_e \omega_{km} \hat{A}_{\hat{q}} \cdot \langle \phi_{k\bar{\imath}} | \boldsymbol{R}_{\bar{\imath}} | \phi_{m\bar{\imath}} \rangle\,. \tag{19.69}$$

Here $\bar{\imath}$ labels the valence orbital.

Equation (19.69), as well as more general formulas (19.61) and (19.58), take the incoming wave's polarization and propagation direction into account. When incoming radiation is isotropic, the evaluation of emission and absorption probabilities involves taking an average as in (19.52). One can show (Exercise 19.9) that the average of $|M_{km}^{\mathrm{E1}}|^2$ on all directions of $\hat{A}_{\hat{q}}$ is equal to

---

[11] Matrix elements of $\boldsymbol{D}$ are dealt with in detail in [59], Chap. 14.

$$\left\langle \left| M_{km}^{\mathrm{E1}} \right|^2 \right\rangle = \frac{1}{3}\omega_{km}^2 \langle E_k|\boldsymbol{D}|E_m\rangle \cdot \langle E_k|\boldsymbol{D}|E_m\rangle^* . \tag{19.70}$$

From (19.56), (19.52), (19.58) and (19.70), one sees that the probability of spontaneous emission from $|E_k\rangle$ to $|E_m\rangle$, per unit time, is given in the electric dipole approximation by

$$w_{k\to m}^s = \frac{\omega_{km}^3}{3\pi\varepsilon_0 \hbar c^3} \langle E_k|\boldsymbol{D}|E_m\rangle \cdot \langle E_k|\boldsymbol{D}|E_m\rangle^* . \tag{19.71}$$

For an atom with just one valence electron, whose wave function is described by just one configuration and where only the valence electron is excited, this formula reduces to

$$w_{k\to m}^s = \frac{4}{3}\frac{\omega_{km}^3 e^2}{\hbar c^3} \langle \phi_{k\bar{\imath}}|\boldsymbol{R}_{\bar{\imath}}|\phi_{m\bar{\imath}}\rangle \cdot \langle \phi_{k\bar{\imath}}|\boldsymbol{R}_{\bar{\imath}}|\phi_{m\bar{\imath}}\rangle^* , \tag{19.72}$$

where (9.4) was used.

Table 19.1 displays the spontaneous emission probability for a transition between the first excited state (np) and the ground state (ns) of several alkali atoms. Theoretical values are computed from (19.72) and radial functions obtained from the central-field model (Sect. 9.6). Experimental values are drawn from [154]. The best configuration interaction calculations bring the relative difference between theoretical and experimental values down to a few parts in 10^3 [87].

Table 19.1. Spontaneous emission probabilities in alkali atoms. Values are given in multiples of 10^8 s^{-1}

	Transition	w^s (theory)	w^s (experiment)
Li	$2p \to 2s$	0.264	0.369
Na	$3p \to 3s$	0.444	0.610
K	$4p \to 4s$	0.282	0.382

19.6 Higher-Order Transitions

The electric dipole approximation of the matrix element M_{km} consists in replacing the exponential of $i\boldsymbol{q}\cdot\boldsymbol{R}_i$ by the identity operator. That substitution may be inadequate in two different situations.

The first one occurs when the radiation wavelength is of the order of atomic dimensions. This happens, for instance, in transitions involving inner orbitals of big atoms. The matrix element should then be evaluated by keeping the whole exponential.

In the second situation, the radiation wavelength is much larger than atomic dimensions. Let us expand the exponential in a power series, so that

$$\exp\{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_i\} = I+\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_i-\frac{1}{2}\left(\boldsymbol{q}\cdot\boldsymbol{R}_i\right)^2+\cdots \tag{19.73}$$

The electric dipole approximation consists in keeping only the first term in this expansion. If the resulting matrix element does not vanish, the approximation is usually very good. If, however, M_{km}^{E1} vanishes, then the correct value of M_{km} is obtained by using higher-order terms in the expansion of the exponential. Here we will only examine the second term and set

$$M'_{km} = \langle E_k|\sum_i \frac{q_i}{m_i}\mathrm{i}\left(\boldsymbol{q}\cdot\boldsymbol{R}_i\right)\left(\hat{A}_{\hat{q}}\cdot\boldsymbol{P}_i\right)|E_m\rangle. \tag{19.74}$$

One can show that (Exercise 19.12)

$$\left(\boldsymbol{q}\cdot\boldsymbol{R}_i\right)\left(\hat{A}_{\hat{q}}\cdot\boldsymbol{P}_i\right) = \frac{1}{2}\left(\boldsymbol{q}\times\hat{A}_{\hat{q}}\right)\cdot\boldsymbol{L}_i+\frac{\mathrm{i}m_i}{2\hbar}\left[H_0,\left(\boldsymbol{q}\cdot\boldsymbol{R}_i\right)\left(\hat{A}_{\hat{q}}\cdot\boldsymbol{R}_i\right)\right]. \tag{19.75}$$

The matrix element is therefore given by

$$M'_{km} = M_{km}^{\mathrm{E2}}+M_{km}^{\mathrm{M1}}, \tag{19.76}$$

where

$$M_{km}^{\mathrm{E2}} = -\frac{1}{2}\omega_{km}\langle E_k|\sum_i q_i\left(\boldsymbol{q}\cdot\boldsymbol{R}_i\right)\left(\hat{A}_{\hat{q}}\cdot\boldsymbol{R}_i\right)|E_m\rangle, \tag{19.77}$$

$$M_{km}^{\mathrm{M1}} = \frac{\mathrm{i}}{2}\left(\boldsymbol{q}\times\hat{A}_{\hat{q}}\right)\cdot\langle E_k|\sum_i \frac{q_i}{m_i}\boldsymbol{L}_i|E_m\rangle. \tag{19.78}$$

In Exercise (19.13) one shows that for a neutral atom whose center of mass coincides with the origin and for which M_{km}^{E1} vanishes, operators $\boldsymbol{R}_i$ and $\boldsymbol{L}_i$ can be replaced by relative coordinates and the associated angular momentum. The error made is of the order of the ratio of the electron's mass to the atom's mass.

Before looking more closely at the matrix elements, let us come back to the coupling terms between spin and magnetic field that have hitherto been left out. For electrons these terms are given by (Exercise 15.9).

$$-\sum_i \frac{q_i}{m_i}\boldsymbol{S}_i\cdot\boldsymbol{B}\left(\boldsymbol{R}_i,t\right). \tag{19.79}$$

It is not difficult to include these terms in the perturbation $\bar{H}(t)$ of Sect. 19.3, and compute their effect. If $\exp(\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_i)$ is replaced by the identity operator, one shows (Exercise 19.14) that the effect of these terms simply amounts to replace, in the expression of M_{km}^{M1}, the operator $\boldsymbol{L}_i$ by $\boldsymbol{L}_i + 2\boldsymbol{S}_i$.

So let us examine the element M_{km}^{M1} modified in this way. The operator $(q_i/m_i)(\boldsymbol{L}_i + 2\boldsymbol{S}_i)$ is proportional to the total magnetic moment of electron i. For this reason a transition governed by M_{km}^{M1} is called a *magnetic dipole transition.*[12] Furthermore

$$\sum_i \frac{q_i}{m_i}\left(\boldsymbol{L}_i + 2\boldsymbol{S}_i\right) = \frac{q_e}{m}\left(\boldsymbol{L} + 2\boldsymbol{S}\right) = \frac{q_e}{m}\left(\boldsymbol{J} + \boldsymbol{S}\right). \tag{19.80}$$

Here q_e and m are the charge and mass of an electron and $\boldsymbol{L}$, $\boldsymbol{J}$ and $\boldsymbol{S}$ have their usual meaning. Since $\boldsymbol{J}$ commutes with H_0, matrix elements $\langle E_k|\boldsymbol{J}|E_m\rangle$ vanish if $E_k \neq E_m$. Thus

$$M_{km}^{\mathrm{M1}} = \frac{\mathrm{i}}{2}\frac{q_e}{m}\left(\boldsymbol{q} \times \hat{A}_{\hat{q}}\right) \cdot \langle E_k|\boldsymbol{S}|E_m\rangle. \tag{19.81}$$

Selection rules of magnetic dipole transitions are easily obtained. Clearly the parity does not change. Since $\boldsymbol{S}$ is an irreducible tensor operator of rank 1

$$\begin{aligned} &J_k - J_m = 0\,, \pm 1 \qquad (\text{except } J_k = 0 = J_m)\,; \\ &M_k - M_m = 0\,, \pm 1\,. \end{aligned} \tag{19.82}$$

Note that in the limit of LS coupling, M_{km}^{M1} vanishes if $E_k \neq E_m$. In every case, if each atomic wave function is adequately represented by just one configuration, $|E_k\rangle$ and $|E_m\rangle$ must involve the same radial functions, i.e. must belong to the same configuration.

Let us now turn to the element M_{km}^{E2}. Since $\boldsymbol{q} \cdot \hat{A}_{\hat{q}} = 0$, one can see that $\sum_i q_i(\boldsymbol{q} \cdot \boldsymbol{R}_i)(\hat{A}_{\hat{q}} \cdot \boldsymbol{R}_i)$ is a component of the total quadrupole moment operator. For this reason a transition governed by M_{km}^{E2} is called an *electric quadrupole transition.* Clearly parity does not change in an electric quadrupole transition. Since the quadrupole moment is an irreducible tensor operator of rank 2, one has from the Wigner–Eckart theorem

$$\begin{aligned} &J_k - J_m = 0\,, \pm 1\,, \pm 2 \qquad (J_k + J_m \geq 2)\,; \\ &M_k - M_m = 0\,, \pm 1\,, \pm 2\,. \end{aligned} \tag{19.83}$$

Selection rules applicable to LS coupling are developed in [59], Sect. 15.2, where general formulas for the evaluation of matrix elements are given.

Let a be a length of the order of atomic dimensions. Insofar as $M_{km}^{\mathrm{E1}} \neq 0$, the order of magnitude of the probability of spontaneous emission from $|E_k\rangle$ to $|E_m\rangle$, per unit time, is given by (19.71) as

$$w_{k \to m}^{s} \approx \frac{\omega_{km}^3 e^2 a^2}{\hbar c^3}\,. \tag{19.84}$$

---

[12] The matrix element M_{km}^{M1} also governs the phenomena of electron spin resonance and, if nuclear spin is included, of nuclear magnetic resonance. They involve a transition, induced by a harmonic magnetic field, between Zeeman levels induced by a static magnetic field. See [15], Chap. 7, [53], Compl. $\mathrm{F_{IV}}$ and [170], Sect. 6.4.

For a visible light transition, one has $w^s_{k\to m} \approx 10^7\,\mathrm{s}^{-1}$. If $M^{\mathrm{E1}}_{km} = 0$ and the transition goes through electric quadrupole, the probability (19.84) is reduced by a factor of order $(2\pi a/\lambda_{km})^2$. If it goes through magnetic dipole, the probability is reduced by a factor of order α^2, where α is the fine structure constant.

Suppose an atom is in an excited state. In general the atom has various probabilities to go, by spontaneous emission, to various lower energy levels. Let w^s be the total probability, per unit time, that there be a spontaneous emission transition. Then $(w^s)^{-1}$ has dimensions of time and is called the *lifetime* of the excited state.

As a matter of fact, excited states often live for much less than their lifetime. This comes from the fact that de-excitation is due not only to photon emission, but also to collisions with neighboring particles.

19.7 The Shape of Spectral Lines

Bound-state energies of atoms and molecules coincide with the Hamiltonian's discrete spectrum. Thus one could expect that in a radiative transition from a state $|E_m\rangle$ to a state $|E_k\rangle$, the angular frequency of emitted or absorbed radiation exactly coincides with the difference between eigenvalues, divided by $\hbar$. Experimentally, we observe that radiation covers a frequency spectrum that is usually rather narrow, but never infinitely so.

Various phenomena can contribute to a widening of spectral lines.[13] The thermal motion of atoms, for example, produces a Doppler shift of the radiation frequency measured in the laboratory frame, in one or the other direction. Since different atoms move at different speeds, the measured frequency varies on an interval whose width $\delta\omega$ is of order $(v/c)\omega$, where v is the average speed of atoms. Furthermore, the presence of several other atoms in the neighborhood of a given atom produces additional, more or less random, terms in the Hamiltonian. Eigenvalues of the Hamiltonian change and, consequently, transition frequencies.

Of course, thermal agitation can be reduced by lowering temperature, and the effect of neighboring atoms turned down by using low-density gases. Even then, spectral lines retain a characteristic width. This remaining width, due to no external phenomenon, corresponds to the *natural shape* of spectral lines.

The natural line width comes from the fact that even in the absence of external factors, excited states of an atom or a molecule have a finite lifetime. Let us look at this through a simple example.

Let $|E_m\rangle$ and $|E_k\rangle$ be two excited states of an atom, with $E_k > E_m$. Let γ^{-1} be the lifetime of $|E_m\rangle$ and suppose that the lifetime of $|E_k\rangle$ is much larger than γ^{-1}. Furthermore suppose that at $t = 0$ the atom is in state $|E_m\rangle$

---

[13] A good discussion of line shape can be found in [64], Chap. 3.

and that it is subject to low-intensity monochromatic radiation of frequency $\omega \approx \omega_{km}$.

Insofar as radiation is low, it is clear that $|C_m(t)|^2 = \mathrm{e}^{-\gamma t}$, where $|C_m(t)|^2$ is the probability that at $t > 0$ the atom is in state $|E_m\rangle$. If, on the other hand, state $|E_k\rangle$ can only be reached by radiative transition from $|E_m\rangle$, one finds from time-dependent perturbation theory, in particular (19.5), that

$$\begin{aligned} C_k(t) &= \frac{1}{\mathrm{i}\hbar}\int_0^t \mathrm{d}t'\, C_m(t')\mathrm{e}^{\mathrm{i}\omega_{km}t'}\langle E_k|\bar{H}(t')|E_m\rangle \\ &= \frac{1}{\mathrm{i}\hbar}\int_0^t \mathrm{d}t'\, \exp\left\{\left(\mathrm{i}\omega_{km} - \frac{\gamma}{2}\right)t'\right\}\langle E_k|\bar{H}(t')|E_m\rangle\,, \end{aligned} \tag{19.85}$$

provided t is much smaller than the lifetime of $|E_k\rangle$. For monochromatic radiation

$$\bar{H}(t) = -\sum_i \frac{q_i}{m_i}\left\{\exp\left[\mathrm{i}\left(\boldsymbol{q}\cdot\boldsymbol{R}_i - \omega t\right)\right] + \exp\left[-\mathrm{i}\left(\boldsymbol{q}\cdot\boldsymbol{R}_i - \omega t\right)\right]\right\}\boldsymbol{A}_{\boldsymbol{q}}\cdot\boldsymbol{P}_i. \tag{19.86}$$

Carrying out the integral over $\mathrm{d}t'$ and keeping only the term with small denominator, one finds that

$$C_k(t) = -\frac{1}{\hbar}A_{\boldsymbol{q}}M_{km}\frac{1}{\omega - \omega_{km} - \mathrm{i}\gamma/2}\left\{\exp\left[\left(\mathrm{i}\omega_{km} - \mathrm{i}\omega - \frac{\gamma}{2}\right)t\right] - 1\right\}.$$

Finally we get, if T is much larger than γ^{-1} and much smaller than the lifetime of $|E_k\rangle$

$$|C_k(T)|^2 = \frac{A_{\boldsymbol{q}}^2|M_{km}|^2}{\hbar^2}\frac{1}{(\omega - \omega_{km})^2 + \gamma^2/4} \equiv \frac{A_{\boldsymbol{q}}^2|M_{km}|^2}{\hbar^2}f(\omega). \tag{19.87}$$

Due to the finite lifetime of $|E_m\rangle$, the probability that a photon is absorbed depends on its frequency. The probability is a Lorentzian function of frequency, centered on ω_{km} and with width at half maximum equal to γ (Fig. 19.3). Because $|E_m\rangle$ has finite lifetime, the transition from $|E_m\rangle$ to $|E_k\rangle$ can be induced by photons of different frequencies. The most natural interpretation consists in saying that the energy of $|E_m\rangle$ is not perfectly well-defined, but that it is equal to $E_m \pm \hbar\gamma$.

We should point out that the Lorentzian shape is much more general than the context in which we derived it. More precisely let γ_m^{-1} and γ_k^{-1} be the lifetimes of $|E_m\rangle$ and $|E_k\rangle$, due not only to spontaneous emission but to every form of decay. Then the Lorentzian shape is still valid, provided that γ is replaced by $\gamma_m + \gamma_k$. Note finally that if one tries, in any way whatsoever, to accurately measure the energy of state $|E_m\rangle$ on a large number of similarly prepared systems, the value E_m will be found up to $\pm\hbar\gamma_m$ only. The same remark applies, of course, to the energy of state $|E_k\rangle$. This constitutes a kind of uncertainty relation on time and energy.

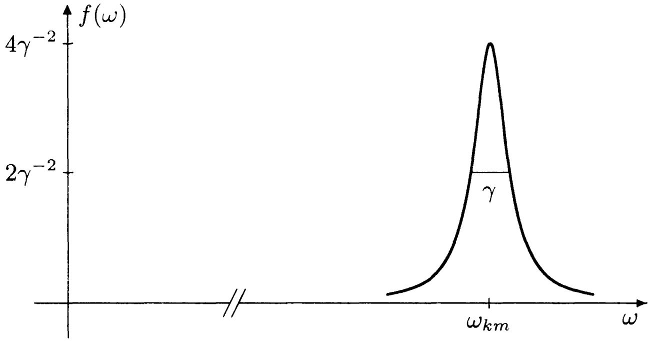

Fig. 19.3. Representation of the Lorentzian function

Exercises

19.1. A quantum system is initially in state $|E_m\rangle$. It is subject to a perturbation given by

$$\bar{H}(t) = \varepsilon f(t) W ,$$

where ε is small, $f(t)$ is a real function that vanishes when $t < 0$ and W is a Hermitian operator. Compute the probability that at $t > 0$ the system is in state $|E_k\rangle$. Consider specific cases where $f(t) = \delta(t - t_0)$ and where $f(t) = \theta(t - t_0)$. [$\theta(t - t_0)$ is the step function defined in (5.120).]

19.2. A hydrogen atom in its ground state is subject to an electric field $\boldsymbol{E}(t)$ such that

$$\boldsymbol{E}(t) = \begin{cases} 0 & \text{if } t < 0 , \\ E_0 \hat{z} \mathrm{e}^{-t/\tau} & \text{if } t \geq 0 . \end{cases}$$

a) Evaluate the perturbation $\bar{H}(t)$ produced by the field.
b) Use (19.16) to obtain, in first-order time-dependent perturbation theory, the probability that the atom is in state $2s$ ($n = 2, l = 0, m = 0$) after a long period of time. Find the corresponding probability for each $2p$ state ($n = 2$, $l = 1$, $m = 0$, ± 1).

19.3. Let W be the operator defined in (19.32), where $\boldsymbol{A}_0$ corresponds to an electromagnetic wave of intensity 10^4 W m^{-2} and wavelength 10^3 Å. Find the order of magnitude of matrix elements of W in states whose spatial extension is of order 1 Å. [Use the radiation intensity formula given in Exercise (19.5)].

19.4. Derive (19.37).

19.5.
a) Compute the Poynting vector $\boldsymbol{S} = \varepsilon_0 c^2 \boldsymbol{E} \times \boldsymbol{B}$ associated with the superposition of plane waves (19.41).

b) Compute the time average of $\boldsymbol{S}$ and show that radiation intensity (i.e. the energy per unit area and time) is given by

$$I = 2\varepsilon_0 c \sum_{\boldsymbol{q}} \omega^2(q) \boldsymbol{A}_{\boldsymbol{q}} \cdot \boldsymbol{A}_{\boldsymbol{q}} \,.$$

19.6. Show that the matrix element M_{km} given in (19.58) vanishes if $|E_k\rangle$ and $|E_m\rangle$ are invariant under rotations.

19.7. Suppose an atom is subject to radiation with uniform intensity in the frequency spectrum between $\omega - \frac{1}{2}\Delta\omega$ and $\omega + \frac{1}{2}\Delta\omega$. The frequency ω corresponds to a spectral line of width much smaller than $\Delta\omega$. With what total intensity will the radiation stimulate emission with a probability equal to that of spontaneous emission? Find the numerical value of the intensity for $\omega = 10^{15}\,\mathrm{s}^{-1}$ and $\Delta\omega = 10^{-5}\omega$.

19.8. Consider N particles with masses, electric charges and coordinates denoted by m_i, q_i and $\boldsymbol{R}_i$ ($i = 1, 2, \ldots, N$). Relative coordinates $\boldsymbol{R}_{iN}$ ($i = 1, 2, \ldots, N-1$) and center-of-mass coordinates $\boldsymbol{R}$ are defined as

$$\boldsymbol{R} = \frac{1}{M} \sum_{i=1}^{N} m_i \boldsymbol{R}_i \,, \qquad M = \sum_{i=1}^{N} m_i \,,$$
$$\boldsymbol{R}_{iN} = \boldsymbol{R}_i - \boldsymbol{R}_N \qquad (i = 1, 2, \ldots, N-1) \,.$$

The dipole moment $\boldsymbol{D}$ is defined as

$$\boldsymbol{D} = \sum_{i=1}^{N} q_i \boldsymbol{R}_i \,.$$

Write $\boldsymbol{D}$ in terms of $\boldsymbol{R}$ and $\boldsymbol{R}_{iN}$ and check that $\boldsymbol{D}$ does not depend on $\boldsymbol{R}$ if the sum of all charges vanishes.

19.9. Derive expression (19.70) for $\langle |M_{km}^{\mathrm{E1}}|^2 \rangle$, valid when incoming radiation is isotropic.

19.10. A radiative transition goes between two energy levels E_k and E_m, with angular momenta J_k and J_m. Show that the spontaneous emission probability, per unit time, from one state M_k to all states M_m is given in the electric dipole approximation by

$$\frac{\omega_{km}^3}{3\pi\hbar\varepsilon_0 c^3} \frac{1}{2J_k + 1} |\langle E_k || D || E_m \rangle|^2 \,.$$

19.11. A hydrogen atom is in a $2p$ state. Use the electric dipole approximation to compute the probability, per unit time, that it falls back to the ground state $1s$ by spontaneous emission.

19.12. Derive (19.75), where $\boldsymbol{q} \cdot \hat{A}_{\boldsymbol{q}} = 0$.

19.13. Consider the N particles and the coordinates introduced in Exercise (19.8).
a) Write momenta $\boldsymbol{P}_i$ in terms of momenta $\boldsymbol{P}_{iN}$ ($i = 1, 2, \ldots, N-1$) and $\boldsymbol{P}$ associated with relative and center-of-mass coordinates.
b) Show that if the total charge, the total dipole moment and the center-of-mass coordinates vanish, then

$$\sum_{i=1}^{N} q_i X_i Y_i = \sum_{i=1}^{N-1} q_i X_{iN} Y_{iN} + O(m_i/M)\,,$$

$$\sum_{i=1}^{N} \frac{q_i}{m_i} \boldsymbol{L}_i = \sum_{i=1}^{N-1} \frac{q_i}{m_i} \boldsymbol{L}_{iN} + O(m_i/M)\,,$$

where, in $O(m_i/M)$, $1 \le i \le N-1$.

19.14. In the perturbation $\bar{H}(t)$ of Sect. 19.3, include coupling terms between spin and magnetic field. Writing $\boldsymbol{B}$ as $\boldsymbol{\nabla} \times \boldsymbol{A}$ and letting $\exp(\mathrm{i}\boldsymbol{q} \cdot \boldsymbol{R}_i) \approx I$, show that these terms are taken into account if one replaces $\boldsymbol{L}_i$ by $\boldsymbol{L}_i + 2\boldsymbol{S}_i$ in the matrix element M_{km}^{M1}.

20 Molecules

The investigation of molecular properties by quantum-mechanical methods is a huge field.[1] Only the most elementary results can be presented here. First we will see that the quantum problem of a molecule approximately separates into an electronic problem and one for the motion of nuclei. Next we will examine electronic wave functions of diatomic molecules. The knowledge of electronic states allows to correctly formulate the problem of the motion of nuclei. This yields, in simple cases, rotational and vibrational energies of diatomic molecules, and corresponding selection rules. A few properties of polyatomic molecules will also be discussed.

20.1 The Born–Oppenheimer Approximation

A molecule is an aggregate of two or more atoms, whose energy is lower than the sum of the energies of each isolated atom (or sub-aggregate of atoms). A molecule is therefore a stable system. Left to itself it does not have the energy necessary to dissociate into free atoms or smaller molecules.[2]

The Hamiltonian of a molecule is obtained by a rather simple extension of the atomic Hamiltonian. Suppose a molecule is made of a certain number of atomic nuclei of charge Z_a, here taken as point particles, and of a certain number of electrons. Let $\boldsymbol{R}_a$ be the position operator of nucleus a and $\boldsymbol{R}_i$ the position operator of electron i. Indices a and i assume appropriate values which at this stage will not be specified further. Let M_a be the mass of nucleus a and m the electronic mass. If spin and relativistic effects are neglected, the molecule's Hamiltonian can be written in the coordinate representation as

---

[1] So is the size of the literature. The reader is more specifically referred to [71], [138], [212], [214] and [233].

[2] Strictly speaking our definition of molecule excludes chemically stable systems made of one atom only, such as noble gases. It includes, on the other hand, macroscopic objects like crystals. Actually we will investigate simple molecules only, made up of a small number of atoms.

L. Marchildon, *Quantum Mechanics*

$$H = -\hbar^2 \sum_a \frac{1}{2M_a} \nabla_a^2 - \frac{\hbar^2}{2m} \sum_i \nabla_i^2 + \sum_{a<b} \frac{Z_a Z_b e^2}{|\boldsymbol{r}_a - \boldsymbol{r}_b|} + \sum_{i<j} \frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} - \sum_{i,a} \frac{Z_a e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_a|} . \tag{20.1}$$

Here ∇_a^2 stands for the Laplacian with respect to coordinates $\boldsymbol{r}_a$. Various terms represent the electrostatic potential energy between nuclei, between electrons and between electrons and nuclei.

Denote by $\boldsymbol{S}$, $\boldsymbol{L}$ and $\boldsymbol{J}$ the molecule's total spin, total orbital angular momentum and total angular momentum, that is,

$$\boldsymbol{S} = \sum_a \boldsymbol{S}_a + \sum_i \boldsymbol{S}_i , \tag{20.2}$$

$$\boldsymbol{L} = \sum_a \boldsymbol{L}_a + \sum_i \boldsymbol{L}_i , \tag{20.3}$$

$$\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S} . \tag{20.4}$$

One can see that $\boldsymbol{S}$, $\boldsymbol{L}$ and $\boldsymbol{J}$ commute with H. If, however, spin–orbit interaction terms are added to the Hamiltonian, only the components of total angular momentum commute with it. On the other hand, in all cases the space-inversion operator P commutes with H. Hence the Hamiltonian of any molecule can be diagonalized simultaneously with $\boldsymbol{J}^2$, J_z and P.

Let $V(\boldsymbol{r}_a, \boldsymbol{r}_i)$ be the molecule's total electrostatic potential energy, that is,

$$V(\boldsymbol{r}_a, \boldsymbol{r}_i) = \sum_{a<b} \frac{Z_a Z_b e^2}{r_{ab}} + \sum_{i<j} \frac{e^2}{r_{ij}} - \sum_{i,a} \frac{Z_a e^2}{r_{ia}} . \tag{20.5}$$

Here $r_{ab} = |\boldsymbol{r}_a - \boldsymbol{r}_b|$ and similarly with r_{ij} and r_{ia}. The Hamiltonian's eigenvalue equation can be written as

$$\left\{ -\hbar^2 \sum_a \frac{1}{2M_a} \nabla_a^2 - \frac{\hbar^2}{2m} \sum_i \nabla_i^2 + V(\boldsymbol{r}_a, \boldsymbol{r}_i) \right\} \Psi(\boldsymbol{r}_a, \boldsymbol{r}_i) = E \Psi(\boldsymbol{r}_a, \boldsymbol{r}_i). \tag{20.6}$$

Obviously this equation cannot in general be solved exactly. So use must be made of different approximations. We know that nuclear masses are typically ten thousand times bigger than the electron's mass. Thus one can think that electrons move much faster than nuclei. Actually we will assume that the electronic wave function can be obtained by clamping nuclei at fixed positions. The wave function of nuclei will then be found from an effective potential due to the motion of electrons and the electrostatic interaction of nuclei. This approach is called the *Born–Oppenheimer approximation.* It is used almost universally in molecular physics, at least as a starting point. We now make it more precise.

Assume then that as a first step the positions of nuclei are fixed at points with coordinates $\boldsymbol{r}_a$. Formally this is the same as giving infinite masses to nuclei. The electronic wave function, that can be denoted by ψ, then satisfies the equation

$$\left\{-\frac{\hbar^2}{2m}\sum_i \nabla_i^2 + V(\boldsymbol{r}_a, \boldsymbol{r}_i)\right\}\psi(\boldsymbol{r}_a, \boldsymbol{r}_i) = E'(\boldsymbol{r}_a)\psi(\boldsymbol{r}_a, \boldsymbol{r}_i). \tag{20.7}$$

Here E' is a function of $\boldsymbol{r}_a$, as we expect that eigenvalues depend on the fixed positions of nuclei. The electronic wave function also depends parametrically on $\boldsymbol{r}_a$. And of course it depends on the electronic coordinates $\boldsymbol{r}_i$.

The Born–Oppenheimer approximation first consists in assuming that the molecular wave function $\Psi(\boldsymbol{r}_a, \boldsymbol{r}_i)$ can be written as

$$\Psi(\boldsymbol{r}_a, \boldsymbol{r}_i) = v(\boldsymbol{r}_a)\psi(\boldsymbol{r}_a, \boldsymbol{r}_i), \tag{20.8}$$

where $\psi(\boldsymbol{r}_a, \boldsymbol{r}_i)$ satisfies (20.7). Substituting (20.8) in (20.6) and using (20.7), one gets

$$\left\{-\hbar^2\sum_a \frac{1}{2M_a}\nabla_a^2 + E'(\boldsymbol{r}_a) - E\right\} v(\boldsymbol{r}_a)\psi(\boldsymbol{r}_a, \boldsymbol{r}_i) = 0. \tag{20.9}$$

Next one assumes that the change in the electronic wave function due to a change in the coordinates of nuclei is much smaller than the corresponding change in the function $v(\boldsymbol{r}_a)$. Let us multiply (20.9) by $\psi^*(\boldsymbol{r}_a, \boldsymbol{r}_i)$ and integrate on the $\mathrm{d}\boldsymbol{r}_i$. Neglecting terms involving derivatives of $\psi(\boldsymbol{r}_a, \boldsymbol{r}_i)$ with respect to components of $\boldsymbol{r}_a$, one finds that (Exercise 20.1)[3]

$$-\hbar^2\sum_a \frac{1}{2M_a}\nabla_a^2 v(\boldsymbol{r}_a) + E'(\boldsymbol{r}_a)v(\boldsymbol{r}_a) = Ev(\boldsymbol{r}_a). \tag{20.10}$$

Thus $v(\boldsymbol{r}_a)$ satisfies the Hamiltonian's eigenvalue equation corresponding to a potential $E'(\boldsymbol{r}_a)$. Recall that $E'(\boldsymbol{r}_a)$ represents the eigenvalue associated with the electronic wave function, under the assumption that nuclei are fixed.

To sum up, in the Born–Oppenheimer approximation the molecular wave function $\Psi(\boldsymbol{r}_a, \boldsymbol{r}_i)$ factorizes into an electronic wave function $\psi(\boldsymbol{r}_a, \boldsymbol{r}_i)$ that satisfies (20.7) and a wave function of nuclei $v(\boldsymbol{r}_a)$ that satisfies (20.10). The eigenvalues associated with (20.7) are necessary for the solution of (20.10), hence (20.7) must be solved first.

The function $E'(\boldsymbol{r}_a)$ determines the stability of the molecule and the mean distances between nuclei.[4] Indeed one shows that (Exercise 20.2)

---

[3] For a diatomic molecule the sharper analysis of Sect. 20.4 will lead to a slightly different equation. A more complete discussion of the Born–Oppenheimer approximation can be found in [243].

[4] For given values of $\boldsymbol{r}_a$ (20.7) has several solutions, for ψ as well as for E'. The function $E'(\boldsymbol{r}_a)$ should therefore be carefully constructed, especially if there is degeneracy.

$$\nabla_{\boldsymbol{r}_a} E' = \int \left[\prod_i \mathrm{d}\boldsymbol{r}_i\right] \psi^*(\boldsymbol{r}_a, \boldsymbol{r}_i)\left[\nabla_{\boldsymbol{r}_a} V(\boldsymbol{r}_a, \boldsymbol{r}_i)\right]\psi(\boldsymbol{r}_a, \boldsymbol{r}_i), \tag{20.11}$$

a relation known as the *Hellmann–Feynman theorem.* The right-hand side of (20.11) represents (up to a sign) the total force acting on nucleus a. Consequently, the position of stable equilibrium of nuclei coincides with values of $\boldsymbol{r}_a$ which minimize $E'(\boldsymbol{r}_a)$. The mean distances between nuclei can be obtained from these values, since vibrational amplitudes are much smaller than molecular dimensions.

It is not difficult to see, from the explicit form of the potential $V(\boldsymbol{r}_a, \boldsymbol{r}_i)$, that (20.7) for the electronic wave function is invariant under a rotation and a translation of all coordinates $\boldsymbol{r}_a$ and $\boldsymbol{r}_i$. This entails that $E'(\boldsymbol{r}_a)$ is invariant under a rotation and a translation of the $\boldsymbol{r}_a$. It is very convenient to define the function $\psi(\boldsymbol{r}_a, \boldsymbol{r}_i)$ in such a way that it is also invariant under a rotation and a translation of $\boldsymbol{r}_a$ and $\boldsymbol{r}_i$. Physically this corresponds to picking $\boldsymbol{r}_i$ not as inertial coordinates, but as coordinates defined in a reference frame rigidly attached to the nuclei. Unless otherwise indicated we will always use such coordinates. In Sect. 20.4 we will see how to define them explicitly for a diatomic molecule.

It is clear that $\psi(\boldsymbol{r}_a, \boldsymbol{r}_i)$ is not invariant under a rotation of the $\boldsymbol{r}_a$ only, or of the $\boldsymbol{r}_i$ only. On the other hand assume that the configuration of nuclei is invariant under a point group G. Then (20.7) is invariant under a transformation that leaves the $\boldsymbol{r}_i$ fixed and transforms the $\boldsymbol{r}_a$ according to G. But (20.7) is always invariant under a simultaneous rotation (or space inversion) of $\boldsymbol{r}_a$ and $\boldsymbol{r}_i$. Therefore, it is invariant under a transformation that leaves the $\boldsymbol{r}_a$ fixed and transforms the $\boldsymbol{r}_i$ according to G. From Chap. 13 one conclude that the $\psi(\boldsymbol{r}_a, \boldsymbol{r}_i)$, for $\boldsymbol{r}_a$ and $E'(\boldsymbol{r}_a)$ fixed, make up the space of a representation of G. Similarly, a look at (20.10) leads us to conclude that for E fixed, the functions $v(\boldsymbol{r}_a)$ also make up the space of a representation of G.

In the following sections we will examine (20.7) and (20.10) more closely. But before, let us make a few general remarks on molecular wave functions and energies.

We know that the tendency of atoms to form molecules is largely determined by electrons in outer atomic orbitals. They are responsible for chemical bonds. When a molecule is formed, valence electrons no longer belong to a specific atom, but to the entire molecule. Some molecular orbitals will thus cover the whole molecule, or an important part of it. On the other hand, inner atomic orbitals are little influenced by the formation of a molecule. Thus some molecular orbitals are very similar to atomic orbitals. Ionization energies associated with inner molecular orbitals coincide, for all practical purposes, with corresponding atomic ionization energies. As for ionization energies associated with outer molecular orbitals, they are as in the atomic case of the order of a few eV. The precise values, however, are very different.

The motion of nuclei, governed by (20.10), also gives rise to characteristic molecular energies. This we already saw, for the case of diatomic molecules, in Sect. 10.7. Rotational energies can be estimated from the classical formula $J^2/2I$, where I is the moment of inertia and J the angular momentum. The moment of inertia is of order Ma^2, where M is the molecule's mass and a its linear dimension. The angular momentum is of order $\hbar$. For a molecule of a few Å and a few tens of nucleons, one finds that the first few rotational energies are of order 10^{-4} eV. Associate frequencies fall in the range of microwaves.

To estimate vibrational energies, one starts with the observation that a vibration with amplitude of the order of molecular dimensions would certainly induce major electronic rearrangement. The energy $M\omega^2 a^2/2$ associated with such a vibration should thus be of a few eV. From this one finds that $\hbar\omega$ is of order 10^{-1} eV, with frequencies typical of the infrared region. Vibrational energies are thus much larger than rotational energies and much smaller than electronic energies. Note that for a gas at room temperature, essentially all molecules are in the lowest electronic energy level and most are in the lowest vibrational level. Molecules are found, however, in several different rotational energy levels.

It follows from this analysis that electronic motion is much faster than rotational or vibrational motion. This adds weight to the Born–Oppenheimer assumptions, according to which electronic wave functions can be calculated by taking fixed nuclei, and wave functions of nuclei can be calculated from a mean-value energy associated with the electronic configuration.

20.2 Molecular Orbitals

The Hartree–Fock method can be used to obtain approximate solutions of (20.7). That equation determines electronic wave functions in the Born–Oppenheimer approximation.

Let us first write, with similar notations as in Sect. 17.3,

$$F(i) = -\frac{\hbar^2}{2m}\nabla_i^2 - \sum_a \frac{Z_a e^2}{r_{ia}} , \tag{20.12}$$

$$G(i,j) = \frac{e^2}{r_{ij}} , \tag{20.13}$$

and let us define $\bar{E}$ as

$$\bar{E} = E'(\boldsymbol{r}_a) - \sum_{a<b} \frac{Z_a Z_b e^2}{r_{ab}} . \tag{20.14}$$

Equation (20.7) becomes

$$\left\{ \sum_i F(i) + \sum_{i<j} G(i,j) \right\} \psi(\boldsymbol{r}_i) = \bar{E}\psi(\boldsymbol{r}_i) . \tag{20.15}$$

Since coordinates $\boldsymbol{r}_a$ are fixed, the $\boldsymbol{r}_a$ dependence of ψ and $\bar{E}$ has not been shown explicitly.

Let Q be the number of electrons in the molecule. The total electronic wave function, with spin included, can be written in terms of determinantal functions like

$$\Phi(1,2,\ldots,Q) = \frac{1}{\sqrt{Q!}} \begin{vmatrix} \phi_1(1) \ldots \phi_1(Q) \\ \ldots \quad \ldots \quad \ldots \\ \phi_Q(1) \ldots \phi_Q(Q) \end{vmatrix} . \tag{20.16}$$

Here

$$\phi_i(j) = u_i(\boldsymbol{r}_j)\chi_i(j) , \tag{20.17}$$

where $u_i(\boldsymbol{r}_j)$ is a spatial function and $\chi_i(j)$ a spin 1/2 vector. The Hartree–Fock method looks for the best solution of (20.15) in the form of determinantal functions. Clearly the Hamiltonian associated with (20.15) is of the same type as Hamiltonian (17.36). Use of the variational method, following an argument formally identical with the one in Sect. 17.4, therefore leads to the following equations for the functions u_i:

$$\begin{gathered} \left\{ -\frac{\hbar^2}{2m}\nabla_1^2 - \sum_a \frac{Z_a e^2}{r_{1a}} - \lambda_i \right\} u_i(\boldsymbol{r}_i) \\ + \sum_j \int \mathrm{d}\boldsymbol{r}_2\, u_j^*(\boldsymbol{r}_2)\frac{e^2}{r_{12}} \left[u_i(\boldsymbol{r}_1)u_j(\boldsymbol{r}_2) - \delta(i,j)u_j(\boldsymbol{r}_1)u_i(\boldsymbol{r}_2)\right] = 0 . \end{gathered} \tag{20.18}$$

Functions $u_i(\boldsymbol{r})$ are known as *molecular orbitals.* The difference between molecular and atomic orbitals comes solely from the sum over a inside the curly brackets. In the atomic case this sum reduces to a single angle-independent term. In the molecular case this sum involves two or more terms, so that spherical symmetry disappears.

The term inside curly brackets is invariant under the point group G that transforms into itself the configuration of nuclei. The second term in (20.18) is not invariant under G but as in the atomic case, one can substitute to it a mean value that is. This means that for all practical purposes, orbitals u_i associated with a given λ_i make up the space of a representation of the point group G.

With a self-consistent approach one can, in principle, solve the Hartree–Fock equations for molecular orbitals. But as a matter of fact the problem is very complicated, even with relatively small molecules. One can, however, replace the differential problem by an algebraic problem simpler to deal with [196]. Vectors $|\phi_i\rangle$ are first written in terms of vectors $|\psi_\alpha\rangle$ in the following way:

$$|\phi_i\rangle = \sum_{\alpha=1}^{M} c^i_\alpha |\psi_\alpha\rangle \,. \tag{20.19}$$

The c^i_α are complex coefficients, M is an integer significantly larger than Q and the $|\psi_\alpha\rangle$ are products of spatial wave functions and spin functions. Actually the spatial wave functions are often picked as atomic orbitals. This choice gives rise to the so-called *LCAO method* ("Linear Combinations of Atomic Orbitals").

Vectors $|\psi_\alpha\rangle$ are not necessarily orthogonal or normalized. We let

$$S_{\alpha\beta} = \langle\psi_\alpha|\psi_\beta\rangle \,. \tag{20.20}$$

The mean value in state $|\Phi\rangle$ of the Hamiltonian associated with (20.15) can be written in terms of the $|\psi_\alpha\rangle$ and c^i_α. Using (17.33), (17.34) and (20.19) one finds that

$$\begin{aligned}\langle\Phi|H|\Phi\rangle &= \sum_{i=1}^{Q}\sum_{\alpha,\beta=1}^{M} (c^i_\alpha)^* c^i_\beta \langle\psi_\alpha|F|\psi_\beta\rangle \\ &+ \sum_{i<j}\sum_{\alpha,\beta,\gamma,\delta=1}^{M} (c^i_\alpha)^*(c^j_\gamma)^* c^i_\beta c^j_\delta \left\{\langle\psi_\alpha\psi_\gamma|G|\psi_\beta\psi_\delta\rangle - \langle\psi_\alpha\psi_\gamma|G|\psi_\delta\psi_\beta\rangle\right\} .\end{aligned} \tag{20.21}$$

The orthonormality of the $|\phi_i\rangle$ is written as

$$\delta_{ij} = \langle\phi_i|\phi_j\rangle = \sum_{\alpha,\beta=1}^{M} (c^i_\alpha)^* c^j_\beta S_{\alpha\beta} \,. \tag{20.22}$$

The problem now consists in finding coefficients c^i_α that make the functional $\langle\Phi|H|\Phi\rangle$ stationary with the constraints $\langle\phi_i|\phi_j\rangle = \delta_{ij}$. One shows (Exercise 20.3) that the conditions for stationarity are given by

$$\begin{aligned}\sum_{\beta=1}^{M} c^i_\beta \Bigg\{ &\langle\psi_\alpha|F|\psi_\beta\rangle - \lambda_i S_{\alpha\beta} \\ &+ \sum_{j=1}^{Q}\sum_{\gamma,\delta=1}^{M} (c^j_\gamma)^* c^j_\delta \left[\langle\psi_\alpha\psi_\gamma|G|\psi_\beta\psi_\delta\rangle - \langle\psi_\alpha\psi_\gamma|G|\psi_\delta\psi_\beta\rangle\right] \Bigg\} = 0 \,.\end{aligned} \tag{20.23}$$

For each value of i, (20.23) represents a system of M equations (one equation for each value of α) for M unknowns c^i_β. For nontrivial solutions to exist the determinant of the matrix of coefficients c^i_β must vanish. There results an M-th degree equation for λ_i, which has in general M real solutions. To each of these solutions there corresponds a set of M coefficients c^i_β. Thus one really obtains M (rather than Q) orbitals $|\phi_i\rangle$. Each determinal wave function is, however, built from Q orbitals.

Note that the M-th degree secular equation for λ_i depends on coefficients c^j_α. Like the Hartree–Fock equations, it is solved most naturally in a self-consistent way. One picks initial values of the coefficients c^j_α that one substitutes inside curly brackets in (20.23). Next one computes the λ_i and finds a better approximation for the c^i_β. The computation is carried out iteratively till convergence.

For small molecules Roothaan's algebraic method can be applied as outlined. For larger molecules it soon becomes too costly in computer resources. Different avenues are used to simplify it ([233], Sect. 25.3). Some calculations, for instance, completely neglect integrals involving functions centered on more than two points. Others consider valence electrons only, and represent the effect of nuclei and inner electrons by appropriate potentials. The latter approximation is motivated by the fact that inner electrons contribute very little to the formation of molecules.

To illustrate the LCAO method we will examine the simplest of molecules, that is, the H_2^+ ion.[5] To simplify notations let us denote the two protons by a and b and the distance between the electron and each proton by r_a and r_b respectively. Let $\boldsymbol{r}$ be the vector going from proton a to proton b and let $\boldsymbol{r}_e$ be the electron's coordinate with respect to the proton's center of mass (Fig. 20.1). One has

$$r_a = \left|\boldsymbol{r}_e + \frac{1}{2}\boldsymbol{r}\right|, \qquad r_b = \left|\boldsymbol{r}_e - \frac{1}{2}\boldsymbol{r}\right|. \tag{20.24}$$

Because the protons are fixed the system's Hamiltonian is given by

$$H = -\frac{\hbar^2}{2m}\nabla^2_{r_e} + \frac{e^2}{r} - \frac{e^2}{r_a} - \frac{e^2}{r_b}. \tag{20.25}$$

Let w_a and w_b be $1s$ hydrogen orbitals, centered on protons a and b. One has

$$w_a(\boldsymbol{r}_e) = \frac{1}{(\pi a_0^3)^{1/2}} \mathrm{e}^{-r_a/a_0}, \qquad w_b(\boldsymbol{r}_e) = \frac{1}{(\pi a_0^3)^{1/2}} \mathrm{e}^{-r_b/a_0}. \tag{20.26}$$

Let us write the molecular orbital $|u\rangle$ as a linear combination of atomic orbitals, that is,[6]

---

[5] The problem of the H_2^+ ion, for a fixed distance between nuclei, can be solved exactly. The solution is developed in [212], App. 1, and in references given therein. Unfortunately the solution does not generalize to other molecules.

[6] It may help, at this point, to summarize the relation between the various symbols ϕ, ψ, u and w. Functions ϕ and u stand for molecular orbitals while ψ and w stand for atomic orbitals. On the other hand ϕ and ψ take spin into account while u and w are spatial functions only. Thus $\phi = u\chi$ and $\psi = w\chi$.

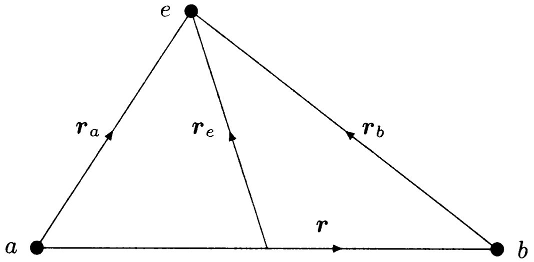

Fig. 20.1. Coordinates used for the H_2^+ ion

$$|u\rangle = c_a|w_a\rangle + c_b|w_b\rangle \,. \tag{20.27}$$

Let $H_{aa} = \langle w_a|H|w_a\rangle$ and similarly with H_{ab}, H_{ba} and H_{bb}. Let $S = \langle w_a|w_b\rangle$. The reality of wave functions and the symmetry of the problem imply that $H_{aa} = H_{bb}$, $H_{ab} = H_{ba}$ and $S = S^*$. One shows (Exercise 20.4) that enforcing the stationarity of the functional $\langle u|H|u\rangle$ with the constraint $\langle u|u\rangle = 1$ yields the following two solutions for $|u\rangle$ and $E = \langle u|H|u\rangle$:

$$|u_1\rangle = \frac{1}{\sqrt{2(1+S)}}\{|w_a\rangle + |w_b\rangle\}\,, \quad E_1 = \frac{1}{1+S}\{H_{aa} + H_{ab}\}\,; \tag{20.28}$$

$$|u_2\rangle = \frac{1}{\sqrt{2(1-S)}}\{|w_a\rangle - |w_b\rangle\}\,, \quad E_2 = \frac{1}{1-S}\{H_{aa} - H_{ab}\}\,. \tag{20.29}$$

Quantities S, H_{aa} and H_{ab} will soon be computed in a more general context. They depend, of course, on the distance r between protons. One finds that for all r, $E_1 < E_2$. Values of E_1 and E_2 as functions of r are plotted in Fig. 3.2, p. 48.

It is instructive to write down the probability density to find the electron at $\boldsymbol{r}_e$, corresponding to each orbital $|u_1\rangle$ and $|u_2\rangle$. Since all these functions are real

$$|u_1(\boldsymbol{r}_e)|^2 = \frac{1}{2(1+S)}\left\{[w_a(\boldsymbol{r}_e)]^2 + [w_b(\boldsymbol{r}_e)]^2 + 2w_a(\boldsymbol{r}_e)w_b(\boldsymbol{r}_e)\right\}, \tag{20.30}$$

$$|u_2(\boldsymbol{r}_e)|^2 = \frac{1}{2(1-S)}\left\{[w_a(\boldsymbol{r}_e)]^2 + [w_b(\boldsymbol{r}_e)]^2 - 2w_a(\boldsymbol{r}_e)w_b(\boldsymbol{r}_e)\right\}. \tag{20.31}$$

Functions $w_a(\boldsymbol{r}_e)$ and $w_b(\boldsymbol{r}_e)$ are everywhere positive. Obviously, the probability to find the electron between the two protons is significantly larger in $|u_1\rangle$ than in $|u_2\rangle$. In $|u_1\rangle$, for values of r large enough, the protons' electrostatic repulsion is compensated by the proton-electron attraction. For small values of r the protons' repulsion wins over. This explains that the plot of E versus r displays a minimum. In $|u_2\rangle$, by contrast, the protons' repulsion wins over for all r. Orbital $|u_1\rangle$ is called *bonding*, while $|u_2\rangle$ is *antibonding*.

Let r_0 be the value of r for which $E_1(r)$ is minimum. Our approximation has expanded molecular orbitals in terms of two atomic orbitals only.

In this approximation $[u_1(\boldsymbol{r}_e)]_{r_0}$ represents the electronic wave function associated with the ground-state energy $E_1(r_0)$ of the H_2^+ ion. The constant r_0 essentially represents the distance between nuclei in that state. Clearly, a better approximation for energy, wave function and internuclear distance in the ground state would be obtained by using a greater number of atomic orbitals.

Instead of increasing the number of atomic orbitals, one can also improve the approximation by introducing an additional variational parameter. Instead of using hydrogen orbitals, let us pick hydrogen-like orbitals associated with a central charge $\alpha|q_e|$ (where α is real but not necessarily an integer):

$$w_a(\boldsymbol{r}_e) = \left(\frac{\alpha^3}{\pi a_0^3}\right)^{1/2} \mathrm{e}^{-\alpha r_a/a_0}, \qquad w_b(\boldsymbol{r}_e) = \left(\frac{\alpha^3}{\pi a_0^3}\right)^{1/2} \mathrm{e}^{-\alpha r_b/a_0}. \tag{20.32}$$

Matrix elements H_{aa} and H_{ab} can be written as

$$H_{aa} = -\frac{e^2}{2a_0}\left\{\langle w_a|a_0^2\nabla_{r_a}^2|w_a\rangle - \frac{2a_0}{r} + \langle w_a|\frac{2a_0}{r_a}|w_a\rangle + \langle w_a|\frac{2a_0}{r_b}|w_a\rangle\right\}, \tag{20.33}$$

$$H_{ab} = -\frac{e^2}{2a_0}\left\{\langle w_a|a_0^2\nabla_{r_b}^2|w_b\rangle - \frac{2a_0}{r}S + 2\langle w_a|\frac{2a_0}{r_a}|w_b\rangle\right\}. \tag{20.34}$$

As in Sect. 17.2,

$$\langle w_a|a_0^2\nabla_{r_a}^2|w_a\rangle = F_1(\alpha) = -\alpha^2\,,$$
$$\langle w_a|\frac{2a_0}{r_a}|w_a\rangle = \frac{1}{2}F_2(\alpha) = 2\alpha\,.$$

Other matrix elements can be evaluated in spheroidal coordinates (Exercise 20.5). Letting $\omega = \alpha r/a_0$ one finds that

$$\langle w_a|w_b\rangle = \mathrm{e}^{-\omega}\left\{1+\omega+\frac{1}{3}\omega^2\right\}, \tag{20.35}$$

$$\langle w_a|\frac{2a_0}{r_b}|w_a\rangle = \frac{2\alpha}{\omega}\left\{1-(1+\omega)\mathrm{e}^{-2\omega}\right\}, \tag{20.36}$$

$$\langle w_a|a_0^2\nabla_{r_b}^2|w_b\rangle = \alpha^2\langle w_a|w_b\rangle - \alpha\langle w_a|\frac{2a_0}{r_a}|w_b\rangle\,, \tag{20.37}$$

$$\langle w_a|\frac{2a_0}{r_a}|w_b\rangle = 2\alpha(1+\omega)\mathrm{e}^{-\omega}\,. \tag{20.38}$$

Thus one gets[7]

---

[7] These expressions are no longer valid if r is large. Indeed the interaction energy between a proton and the dipole it induces in a neighboring hydrogen atom decreases like r^{-4} rather than exponentially. It is inappropriate then to use $1s$ orbitals only.

$$H_{aa} = -\frac{e^2}{2a_0}\left\{-\alpha^2 + 2\alpha\left[1 - \frac{1+\omega}{\omega}\mathrm{e}^{-2\omega}\right]\right\}, \tag{20.39}$$

$$H_{ab} = -\frac{e^2}{2a_0}\mathrm{e}^{-\omega}\left\{\alpha^2\left[-1-\omega+\frac{1}{3}\omega^2\right] + 2\alpha\left[-\frac{1}{\omega}+1+\frac{5\omega}{3}\right]\right\}. \tag{20.40}$$

Substituting (20.39) and (20.40) in (20.28) one gets the eigenvalue E_1 as a function of variables r and α. Figure 20.2 shows that the minimum of E_1 is at $\alpha = \bar{\alpha} = 1.24$ and $r = r_0 = 2.00\,a_0$. One has $E_1(r_0, \bar{\alpha}) = -1.173\,\mathrm{Ry}$. Experimentally one finds ([233], Sect. 26.1) that $r_0 = 2.00\,a_0$ and $E_1 = -1.205\,\mathrm{Ry}$.

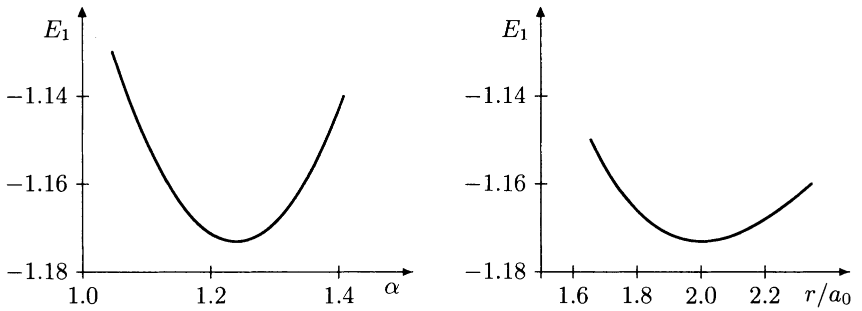

Fig. 20.2. The function $E_1(r, \alpha)$ [Ry] for the H_2^+ ion: left, $E_1(2.00\,a_0, \alpha)$; right, $E_1(r, 1.24)$

20.3 Electronic Terms of Diatomic Molecules

In any diatomic molecule the configuration of nuclei is invariant under rotations about the internuclear axis and under reflection in a plane containing the axis. The symmetry group is therefore $C_{\infty v}$. For identical nuclei there is also invariance under inversion of the coordinates of nuclei. The symmetry group becomes $D_{\infty h}$, which is isomorphic to $C_{\infty v} \otimes C_2$.

Following Sect. 20.1, the electronic Hamiltonian has the same symmetry group as the configuration of nuclei. Corresponding transformations here are rotations of electronic coordinates about the axis, their reflection in a plane containing the axis and, for homonuclear molecules, their spatial inversion (coordinates of nuclei being held fixed).

It is not difficult to obtain the irreducible representations of $C_{\infty v}$ and $D_{\infty h}$. Let us first note that irreducible representations of C_∞, all one-dimensional, have the form $\mathrm{e}^{\mathrm{i}\bar{m}\phi}$ where $\bar{m}$ is an integer and ϕ is the angle of rotation about the symmetry axis. Let $\{M\}$ be an irreducible representation of $C_{\infty v}$. Restricted to the subgroup C_∞ this representation (in a suitable basis) is diagonal, with elements $\mathrm{e}^{\mathrm{i}\bar{m}\phi}$ on the diagonal. Furthermore, under reflection in a plane containing the axis (σ_v), a rotation by $\bar{\phi}$ transforms into a rotation by $-\bar{\phi}$. This means that $\mathrm{e}^{\mathrm{i}\bar{m}\bar{\phi}}$ transforms into $\mathrm{e}^{-\mathrm{i}\bar{m}\bar{\phi}}$. Irreducible representations of $\mathrm{C}_{\infty v}$ therefore have the form

$$M(\bar{\phi}) = \begin{pmatrix} \mathrm{e}^{\mathrm{i}\bar{m}\bar{\phi}} & 0 \\ 0 & \mathrm{e}^{-\mathrm{i}\bar{m}\bar{\phi}} \end{pmatrix}, \qquad M(\sigma_v) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{20.41}$$

Obviously, the representations for $\bar{m}$ and $-\bar{m}$ are equivalent. The case where $\bar{m} = 0$ is an exception, in the sense that the representation is then reducible to two one-dimensional representations:

$$\text{i)} \quad M(\bar{\phi}) = 1\,, \qquad M(\sigma_v) = 1\,; \tag{20.42}$$

$$\text{ii)} \quad M(\bar{\phi}) = 1\,, \qquad M(\sigma_v) = -1\,. \tag{20.43}$$

Denote by i_e the operation that consists in spatially inverting electronic coordinates only. To every irreducible representation of $\mathrm{C}_{\infty v}$ there correspond two irreducible representations of $\mathrm{D}_{\infty h}$, according as $M(i_e) = \pm I$, where I is the identity matrix in the space of one or two dimensions.

These remarks have immediate consequences on the classification of electronic terms of diatomic molecules. The matrix $M(\bar{\phi})$ corresponds to a rotation by an angle $\bar{\phi}$ about the internuclear axis.[8] Hence number $\bar{m}$ corresponds to the eigenvalue of the electronic angular momentum component on this axis. Let $\Lambda = |\bar{m}|$. For $\Lambda = 0, 1, 2, \ldots$, terms are denoted by Σ, Π, Δ, …[9] For $\Lambda \neq 0$ these terms are twofold degenerate.[10] Σ terms, on the other hand, are not degenerate. Because $M(\sigma_v)$ commutes with $M(\bar{\phi})$ if $m = 0$, the Σ terms should be distinguished by the sign of $M(\sigma_v)$. One thus writes Σ^+ and Σ^-. In the specific case of homonuclear molecules all terms must also be distinguished by the sign of $M(i_e)$. An index g or u is used for $M(i_e) = +I$ or $-I$, respectively. One gets such terms as Π_g, Δ_u, Σ_g^+, to write only a few.[11]

In every molecule the total electronic spin operator commutes with the electronic Hamiltonian, inasmuch as spin–orbit interaction is neglected. Electronic terms can thus be classified by the value of the total spin. As in the atomic case one uses the multiplicity $(2S + 1)$, shown as an exponent to the

---

[8] One should not confuse $\bar{\phi}$, the angle of rotation about the internuclear axis, with ϕ, the azimuthal angle of that axis (Sect. 20.4).

[9] Greek letters Σ, Π, Δ, … correspond to Roman letters S, P, D, … of atomic terms.

[10] The degeneracy can be removed by the coupling of electronic angular momentum with the rotation of nuclei. This is called Λ-doubling. See for instance [138], Sect. 88.

[11] The notation g and u comes from "gerade" and "ungerade," German for "even" and "odd." The classification of electronic terms according to the sign of $M(i_e)$ is also largely valid if the nuclei are two different isotopes of the same element. The nuclei then have identical charge and their electrostatic interaction with electrons is similar.

left of the term's main symbol. Thus terms of a diatomic molecule are written as ${}^2\Pi_g$, ${}^1\Delta_u$, ${}^3\Sigma_g^+$, etc.

Empirically, the ground state of the vast majority of diatonic molecules is a term ${}^1\Sigma^+$, or ${}^1\Sigma_g^+$ for homonuclear molecules. The wave function is then invariant under all symmetry transformations of the molecule. Exceptions are NO and O_2 whose ground states are ${}^2\Pi$ and ${}^3\Sigma_g^-$, respectively.

To explicitly construct electronic wave functions of molecules, one uses Roothaan's algebraic approach and the LCAO method. One first writes down a determinantal wave function where each element is a molecular orbital, i.e. the product of a spatial function and a spin vector. Spatial wave functions are taken as linear combinations of atomic orbitals. They belong to irreducible representations of $C_{\infty v}$ or $D_{\infty h}$ and are denoted like molecular terms, with lower-case letters however (π_g, δ_u, σ_g^+, etc.). The electronic wave function (in the absence of configuration interaction) is a linear combination of determinantal functions, corresponding to a specific value of the total electronic spin.

To understand how the molecule's electronic term is obtained from molecular orbitals we will investigate the simplest example, that of the H_2 molecule. Many more are analyzed in [212], [214] and in [233], Chap. 26.

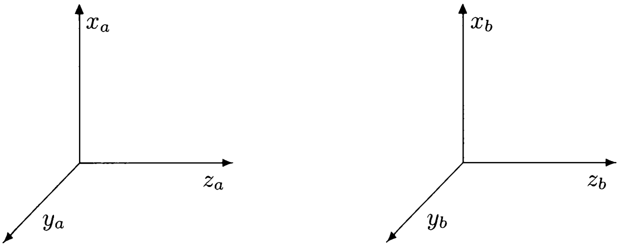

Fig. 20.3. Orientation of coordinate systems centered on nuclei a and b

Figure 20.3 represents two parallel coordinate systems centered on nuclei a and b respectively. Consider an s orbital and a p orbital. Denoting the radial functions by f_s and f_p, one can write the atomic wave functions as[12]

$$\begin{aligned} w^a_{00} &= f_s(r_a)\,, & w^b_{00} &= f_s(r_b)\,, \\ w^a_{11} &= -f_p(r_a)\sin\theta_a\, \mathrm{e}^{\mathrm{i}\phi_a}\,, & w^b_{11} &= -f_p(r_b)\sin\theta_b\, \mathrm{e}^{\mathrm{i}\phi_b}\,, \\ w^a_{10} &= \sqrt{2} f_p(r_a)\cos\theta_a\,, & w^b_{10} &= \sqrt{2} f_p(r_b)\cos\theta_b\,, \\ w^a_{1,-1} &= f_p(r_a)\sin\theta_a\, \mathrm{e}^{-\mathrm{i}\phi_a}\,, & w^b_{1,-1} &= f_p(r_b)\sin\theta_b\, \mathrm{e}^{-\mathrm{i}\phi_b}\,. \end{aligned} \tag{20.44}$$

It is not difficult (Exercise 20.6) to establish the following correspondence between linear combinations of atomic orbitals and molecular orbitals:

---

[12] Note that $\boldsymbol{r}_a$ and $\boldsymbol{r}_b$ depend on one vector only, i.e. $\boldsymbol{r}_e$ in Fig. 20.1. One has $\boldsymbol{r}_a = \boldsymbol{r}_e + \boldsymbol{r}/2$ and $\boldsymbol{r}_b = \boldsymbol{r}_e - \boldsymbol{r}/2$.

$$\begin{aligned} u_1 &= w^a_{00} + w^b_{00} \to \sigma_g^+ , & u_5 &= w^a_{11} + w^b_{11} \\ u_2 &= w^a_{00} - w^b_{00} \to \sigma_u^+ , & u_6 &= w^a_{1,-1} + w^b_{1,-1} \end{aligned} \Bigg\} \to \pi_u ,$$

$$\begin{aligned} u_3 &= w^a_{10} + w^b_{10} \to \sigma_u^+ , & u_7 &= w^a_{11} - w^b_{11} \\ u_4 &= w^a_{10} - w^b_{10} \to \sigma_g^+ , & u_8 &= w^a_{1,-1} - w^b_{1,-1} \end{aligned} \Bigg\} \to \pi_g . \tag{20.45}$$

One can expect that the lowest-energy molecular orbital is a σ_g^+ orbital, built from atomic orbitals much like hydrogen-like $1s$ functions. If two electrons are put in that molecular orbital, their spins must be opposite, so that the term ${}^1\Sigma_g^+$ obtains. The electronic wave function is given by the determinant

$$\begin{aligned} \Phi(1,2) &= \frac{1}{\sqrt{2}} \begin{vmatrix} u_1(\boldsymbol{r}_1)\chi_+(1) & u_1(\boldsymbol{r}_2)\chi_+(2) \\ u_1(\boldsymbol{r}_1)\chi_-(1) & u_1(\boldsymbol{r}_2)\chi_-(2) \end{vmatrix} \\ &= \frac{1}{\sqrt{2}} u_1(\boldsymbol{r}_1) u_1(\boldsymbol{r}_2) \left\{ \chi_+(1)\chi_-(2) - \chi_-(1)\chi_+(2) \right\} . \end{aligned} \tag{20.46}$$

Here $\boldsymbol{r}_1$ stands for the $\boldsymbol{r}_e$ coordinates of the first electron, while $\boldsymbol{r}_2$ stands for the coordinates of the second one.

So much for the lowest-energy term of the hydrogen molecule. The first excited state corresponds to an electron in a σ_g^+ orbital and an electron in a σ_u^+ orbital. Corresponding terms are ${}^1\Sigma_u^+$ and ${}^3\Sigma_u^+$. Their wave functions are easily obtained by means of Slater determinants. Other excited states are similarly obtained by putting the electrons in suitable molecular orbitals.

To get wave functions explicitly one must find the functions f_s and f_p necessary to the expression of atomic orbitals. The latter can be obtained by the variational method. Take for instance the ground state. Let us write the molecular wave function as in (20.46). According to (20.45) the spatial function u_1 is the sum of w^a_{00} and w^b_{00}. For these atomic functions let us pick the hydrogen-like orbitals (20.32) of Sect. 20.2, so that

$$u_1(\alpha; \boldsymbol{r}) = \frac{1}{\sqrt{2(1+S)}} \left(\frac{\alpha^3}{\pi a_0^3} \right)^{1/2} \left\{ \mathrm{e}^{-\alpha r_a / a_0} + \mathrm{e}^{-\alpha r_b / a_0} \right\} . \tag{20.47}$$

The hydrogen molecule's electronic Hamiltonian is given by

$$H = -\frac{\hbar^2}{2m}\nabla_1^2 - \frac{\hbar^2}{2m}\nabla_2^2 + \frac{e^2}{r} - \frac{e^2}{r_{1a}} - \frac{e^2}{r_{1b}} - \frac{e^2}{r_{2a}} - \frac{e^2}{r_{2b}} + \frac{e^2}{r_{12}} , \tag{20.48}$$

where a and b stand for nuclei while 1 and 2 stand for electrons. The trial function can be written as

$$\Phi(\alpha; 1, 2) = \frac{1}{\sqrt{2}} u_1(\alpha; \boldsymbol{r}_1) u_1(\alpha; \boldsymbol{r}_2) \left\{ \chi_+(1)\chi_-(2) - \chi_-(1)\chi_+(2) \right\} . \tag{20.49}$$

The problem thus consists in finding the stationary value of the functional $\langle \Phi(\alpha) | H | \Phi(\alpha) \rangle$. Spin vectors disappear and there remains a certain number of spatial integrals. There are evaluated in [214], Chap. 21, where results of the calculation are also presented.

This, of course, is only a very simple example of the computation of molecular orbitals. A more precise calculation would represent molecular orbitals by means of a larger number of functions and would introduce configuration interaction.

In any case the stability of molecules is first and foremost due to the formation of *bonding* molecular orbitals, i.e. orbitals where the electronic density between the two nuclei is large. In most cases two electrons with opposite spins occupy the orbital, making up what is called a *chemical bond.* Bonds produced by electrons in σ or in π orbitals are called σ bonds and π bonds. In homonuclear molecules electrons have the same probability to be found near anyone of the nuclei. Such a bond is called *covalent.* In heteronuclear molecules, on the other hand, computations show that electrons sometimes have a much larger probability to be found near one nucleus than near the other one. Such a bond is called *ionic.*

20.4 Rotation and Vibration of Diatomic Molecules

In the last two sections we outlined the way to compute electronic energies and wave functions of diatomic molecules. We now turn to the motion of nuclei.

At this stage it is useful to go back to the eigenvalue equation (20.6) of the molecule's Hamiltonian. Separating out center-of-mass coordinates we can write this equation essentially as[13]

$$\left\{ -\frac{\hbar^2}{2M}\nabla_r^2 - \frac{\hbar^2}{2m}\sum_i \nabla_i^2 + V(\boldsymbol{r}, \boldsymbol{r}_i') \right\} \Psi(\boldsymbol{r}, \boldsymbol{r}_i') = E\Psi(\boldsymbol{r}, \boldsymbol{r}_i'). \tag{20.50}$$

Here $\boldsymbol{r}$ is the vector going from one nucleus to the other, M stands for the reduced mass of the nuclei and E is the molecule's total energy in the center-of-mass frame. The prime on electronic coordinates indicates that they are presently defined in a frame at rest.

Let $\boldsymbol{L}_N$ denote the orbital angular momentum of the nuclei. The Laplacian with respect to $\boldsymbol{r}$ can be written as

$$\nabla_r^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r}\right) - \frac{1}{\hbar^2 r^2}\boldsymbol{L}_N \cdot \boldsymbol{L}_N . \tag{20.51}$$

We have $\boldsymbol{L}_N = \boldsymbol{L} - \boldsymbol{L}_e'$, where $\boldsymbol{L}$ and $\boldsymbol{L}_e'$ stand for the molecule's total orbital angular momentum and the total electronic orbital angular momentum, respectively. As in Sect. 20.1 let us write

---

[13] The same approximations as in Sect. 9.1 are made here.

$$\Psi(\boldsymbol{r},\boldsymbol{r}'_i) = v(\boldsymbol{r})\psi(\boldsymbol{r},\boldsymbol{r}'_i)\,, \tag{20.52}$$

where $\psi(\boldsymbol{r},\boldsymbol{r}'_i)$ satisfies (20.7). Substituting (20.52) in (20.50) and using (20.7), one easily obtains

$$\begin{aligned}&\left\{-\frac{\hbar^2}{2Mr^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{2Mr^2}\left[\boldsymbol{L}^2 - 2\boldsymbol{L}'_e\cdot\boldsymbol{L}_N - \boldsymbol{L}'_e\cdot\boldsymbol{L}'_e\right]\right.\\ &\qquad\left. + E'(r) - E\right\} v(\boldsymbol{r})\psi(\boldsymbol{r},\boldsymbol{r}'_i) = 0\,.\end{aligned} \tag{20.53}$$

The term inside the square brackets coincides with $\boldsymbol{L}_N^2$. Due to axial symmetry the function E' depends only on r.

Inasmuch as spin–orbit interaction is neglected, $\boldsymbol{L}$ commutes with the molecule's Hamiltonian. Therefore the wave function $v(\boldsymbol{r})\psi(\boldsymbol{r},\boldsymbol{r}'_i)$ is an eigenfunction of $\boldsymbol{L}^2$ with eigenvalue $\hbar^2L(L+1)$. As in Sect. 20.1 let us multiply (20.53) by $\psi^*(\boldsymbol{r},\boldsymbol{r}'_i)$ and let us integrate on the $\mathrm{d}\boldsymbol{r}'_i$. Once again derivatives of ψ with respect to components of $\boldsymbol{r}$ can be neglected. Since $\psi(\boldsymbol{r},\boldsymbol{r}'_i)$ is normalized for any $\boldsymbol{r}$ one finds that

$$\begin{aligned}&\left\{-\frac{\hbar^2}{2Mr^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\hbar^2L(L+1)}{2Mr^2} + E'(r) - E\right\} v(\boldsymbol{r})\\ &\quad = \frac{1}{2Mr^2}\left\{\int\left[\prod_i \mathrm{d}\boldsymbol{r}'_i\right]\psi^*(\boldsymbol{r},\boldsymbol{r}'_i)\left[2\left(\boldsymbol{L}'_e\psi\right)\cdot\boldsymbol{L}_N + \boldsymbol{L}'_e\cdot\boldsymbol{L}'_e\psi\right]\right\} v(\boldsymbol{r})\,.\end{aligned} \tag{20.54}$$

The electronic wave function ψ is characterized by a quantum number Λ which corresponds to the projection of electronic angular momentum on the axis between nuclei. In a state Λ, the mean value of components of $\boldsymbol{L}'_e$ perpendicular to the axis vanishes (Exercise 20.7). Hence after integration, the term $\psi^*(\boldsymbol{L}'_e\psi)\cdot\boldsymbol{L}_N$ only involves the perpendicular component of $\boldsymbol{L}_N$ along the axis. But the latter vanishes, since only derivatives with respect to the axis's polar and azimuthal angles appear in $\boldsymbol{L}_N$.[14] There remains the term $\psi^*\boldsymbol{L}'_e\cdot\boldsymbol{L}'_e\psi$. Its integral simply produces a function of the coordinate r which adds up to $E'(r)$. Equation (20.54) thus becomes

$$\left\{-\frac{\hbar^2}{2Mr^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - L(L+1)\right] + V(r)\right\} v(\boldsymbol{r}) = Ev(\boldsymbol{r}), \tag{20.55}$$

where $V(r)$ has been defined as

$$V(r) = E'(r) - \frac{1}{2Mr^2}\int\left[\prod_i \mathrm{d}\boldsymbol{r}'_i\right]\psi^*(\boldsymbol{r},\boldsymbol{r}'_i)\boldsymbol{L}'_e\cdot\boldsymbol{L}'_e\psi(\boldsymbol{r},\boldsymbol{r}'_i). \tag{20.56}$$

---

[14] This corresponds to the classical result that the angular momentum of a rigid body made up of two point masses has no component along the axis going from one mass to the other.

We will now derive the angular dependence of $v(\boldsymbol{r})$. For this purpose we introduce coordinates rigidly fixed in the molecule. Let θ and ϕ be the polar and azimuthal angles of the axis. The rigid coordinates $\boldsymbol{r}_i$ are defined as

$$x_i = -x_i' \sin\phi + y_i' \cos\phi \,, \tag{20.57}$$

$$y_i = -x_i' \cos\theta\cos\phi - y_i' \cos\theta\sin\phi + z_i' \sin\theta \,, \tag{20.58}$$

$$z_i = x_i' \sin\theta\cos\phi + y_i' \sin\theta\sin\phi + z_i' \cos\theta \,. \tag{20.59}$$

Clearly the z_i are oriented along the axis. One shows in Exercise (20.8) that the components of $\boldsymbol{L}$ can be written in terms of θ, ϕ and $\boldsymbol{r}_i$ as

$$L_\pm = \mathrm{e}^{\pm\mathrm{i}\phi}\left\{\hbar\left[\pm\frac{\partial}{\partial\theta} + \mathrm{i}\cot\theta\frac{\partial}{\partial\phi}\right] + \frac{1}{\sin\theta}L_{ez}\right\}, \qquad L_z = -\mathrm{i}\hbar\frac{\partial}{\partial\phi}, \tag{20.60}$$

where $L_{ez} = \sum_i L_{iz}$ is the component of $\boldsymbol{L}_e$ along the molecular axis. For an electronic term Λ

$$L_{ez} v(\boldsymbol{r})\psi(\boldsymbol{r},\boldsymbol{r}_i) = v(\boldsymbol{r}) L_{ez}\psi(\boldsymbol{r},\boldsymbol{r}_i) = \bar{m}\hbar v(\boldsymbol{r})\psi(\boldsymbol{r},\boldsymbol{r}_i), \tag{20.61}$$

where $|\bar{m}| = \Lambda$. From (20.60) and (20.61) one shows in Exercise (20.9) that the angular dependence of $v(\boldsymbol{r})$ can be expressed in terms of rotation matrices as[15]

$$v(\boldsymbol{r}) \sim \left[D^{(L)}_{m\bar{m}}(0,\theta,\phi)\right]^* . \tag{20.62}$$

The quantum number L refers to the molecule's total angular momentum. Due to complex conjugation the ϕ dependence of $v(\boldsymbol{r})$ is given by $\mathrm{e}^{\mathrm{i}m\phi}$. Note that $v(\boldsymbol{r})$ depends on the axis's azimuthal angle and does not depend on the angle of rotation about the axis.

The angular dependence of $v(\boldsymbol{r})$ is given by relation (20.62), whereas its radial dependence is determined by (20.55). It is easy to see that (20.55) is equivalent to (10.62) if one lets $\bar{u}(r) = rv(r)$. The whole discussion of Sect. 10.7 about rotational and vibrational energies therefore applies to the present situation.

When the electronic wave function is a Σ term (i.e. $\Lambda = 0$), the angular dependence of $v(\boldsymbol{r})$ is a spherical harmonic. All integral values of L are allowed.

For an electronic term other than Σ (i.e. $\Lambda = |\bar{m}| > 0$), the angular dependence of $v(\boldsymbol{r})$ is a rotation matrix rather than a spherical harmonic. The properties of rotation matrices then entail that values of L must satisfy the inequality $L \geq \Lambda$. That could be expected intuitively. The molecule's orbital angular momentum is the sum of angular momenta of electrons and of nuclei. But on the average $\boldsymbol{L}_N$ and $\boldsymbol{L}_e$ are perpendicular. It should come

---

[15] See also [119].

as no surprise that the total orbital angular momentum is larger than the component of $\boldsymbol{L}_e$ along the molecule's axis.

As mentioned above, our analysis is valid provided that spin–orbit interaction can be neglected. The effect of spin–orbit interaction is studied in detail in [109], Chap. 5 and in [138], Sects. 83–85. We simply quote a few results. We assume that the energy difference between electronic terms is much larger than rotational energies and spin–orbit interaction energies. In the case where rotational energies are much larger than spin–orbit energies, the results just obtained remain valid as a first approximation. The spin–orbit interaction partly removes the degeneracy in a level with a given L. In the opposite case, electronic spin and orbital angular momentum add up to a total angular momentum whose component along the axis can be denoted by Ω. Rotational energies then have a term involving $J(J+1)$ (where $J \geq \Omega$) as well as terms proportional to Ω and to Ω^2.

There are also more complicated intermediate cases where spin–orbit energies are comparable to rotational energies. Actually, rotational energies are larger in small molecules, while spin–orbit energies are larger in molecules made up of big atoms.

20.5 Electric Dipole Transitions

The general structure of a diatomic molecule's energy levels is represented in a simplified way in Fig. 20.4. The two electronic levels are separated by a few eV. Each electronic level splits into vibrational levels. As a first approximation these levels are equidistant. Their separation (of order 0.1 eV) depends however on the electronic level. Each vibrational level subdivides into rotational levels. Their separation (of order 10^{-4} eV) is roughly proportional to J and changes little from one vibrational level to another.

In this section we will investigate possible radiative transitions between levels in Fig. 20.4. This problem is very complicated and, for simplicity, we will entirely neglect electronic and nuclear spins. We will consider electric dipole transitions only. A detailed treatment can be found in [109].

As in atoms, electric dipole transitions in molecules are governed by matrix element (19.61):

$$M_{km}^{\mathrm{E1}} = \mathrm{i}\omega_{km} \hat{A}_{\hat{q}} \cdot \langle E_k | \boldsymbol{D} | E_m \rangle \,.$$

Here $|E_k\rangle$ and $|E_m\rangle$ stand for molecular states and $\boldsymbol{D}$ is the total electric dipole moment operator. In neutral molecules the latter can be written entirely in terms of relative coordinates.

In Sect. 20.1 we saw that total angular momentum and space inversion commute with the Hamiltonian of a molecule. As in Sect. 19.5, one finds that parity must change in an electric dipole transition and that $J_k - J_m = 0$ or ± 1 (except $J_k = 0 = J_m$).

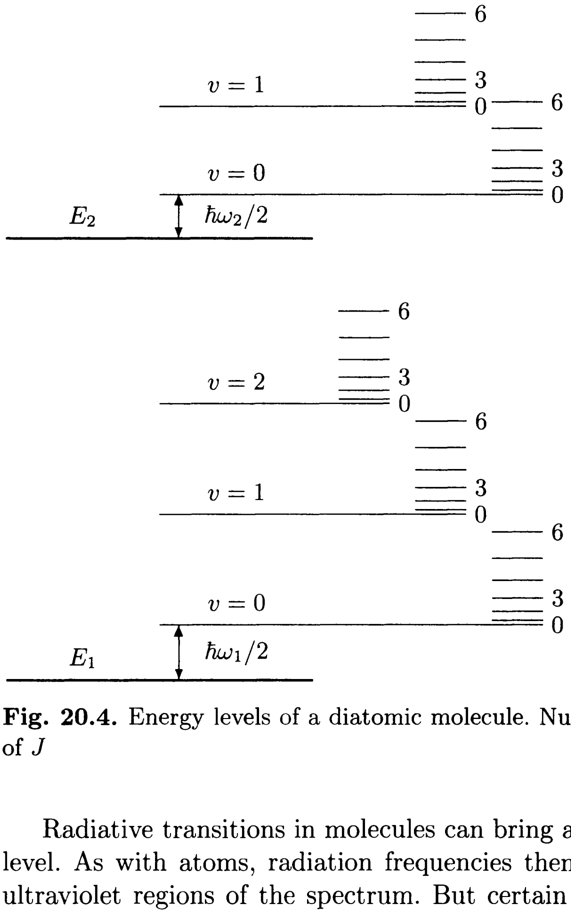

Fig. 20.4. Energy levels of a diatomic molecule. Numbers on the right are values of J

Radiative transitions in molecules can bring about a change of electronic level. As with atoms, radiation frequencies then fall in the visible or near ultraviolet regions of the spectrum. But certain transitions also occur between different rotational or vibrational levels in a given electronic level. They are the ones we will first investigate. Radiation frequencies then fall in the microwave or infrared regions.

Consider a transition between two molecular states with wave functions given in the Born–Oppenheimer approximation by $v_m(\boldsymbol{r})\psi(\boldsymbol{r},\boldsymbol{r}_i)$ and $v_k(\boldsymbol{r})\psi(\boldsymbol{r},\boldsymbol{r}_i)$, respectively. The wave functions of nuclei are different while electronic wave functions are the same. Then

$$\langle E_k|\boldsymbol{D}|E_m\rangle = \int \mathrm{d}\boldsymbol{r}\, v_k^*(\boldsymbol{r}) v_m(\boldsymbol{r}) \int \left[\prod_i \mathrm{d}\boldsymbol{r}_i\right] \psi^*(\boldsymbol{r},\boldsymbol{r}_i)\boldsymbol{D}\psi(\boldsymbol{r},\boldsymbol{r}_i). \quad (20.63)$$

The second integral represents the molecule's total dipole moment for a relative position of nuclei equal to $\boldsymbol{r}$. Since $|\psi|^2$ does not depend on the azimuthal angle $\bar{\phi}$, this moment must (up to a sign) be parallel to $\boldsymbol{r}$. One can thus write

$$\int \left[\prod_i \mathrm{d}\boldsymbol{r}_i\right] \psi^*(\boldsymbol{r},\boldsymbol{r}_i)\boldsymbol{D}\psi(\boldsymbol{r},\boldsymbol{r}_i) = \{D_0 + (r - r_0)D_1\}\,\hat{r}, \quad (20.64)$$

where r_0 is the equilibrium distance between nuclei, D_0 is the total dipole moment at the equilibrium distance and D_1 is the derivative of the dipole moment with respect to the distance between nuclei. Higher-order terms in the series expansion have been neglected.

Homonuclear diatomic molecules (H_2, N_2, O_2, etc.) have no permanent dipole moment. Hence they cannot give rise to radiative transitions within a given electronic level.

Recall that the wave function of nuclei separates into a radial function associated with vibrational motion and an angular function associated with rotation. One can thus write $v_m(\boldsymbol{r}) = v_m(r)v_m(\hat{r})$, and similarly with v_k. From (20.63) and (20.64) one gets

$$\begin{aligned}&\langle E_k|\boldsymbol{D}|E_m\rangle\\&= \int_0^\infty \mathrm{d}r\, r^2 v_k^*(r)\left\{D_0 + (r-r_0)D_1\right\} v_m(r) \int \mathrm{d}\hat{r}\, v_k^*(\hat{r})\hat{r}v_m(\hat{r})\,. \qquad (20.65)\end{aligned}$$

Consider first the case where the radial function does not change, i.e. $v_k(r) = v_m(r)$. One then has a *rotational* transition. The integral over $\mathrm{d}r$ is proportional to the total dipole moment and the integral over $\mathrm{d}\hat{r}$ can be computed if we recall that angular functions coincide with rotation matrices. One finds the selection rule $\Delta J = 0$ or ± 1 (except $0 \to 0$). If $\Lambda = 0$, rotation matrices reduce to spherical harmonics, and it follows that the change of parity forbids the transition $\Delta J = 0$.

Let us now turn to the case where the radial function changes in the transition. Then we have a *vibration-rotation* transition. The angular integral constrains values of ΔJ in the same way as in a rotational transition. In the radial integral the term involving D_0 vanishes due to the orthogonality of radial functions, while the term involving D_1 is given by

$$D_1 \int_0^\infty \mathrm{d}r\, r^2 (r-r_0) v_k^*(r) v_m(r)\,. \qquad (20.66)$$

The radial potential is associated with a harmonic oscillator perturbed by weak anharmonic terms. Assume the latter can be neglected. Then $rv_m(r)$ is an eigenfunction of the harmonic oscillator and (20.66) vanishes, unless the quantum numbers satisfy the relation $v_k - v_m = \pm 1$. Vibration-rotation transitions for which this selection rule applies are called *fundamental.* As a matter of fact, radial functions have more or less significant contributions coming from anharmonic terms in the potential. One thus observes transitions $v_k - v_m = \pm 2, \pm 3, \ldots$, with intensities much weaker than in fundamental transitions.

So much for transitions where only the wave function of nuclei changes. We turn to the case where the electronic wave function also changes. It is then convenient to write

$$\begin{aligned}\langle E_k|\boldsymbol{D}|E_m\rangle &= \int \mathrm{d}\boldsymbol{r}\left[\prod_i \mathrm{d}\boldsymbol{r}_i\right] v_k^*(\boldsymbol{r})\psi_k^*(\boldsymbol{r},\boldsymbol{r}_i)\\ &\quad\times\left\{\sum_j(q_e\boldsymbol{r}_j)+q_N\boldsymbol{r}\right\}\psi_m(\boldsymbol{r},\boldsymbol{r}_i)v_m(\boldsymbol{r})\\ &= \int \mathrm{d}\boldsymbol{r}\, v_k^*(\boldsymbol{r})v_m(\boldsymbol{r})\int\left[\prod_i \mathrm{d}\boldsymbol{r}_i\right]\psi_k^*(\boldsymbol{r},\boldsymbol{r}_i)\sum_j(q_e\boldsymbol{r}_j)\psi_m(\boldsymbol{r},\boldsymbol{r}_i),\end{aligned} \tag{20.67}$$

where the second term disappeared due to the orthogonality of electronic wave functions.

In Sect. 20.1 we pointed out that the change in electronic wave functions due to a change in the coordinates of nuclei is much smaller than corresponding changes in wave functions of nuclei. Hence one can write

$$\begin{aligned}\langle E_k|\boldsymbol{D}|E_m\rangle &= \left\{\int_0^\infty \mathrm{d}r\, r^2 v_k^*(r)v_m(r)\right\}\\ &\quad\times\left\{\int \mathrm{d}\hat{r}\, v_k^*(\hat{r})v_m(\hat{r})\int\left[\prod_i \mathrm{d}\boldsymbol{r}_i\right]\psi_k^*(\boldsymbol{r},\boldsymbol{r}_i)\sum_j(q_e\boldsymbol{r}_j)\psi_m(\boldsymbol{r},\boldsymbol{r}_i)\right\},\end{aligned} \tag{20.68}$$

where it is understood that the integral on $\mathrm{d}\boldsymbol{r}_i$ is evaluated at values of $\boldsymbol{r}$ that correspond to the equilibrium position of nuclei. The first factor on the right-hand side of (20.68) is called the *overlap integral* and depends on wave functions of nuclei only.

Since v_k and v_m are associated with different electronic wave functions, and thus with different potentials, they do not satisfy orthogonality relations. Hence there are no selection rules on v. The value of the overlap integral varies as a function of the potential wells' relative position and the more or less rapid oscillation of functions v_k and v_m.

Selection rules on parity change and ΔJ remain valid, as with every molecule. The integral on $\mathrm{d}\boldsymbol{r}_i$ in the second curly brackets on the right-hand side of (20.68) yields additional rules that we will now develop. For this purpose note that this integral can be written as

$$\begin{aligned}&\int\left[\prod_i \mathrm{d}\boldsymbol{r}_i\right]\psi_k^*(\boldsymbol{r},\boldsymbol{r}_i)U^\dagger U\sum_j(q_e\boldsymbol{r}_j)U^\dagger U\psi_m(\boldsymbol{r},\boldsymbol{r}_i)\\ &\quad=\int\left[\prod_i \mathrm{d}\boldsymbol{r}_i\right]\{U\psi_k(\boldsymbol{r},\boldsymbol{r}_i)\}^*\left\{U\sum_j(q_e\boldsymbol{r}_j)U^\dagger\right\}\{U\psi_m(\boldsymbol{r},\boldsymbol{r}_i)\},\end{aligned} \tag{20.69}$$

where U is an arbitrary unitary operator.

Suppose first that U represents a rotation of electronic coordinates by an angle $\bar{\phi}$ about vector $\boldsymbol{r}$. Then $U\psi_m = \mathrm{e}^{-\mathrm{i}\bar{m}_m\bar{\phi}}\psi_m$ and $U\psi_k = \mathrm{e}^{-\mathrm{i}\bar{m}_k\bar{\phi}}\psi_k$, while components of $\boldsymbol{r}_i$ transform by factors 1, $\mathrm{e}^{\pm\mathrm{i}\bar{\phi}}$. But (20.69) should not depend on $\bar{\phi}$. This is possible only if $\Lambda_k - \Lambda_m = |\bar{m}_k| - |\bar{m}_m| = 0$ or ± 1. This means that transitions like $\Sigma \rightarrow \Pi$, $\Pi \rightarrow \Pi$, $\Pi \rightarrow \Delta$ are possible whereas $\Sigma \rightarrow \Delta$ is forbidden.

Suppose next that U represents a reflection in a plane that contains $\boldsymbol{r}$. Then $U\psi_m = \varepsilon_m\psi_m$ and $U\psi_k = \varepsilon_k\psi_k$, where ε_m and ε_k are equal to ± 1. Let us pick the z axis parallel to $\boldsymbol{r}$. We will show that in every case the x component of (20.69) vanishes. Suppose the reflection plane contains the x axis. Then $Ux_iU^\dagger = x_i$, from which we find that the component vanishes if $\varepsilon_k\varepsilon_m = -1$. If, on the other hand, the reflection is chosen so as to contain the y axis, then $Ux_iU^\dagger = -x_i$, from which we find that the component vanishes if $\varepsilon_k\varepsilon_m = 1$. By a similar argument one shows that the y component of (20.69) vanishes. As for the z component, it vanishes if $\varepsilon_k\varepsilon_m = -1$, since the reflection always contains the z axis. To sum up, (20.69) vanishes if $\varepsilon_k\varepsilon_m = -1$. For states with $\Lambda > 0$ this does not really give rise to a new selection rule, since to each energy there correspond two states with opposite ε. For Σ states, on the other hand, this degeneracy does not exist. The only possible transitions are $\Sigma^+ \leftrightarrow \Sigma^+$ and $\Sigma^- \leftrightarrow \Sigma^-$.

We should emphasize that the foregoing discussion is based on the possibility of factorizing the molecular wave function into an electronic function and a wave function of nuclei, the latter itself factorized into a rotational and a vibrational wave function. Most often this is a good approximation. Nevertheless correlation between different molecular motions produces more or less frequent exceptions to the selection rules we have derived.

Parity and angular momentum selection rules are more general than the wave function factorization. They can, however, be formulated rather simply in terms of factorized functions. As an example let us examine the parity associated with a Σ state. The molecular wave function can be written as

$$\Psi_k(\boldsymbol{r}, \boldsymbol{r}_i) = v_k(r)v_k(\hat{r})\psi_k(\boldsymbol{r}, \boldsymbol{r}_i). \tag{20.70}$$

Under space inversion $\boldsymbol{r} \rightarrow -\boldsymbol{r}$, which means that $r \rightarrow r$, $\theta \rightarrow \pi - \theta$ and $\phi \rightarrow \phi + \pi$. Inertial coordinates $\boldsymbol{r}'_i$ transform into $-\boldsymbol{r}'_i$, which, owing to (20.57)–(20.59), entail that rigid coordinates transform as $x_i \rightarrow x_i$, $y_i \rightarrow -y_i$ and $z_i \rightarrow z_i$. In other words, space inversion transforms rigid coordinates like a σ_v reflection. For a Σ state, $v_k(\hat{r})$ is a spherical harmonic, with parity $(-1)^L$, while $v_k(r)$ is invariant under space inversion. From this one concludes that Σ^+ states have parity $(-1)^L$, whereas Σ^- states have parity $-(-1)^L$. Allowed transitions are easily obtained.

All this applies to general diatomic molecules. Additional rules, however, apply to homonuclear molecules. Somewhat more generally, consider first two nuclei with the same electric charge. The electronic Hamiltonian is then invariant under the inversion of rigid electronic coordinates. From this one immediately derives the selection rule $g \leftrightarrow u$.

For a strict homonuclear molecule, the total Hamiltonian is invariant under exchange of the coordinates of nuclei. To the total wave function is then appended an additional index s or a (for "symmetric" or "antisymmetric"), depending on the eigenvalue associated with the exchange operator being $+1$ or -1. One finds the selection rule $s \leftrightarrow s$ and $a \leftrightarrow a$, which essentially comes from the fact that the $\boldsymbol{r}$ term in (20.67) vanishes. A detailed discussion of this rule can be found in [109], Chaps. 3 and 5 and in [233], Sect. 27.2.

20.6 Polyatomic Molecules

Polyatomic molecules have more than two atoms. Their size spans several orders of magnitude, from the smallest (H_2O, NH_3, CH_4, ...) made of a few atoms only to the largest (proteins, DNA) that can contain millions of atoms. The small ones lend themselves to a more complete theoretical analysis, due of course to the restricted number of particles but also to the fact that the configuration of nuclei often displays a point symmetry.

Specifically, suppose that the equilibrium configuration of nuclei is invariant under a point group G. Electronic wave functions, for a given energy, then make up the space of a (generally irreducible) representation of G. For the vast majority of molecules one observes that the ground state is associated with the identity representation. In other words the ground state has the same symmetries as the equilibrium configuration of nuclei. Recall that the wave function of nuclei satisfies (20.10). The function $E'(\boldsymbol{r}_a)$ represents the eigenvalue associated with the electronic wave function, for nuclei at rest at points $\boldsymbol{r}_a$. The coordinates $\boldsymbol{r}_a$ ($a = 1, 2, \ldots, N$) can be defined with respect to the equilibrium position of the N nuclei. The function E' then has a minimum at $\boldsymbol{r}_a = 0$. Hence its first derivatives vanish, so that

$$E'(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = E'(0, \ldots, 0) + \sum_{a,b} \frac{\partial^2 E'}{\partial \boldsymbol{r}_a \partial \boldsymbol{r}_b} \cdot \delta\boldsymbol{r}_a \delta\boldsymbol{r}_b + \cdots \tag{20.71}$$

where the notation means that second derivatives of E' are taken with respect to each component of vectors $\boldsymbol{r}_a$ and $\boldsymbol{r}_b$. Second derivatives are evaluated at the equilibrium position and their values are such that E' decreases in no direction around $\boldsymbol{r}_a = 0$.

The most general motion of N bound particles resolves into three translations, three rotations and $3N - 6$ vibrations.[16] The formal analysis of the quantum problem is similar to that of the classical problem ([93], Chap. 6;

---

[16] If particles are laid out on a straight line at equilibrium, there are three different translations, two rotations and $3N - 5$ vibrations. This comes from the fact that rotation of a point particle about itself is meaningless.

[136], Chap. 5). New coordinates are defined so as to absorb mass factors and diagonalize the matrix of second derivatives. Six eigenvalues of the matrix vanish, corresponding to translations and rotations. Translational terms in the Hamiltonian separate as in Sect. 9.1. To a large extent rotational terms also separate, and they will soon be examined explicitly. There remain vibrational terms which insofar as higher powers are neglected can be written as

$$H_{\text{vib}} = \frac{1}{2} \sum_{\alpha=1}^{3N-6} \{P_\alpha^2 + \omega_\alpha^2 Q_\alpha^2\} . \tag{20.72}$$

Thus the vibrational Hamiltonian is a sum of $3N - 6$ terms, each one representing an independent harmonic oscillator called a *normal mode of vibration.*

The molecule's total vibrational energy is given by

$$E_{\text{vib}} = \sum_\alpha \hbar\omega_\alpha \left(v_\alpha + \frac{1}{2}\right) , \qquad v_\alpha = 0, 1, 2, \ldots \tag{20.73}$$

The vibrational wave function is the product of N harmonic oscillator wave functions. Note however that (20.73) has to be modified if anharmonic terms in H_{vib} are taken into account.

When the equilibrium configuration of nuclei is invariant under a point group G, normal coordinates make up the space of a (generally reducible) representation of G. This representation can be decomposed into irreducible representations. All normal modes associated with a given irreducible representation have the same frequency. Thus degeneracy of normal frequencies of vibration follows purely from symmetry considerations. The analysis is carried out in detail in [138], Sect. 100.

Let us now examine, in the Hamiltonian of nuclei, terms associated with rotations. As with diatomic molecules rotational and vibrational motions influence each other, particularly if quantum numbers are high. We will however treat the problem in a simple way, by assuming that the molecule's rotation corresponds to the one of a rigid body.

A rigid body is characterized by three orthogonal axes A, B and C about which its principal moments of inertia I_A, I_B and I_C are defined. In quantum mechanics as in classical mechanics, the rotational Hamiltonian is given by ([93], Chap. 5; [136], Chap. 6)

$$H_{\text{rot}} = \frac{J_A^2}{2I_A} + \frac{J_B^2}{2I_B} + \frac{J_C^2}{2I_B} . \tag{20.74}$$

In quantum mechanics J_A, J_B and J_C represent angular momentum operators about the A, B and C axes. These axes are rigidly fixed to the nuclei and thus they rotate with the molecule.

We now evaluate the commutation rules of J_A, J_B and J_C. One should not a priori conclude that they coincide with the commutation rules of angular momenta defined in a fixed reference frame. Introduce three unit vectors $\hat{n}_A$,

$\hat{n}_B$ and $\hat{n}_C$ parallel to the A, B and C axes. Let J_x, J_y and J_z be angular momentum components in the fixed reference frame. One has

$$J_A = \hat{n}_A \cdot \boldsymbol{J}, \qquad J_B = \hat{n}_B \cdot \boldsymbol{J}, \qquad J_C = \hat{n}_C \cdot \boldsymbol{J}. \tag{20.75}$$

Since $\hat{n}_A$, $\hat{n}_B$ and $\hat{n}_C$ rotate with the molecule they do not commute with $\boldsymbol{J}$, but they have with this operator commutation relations identical with the ones of position vectors. Hence[17]

$$\begin{aligned}
[J_A, J_B] &= \left[\sum_{i=1}^{3} n_{Ai} J_i, \sum_{j=1}^{3} n_{Bj} J_j\right] \\
&= \sum_{i,j=1}^{3} \{n_{Ai}\left[J_i, n_{Bj}\right] J_j + n_{Ai} n_{Bj}\left[J_i, J_j\right] \\
&\qquad + \left[n_{Ai}, n_{Bj}\right] J_j J_i + n_{Bj}\left[n_{Ai}, J_j\right] J_i\} \\
&= \mathrm{i}\hbar \sum_{i,j,k=1}^{3} \varepsilon_{ijk}\{n_{Ai} n_{Bk} J_j + n_{Ai} n_{Bj} J_k + n_{Bj} n_{Ak} J_i\} \\
&= \mathrm{i}\hbar\{-(\hat{n}_A \times \hat{n}_B) + (\hat{n}_A \times \hat{n}_B) - (\hat{n}_A \times \hat{n}_B)\} \cdot \boldsymbol{J} \\
&= -\mathrm{i}\hbar \hat{n}_C \cdot \boldsymbol{J} = -\mathrm{i}\hbar J_C\,.
\end{aligned} \tag{20.76}$$

Similarly

$$[J_B, J_C] = -\mathrm{i}\hbar J_A, \qquad [J_C, J_A] = -\mathrm{i}\hbar J_B. \tag{20.77}$$

Thus commutation relations of angular momentum components in the rotating frame differ by a sign from the ones of components in the fixed reference frame. Formally, the ones can be obtained from the others by changing the sign of the three operators.

The argument in Sect. 14.3 shows that rotation matrices are eigenfunctions of $J_A^2 + J_B^2 + J_C^2$ and J_C (say). The eigenvalue of J_C differs by a sign from what it would be in a fixed reference frame.

A rigid body whose three principal moments of inertia are equal is called a *spherical top.* Such are molecules invariant with respect to a cubic group, since the three axes are then transformed into each other by a symmetry transformation. For a spherical top

$$H_{\text{rot}} = \frac{J_A^2 + J_B^2 + J_C^2}{2I_A} = \frac{\boldsymbol{J}^2}{2I_A}. \tag{20.78}$$

Eigenvalues of the rotational Hamiltonian are equal to $\hbar^2 J(J+1)/2I_A$. The degree of degeneracy is equal to $(2J+1)^2$, corresponding to values of indices m and m' in rotation matrices.

---

[17] We assume that the A, B and C axes are right-handedly oriented. For a different way of computing the commutation relations see [96].

A *symmetric top* is a rigid body with two equal principal moments of inertia (say $I_A = I_B \neq I_C$). Such are molecules with a symmetry axis of order higher than two, since then one of the axes transforms into a combination of itself and another one under a symmetry transformation. One has

$$H_{\text{rot}} = \frac{J_A^2 + J_B^2}{2I_A} + \frac{J_C^2}{2I_C} = \frac{\boldsymbol{J}^2}{2I_A} + \left\{ \frac{1}{2I_C} - \frac{1}{2I_A} \right\} J_C^2. \tag{20.79}$$

Eigenvalues of the Hamiltonian are equal to

$$E_{\text{rot}} = \frac{\hbar^2}{2I_A} J(J+1) + \hbar^2 \left\{ \frac{1}{2I_C} - \frac{1}{2I_A} \right\} m^2. \tag{20.80}$$

The m degeneracy is partly removed whereas the m' degeneracy remains as it is.

A rigid body whose three principal moments of inertia are different is called an *asymmetric top*. In that case there is no general formula for eigenvalues of H_{rot}. They must be calculated by diagonalizing H_{rot} in suitable spaces. Examples can be found in [3], Chap. 2 and in [138], Sect. 103.

We will not undertake the study of selection rules valid for polyatomic molecules. The interested reader is referred to the very general analysis of [121].

Exercises

20.1. Carry out the integral over the $\mathrm{d}\boldsymbol{r}_i$ leading to (20.10).

20.2. Use (20.7) to show that

$$E'(\boldsymbol{r}_a) = \int \left[\prod_i \mathrm{d}\boldsymbol{r}_i \right] \psi^*(\boldsymbol{r}_a, \boldsymbol{r}_i) \left\{ -\frac{\hbar^2}{2m} \sum_i \nabla_i^2 + V(\boldsymbol{r}_a, \boldsymbol{r}_i) \right\} \psi(\boldsymbol{r}_a, \boldsymbol{r}_i).$$

From this derive (20.11). [The electronic wave function is normalized for all values of coordinates $\boldsymbol{r}_a$.]

20.3. Derive (20.23) from the stationarity of functional (20.21).

20.4. Derive (20.28) and (20.29).

20.5. Derive (20.35)–(20.38). For this purpose it is convenient to carry out integrals in spheroidal coordinates $(\lambda, \mu, \bar{\phi})$, where $\bar{\phi}$ is the angle about the molecule's axis and where

$$\lambda = \frac{1}{r}(r_a + r_b), \qquad \mu = \frac{1}{r}(r_a - r_b).$$

One has $1 \leq \lambda < \infty$, $-1 \leq \mu \leq 1$ and the Jacobian of the transformation is given by

$$J\left(\frac{x, y, x}{\lambda, \mu, \bar{\phi}}\right) = \frac{r^3}{8}(\lambda^2 - \mu^2)\,.$$

One can make use of the following formula:

$$\int_1^\infty \mathrm{d}\lambda\, \lambda^n \mathrm{e}^{-\omega\lambda} = \frac{n!\mathrm{e}^{-\omega}}{\omega^{n+1}} \sum_{l=0}^{n} \frac{1}{l!}\omega^l\,.$$

20.6. Check that orbitals $u_1, \ldots, u_8$ given in (20.45) have the transformation properties shown.

20.7. Show that in an electronic state Λ the mean value of components of $\boldsymbol{L}_e$ perpendicular to the molecule's axis vanish. Here $\boldsymbol{L}_e$ is defined in terms of rigid coordinates. Explain why, for all practical purposes, the result also holds for the mean value of components of $\boldsymbol{L}'_e$ perpendicular to the axis.

20.8. Let θ and ϕ be the polar and azimuthal angles of a diatomic molecule's axis. In a motionless frame with coordinates $\boldsymbol{r}'_i$, the components of the molecule's total orbital angular momentum can be written as

$$\begin{aligned}
L_+ &= \hbar \mathrm{e}^{\mathrm{i}\phi}\left\{\frac{\partial'}{\partial\theta} + \mathrm{i}\cot\theta\frac{\partial'}{\partial\phi}\right\} + \sum_{i=1}^{Q} L'_{i+}\,, \\
L_- &= \hbar \mathrm{e}^{-\mathrm{i}\phi}\left\{-\frac{\partial'}{\partial\theta} + \mathrm{i}\cot\theta\frac{\partial'}{\partial\phi}\right\} + \sum_{i=1}^{Q} L'_{i-}\,, \\
L_z &= -\mathrm{i}\hbar\frac{\partial'}{\partial\phi} + \sum_{i=1}^{Q} L'_{iz}\,,
\end{aligned}$$

where primed derivatives with respect to θ and ϕ are evaluated at fixed $\boldsymbol{r}'_i$ and where

$$L'_{ix} = -\mathrm{i}\hbar\left\{y'_i\frac{\partial}{\partial z'_i} - z'_i\frac{\partial}{\partial y'_i}\right\}, \quad \text{etc.}$$

Carry out the change of variables (20.57)–(20.59) and show that in terms of variables x_i, y_i and z_i, components L_+, L_- and L_z can be written as

$$\begin{aligned}
L_+ &= \mathrm{e}^{\mathrm{i}\phi}\left\{\hbar\left[\frac{\partial}{\partial\theta} + \mathrm{i}\cot\theta\frac{\partial}{\partial\phi}\right] + \frac{1}{\sin\theta}\sum_{i=1}^{Q} L_{iz}\right\}, \\
L_- &= \mathrm{e}^{-\mathrm{i}\phi}\left\{\hbar\left[-\frac{\partial}{\partial\theta} + \mathrm{i}\cot\theta\frac{\partial}{\partial\phi}\right] + \frac{1}{\sin\theta}\sum_{i=1}^{Q} L_{iz}\right\}, \\
L_z &= -\mathrm{i}\hbar\frac{\partial}{\partial\phi}\,,
\end{aligned}$$

where derivatives with respect to θ and ϕ are now evaluated at fixed $\boldsymbol{r}_i$ and where

$$L_{iz} = -\mathrm{i}\hbar \left\{ x_i \frac{\partial}{\partial y_i} - y_i \frac{\partial}{\partial x_i} \right\}.$$

20.9.

a) Use results of Exercise (20.8) to show that in an electronic state where the z projection of orbital angular momentum is $\bar{m}\hbar$,

$$\begin{aligned} \hbar^{-2} \boldsymbol{L} \cdot \boldsymbol{L} &= \hbar^{-2} \left\{ L_+ L_- - \hbar L_z + L_z^2 \right\} \\ &= -\frac{\partial^2}{\partial \theta^2} - \cot\theta \frac{\partial}{\partial \theta} - \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial \phi^2} + \frac{2\mathrm{i}\bar{m}\cot\theta}{\sin\theta} \frac{\partial}{\partial \phi} + \frac{\bar{m}^2}{\sin^2\theta} . \end{aligned}$$

b) Check that $[D^{(L)}_{m\bar{m}}(0, \theta, \phi)]^*$ is an eigenfunction of $\boldsymbol{L} \cdot \boldsymbol{L}$.

20.10.

a) Use results of Exercise (20.8) to check that the component of total angular momentum along the molecule's axis, L_r, is given by

$$L_r \equiv \sin\theta\cos\phi L_x + \sin\theta\sin\phi L_y + \cos\theta L_z = \sum_{i=1}^{Q} L_{iz} .$$

b) Check that L_+, L_- and L_z, written in terms of coordinates $\boldsymbol{r}_i$, satisfy the commutation relations of angular momentum.

20.11. An electron has a potential energy given by

$$V(\boldsymbol{R}) = V_0 \left(X^2 + Y^2 + \varepsilon Z^2 \right), \quad \text{where} \quad 1 < \varepsilon < 2 .$$

Here V_0 is a constant and $\boldsymbol{R}$ stands for the position operator. Show that the L_z component of the electron's angular momentum commutes with $V(\boldsymbol{R})$. Obviously $V(\boldsymbol{R})$ commutes with reflections in a plane that contains the z axis and with space inversion. Thus this problem shares the symmetries of a homonuclear diatomic molecule. Moreover $V(\boldsymbol{R})$ represents the potential of an anisotropic harmonic oscillator. Eigenfunctions and eigenvalues of the eigenvalue equation of H can easily be determined. Find all eigenfunctions (seven in all) that correspond to the four lowest values of energy. For each value of energy, diagonalize one after another the Hamiltonian's three symmetry operators, and give their eigenvalues. Note that these three operators cannot be diagonalized simultaneously, that is, eigenfunctions of one are not always eigenfunctions of another.

Answers to Selected Exercises

Answers or hints are provided here for most of the exercises given at the end of chapters. Answers proposed are not necessarily unique.

2.1
a) $\langle\phi_1|\phi_2\rangle = 1$, $\langle\phi_2|\phi_3\rangle = -\mathrm{i}$, $\langle\phi_3|\phi_1\rangle = 0$.
b) $c_1 = 1/\sqrt{5}$, $c_2 = 1/\sqrt{2}$, $c_3 = 1$.

2.2
b) $\alpha = 1/\sqrt{3}$, $\beta = 1/\sqrt{2}$.
c) $|\phi_3\rangle = (1/\sqrt{6})\,\{2\mathrm{i}|u_1\rangle - |u_2\rangle + |u_3\rangle\}$.

2.3 $|\langle\phi|\psi\rangle|^2 = 41 < 231 = \langle\phi|\phi\rangle\langle\psi|\psi\rangle$.

2.4 To prove that equality implies proportionality, write $|\phi\rangle = \lambda|\psi\rangle + |\alpha\rangle$, where $\langle\alpha|\psi\rangle = 0$, and show that $|\alpha\rangle = 0$.

2.5 O is not linear and does not have an inverse; P is linear but does not have an inverse.

2.6
a) Evaluate $(O^\dagger\,\{\alpha_1|\phi_1\rangle + \alpha_2|\phi_2\rangle\}\,, |\psi\rangle)$ for α_1, α_2, $|\phi_1\rangle$, $|\phi_2\rangle$ and $|\psi\rangle$ arbitrary.
b) Evaluate $(|\phi\rangle, (O^\dagger)^\dagger|\psi\rangle)$.
c) Evaluate $((O_1O_2)^\dagger|\phi\rangle, |\psi\rangle)$.

2.7 $\langle\phi|O|\psi\rangle = \mathrm{i} - 7 = \langle\psi|O|\phi\rangle^*$.

2.8 Evaluate $\langle\psi|\,\{|\phi\rangle\langle\phi|\}\,|\chi\rangle$ for $|\psi\rangle$ and $|\chi\rangle$ arbitrary.

2.9 Evaluate $[U]^\dagger\,[U]$.

2.10
a) Evaluate $\langle u_k|L|u_l\rangle$.
b) $\mathcal{L}_{kl} = \sum_j L_{kj}L_{jl}$.

2.11

$$\lambda_1 = 0\,, \qquad |\phi_1\rangle = (1/2)\left\{-\sqrt{2}\mathrm{i}|u_1\rangle - \mathrm{i}|u_2\rangle + |u_3\rangle\right\};$$

$$\lambda_2 = 4\,, \qquad |\phi_2\rangle = \sqrt{2/3}\left\{(\mathrm{i}/\sqrt{2})|u_1\rangle + |u_3\rangle\right\};$$

$$\lambda_3 = 4\,, \qquad |\phi_3\rangle = \left\{1/\left(2\sqrt{3}\right)\right\}\left\{-\sqrt{2}\mathrm{i}|u_1\rangle + 3\mathrm{i}|u_2\rangle + |u_3\rangle\right\}.$$

2.12
a)

$$\begin{aligned}
\lambda_1 &= 0\,, & |\phi_1\rangle &= \left(1/\sqrt{2}\right)\{\mathrm{i}|u_1\rangle + |u_3\rangle\}\,;\\
\lambda_2 &= 1\,, & |\phi_2\rangle &= |u_2\rangle\,;\\
\lambda_3 &= 2\,, & |\phi_3\rangle &= \left(1/\sqrt{2}\right)\{-\mathrm{i}|u_1\rangle + |u_3\rangle\}\,.
\end{aligned}$$

2.13
a)

$$\begin{aligned}
\lambda_1 &= -1\,, & |\phi_1\rangle &= (1/2)\left\{-\sqrt{2}\mathrm{i}|u_1\rangle + |u_2\rangle - |u_3\rangle\right\};\\
\lambda_2 &= 1\,, & |\phi_2\rangle &= (1/2)\left\{\sqrt{2}|u_1\rangle - \mathrm{i}|u_2\rangle + \mathrm{i}|u_3\rangle\right\};\\
\lambda_3 &= 2\,, & |\phi_3\rangle &= \left(1/\sqrt{2}\right)\{|u_2\rangle + |u_3\rangle\}\,.
\end{aligned}$$

2.14

$$\begin{aligned}
\lambda_1 &= E - |a|\,, & |\phi_1\rangle &= \left(1/\sqrt{2}\right)\left\{\frac{a}{|a|}|u_1\rangle - |u_2\rangle\right\};\\
\lambda_2 &= E + |a|\,, & |\phi_2\rangle &= \left(1/\sqrt{2}\right)\left\{\frac{a}{|a|}|u_1\rangle + |u_2\rangle\right\}.
\end{aligned}$$

2.15

$$\begin{aligned}
\lambda_1 &= E\,, & |\phi_1\rangle &= \left(1/\sqrt{2}\right)\{|u_1\rangle - |u_3\rangle\}\,;\\
\lambda_2 &= E + \sqrt{2}a\,, & |\phi_2\rangle &= (1/2)\left\{|u_1\rangle + \sqrt{2}|u_2\rangle + |u_3\rangle\right\};\\
\lambda_3 &= E - \sqrt{2}a\,, & |\phi_3\rangle &= (1/2)\left\{|u_1\rangle - \sqrt{2}|u_2\rangle + |u_3\rangle\right\}.
\end{aligned}$$

2.16 Direct verification.

2.17 Direct verification.

2.18
a) Check the formula for $n = 1$, and differentiate the formula for arbitrary $n - 1$.
b) Expand $F(1)$ in a Taylor series around $t = 0$.

2.19 The lemma is true if B is unitary. It is not true in general, as can be seen by letting

$$A=\begin{pmatrix}1&0\\0&1\end{pmatrix}, \qquad B=\begin{pmatrix}1&2\\0&3\end{pmatrix}.$$

2.20 The matrices make up a complete set of commuting operators. Here are common eigenvectors:

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1\\-\mathrm{i}\\0\\0\end{pmatrix}, \quad \frac{1}{\sqrt{2}}\begin{pmatrix}1\\\mathrm{i}\\0\\0\end{pmatrix}, \quad \frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\1\\1\end{pmatrix}, \quad \frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\1\\-1\end{pmatrix}.$$

2.21

a) Common eigenvectors are:

$$\begin{pmatrix}1\\0\\0\end{pmatrix}, \quad \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\1\end{pmatrix}, \quad \frac{1}{\sqrt{2}}\begin{pmatrix}0\\1\\-1\end{pmatrix}.$$

b) $\{[H],[K]\}$.

3.1

a) Possible results: -1, 1, 2; probabilities: $1/4$, $1/4$, $1/2$.

b) $\langle A\rangle = 1$; $\Delta A = \sqrt{3/2}$.

3.2

b) $P(A\to 0) = 1/6$; $P(A \to 6) = 5/6$; $\langle A\rangle = 5$; $\Delta A = \sqrt{5}$.

3.3

$$\begin{aligned}\mathrm{Var}(H) &= \frac{|C_1|^2E_1^2+|C_2|^2E_2^2}{|C_1|^2+|C_2|^2} - \left\{\frac{|C_1|^2E_1+|C_2|^2E_2}{|C_1|^2+|C_2|^2}\right\}^2 \\ &= \frac{|C_1|^2|C_2|^2}{\{|C_1|^2+|C_2|^2\}^2}(E_1-E_2)^2\,.\end{aligned}$$

3.4 $|\psi(t)\rangle = \exp(-\mathrm{i}Et/\hbar)\,\{\cos(at/\hbar)|u_1\rangle + \mathrm{i}\sin(at/\hbar)|u_2\rangle\}$.

3.5

a) $|\psi(t=0)\rangle = (1/\sqrt{2})\,\{|u_2\rangle - |u_3\rangle\}$.

b) $P(K\to -b; t) = 1$.

3.6

a) $|\psi(t)\rangle = (1/\sqrt{2})\,\{\exp(-\mathrm{i}Et/\hbar)|u_1\rangle - \exp(\mathrm{i}Et/\hbar)|u_2\rangle\}$.

b) $\langle\psi(t)|K|\psi(t)\rangle = b/2$.

3.7

b) $|\psi(t)\rangle = -(1/\sqrt{2})\exp(-8\mathrm{i}t/\hbar)|E_1\rangle + (1/\sqrt{2})\exp(4\mathrm{i}t/\hbar)|E_3\rangle$.

c) $\langle H\rangle = 2$; $\Delta H = 6$.

3.8

a) $\eta_1 = 1$, $|\eta_1\rangle = |E_1\rangle$; $\eta_2 = 1$, $|\eta_2\rangle = |E_2\rangle$; $\eta_3 = 0$, $|\eta_3\rangle = |E_3\rangle$.

c)

$$[K] = \frac{1}{4}\begin{pmatrix} 2 & -\sqrt{2} & \sqrt{2} \\ -\sqrt{2} & 3 & 1 \\ \sqrt{2} & 1 & 3 \end{pmatrix}.$$

3.9

b) $|\psi(t)\rangle = (1/\sqrt{2})\exp(-\mathrm{i}E't/\hbar)\,\{\exp(-\mathrm{i}bt/\hbar)|E_2\rangle + \exp(\mathrm{i}bt/\hbar)|E_3\rangle\}$.

c) $\langle H\rangle = E'$; $\Delta H = |b|$.

3.10

a) 2, 1, −1.

b) $P(A \to 2; t_0) = 1/2$.

c) $\langle A\rangle = 1$

d) Several possibilities.

4.2

b) Set $\alpha A = \mathrm{i}\delta\theta[I_y]$ and $\beta B = \mathrm{i}\delta\theta[I_x]$ in (a).

4.3 Note that (I_x, I_y, I_z) is irreducible if and only if (I_+, I_-, I_z) is.

4.9

a) Let $\bar{\theta}$ and $\bar{\phi}$ be the polar and azimuthal angles of $\hat{u}$. Then

$$|1;\hat{u}\rangle = \cos^2\frac{\bar{\theta}}{2}\mathrm{e}^{-\mathrm{i}\bar{\phi}}|1\rangle + \frac{1}{\sqrt{2}}\sin\bar{\theta}|0\rangle + \sin^2\frac{\bar{\theta}}{2}\mathrm{e}^{\mathrm{i}\bar{\phi}}|-1\rangle\,,$$

$$|0;\hat{u}\rangle = -\frac{1}{\sqrt{2}}\sin\bar{\theta}\mathrm{e}^{-\mathrm{i}\bar{\phi}}|1\rangle + \cos\bar{\theta}|0\rangle + \frac{1}{\sqrt{2}}\sin\bar{\theta}\mathrm{e}^{\mathrm{i}\bar{\phi}}|-1\rangle\,,$$

$$|-1;\hat{u}\rangle = -\sin^2\frac{\bar{\theta}}{2}\mathrm{e}^{-\mathrm{i}\bar{\phi}}|1\rangle + \frac{1}{\sqrt{2}}\sin\bar{\theta}|0\rangle - \cos^2\frac{\bar{\theta}}{2}\mathrm{e}^{\mathrm{i}\bar{\phi}}|-1\rangle\,.$$

b) 1/4, 1/2, 1/4.

5.1 Integrate by parts to obtain the first identity. To obtain the second one, take $f(x) = F(x)(x - x_0)^2$.

5.2 Δ_ε can represent a delta function, but not its derivatives.

5.3 Evaluate $\int \mathrm{d}p\, I|p\rangle\langle p|I$ where I is represented as in (5.1).

5.4 Evaluate the (xx') matrix element of (5.18) and use properties of the delta function.

5.5 First show, by induction, that $\langle x|P^n|\psi(t)\rangle = \left(-\mathrm{i}\hbar\frac{\partial}{\partial x}\right)^n \psi(x,t)$. To obtain the second equality, use the Fourier transform.

5.7 $\langle\psi|X|\psi\rangle = 0 = \langle\psi|P|\psi\rangle$.

5.8

$$A = \left\{a + \frac{1}{2k}\sin(2ka) + \frac{1}{\kappa}\cos^2(ka)\right\}^{-1/2} = \left\{\frac{\kappa}{1+\kappa a}\right\}^{1/2}.$$

5.9 -0.9796; -0.8176.

5.11 For $x > a$ write $\psi(x) = A_+\mathrm{e}^{ikx} + B_+\mathrm{e}^{-ikx}$, where $k^2 = 2mE/\hbar^2$.

5.12
c) $(k-k')^2/(k+k')^2$.

5.14 Substitute functions in the differential equation, with appropriate values of E.

5.15

$$\int_{-\infty}^{\infty} \mathrm{d}\xi\, H_m(\xi)H_n(\xi)\xi\mathrm{e}^{-\xi^2} = \begin{cases} \sqrt{\pi}2^n(n+1)! & \text{if } m = n+1, \\ 0 & \text{if } m \neq n \pm 1, \\ \sqrt{\pi}2^{n-1}n! & \text{if } m = n-1. \end{cases}$$

5.16
c) $\Delta X = \{(2n+1)\hbar/2\omega m\}^{1/2}$, $\Delta P = \{(2n+1)m\hbar\omega/2\}^{1/2}$.

5.19
a) Use the definition of the propagator and properties of the evolution operator.
b) Replace U by its value and carry out the integral.

5.20 Use properties of the evolution operator.

6.1 The solid angle corresponding to $\lambda_z > 0$ is 2π. The region where ξ is 1 is proportional to $\pi - \theta'$, whereas the region where ξ is -1 is proportional to θ'.

6.2 Expand $\psi(x_0, t_0)$ in terms of harmonic oscillator eigenfunctions and pinpoint the most important coefficient.

7.1
b) Express $\langle\phi|H|\psi\rangle - \langle\psi|H|\phi\rangle^*$ as the integral of a divergence, and transform it into a surface integral.

7.7 Separate terms that depend on r from those that depend on θ and ϕ.

7.10 $\Delta R = \sqrt{3} a_0 / 2Z$.

7.11
b) $\langle R^2 \rangle = 3a_0^2/Z^2$, $\langle \boldsymbol{P} \cdot \boldsymbol{P} \rangle = (\hbar Z/a_0)^2$, $\Delta X \Delta P_x = \hbar/\sqrt{3}$.

7.12 $E = \hbar\omega \{n_x + n_y + n_z + 3/2\}$ where n_x, n_y and n_z are nonnegative integers.

7.14
b) Estimate the mean value of $\mathcal{L}_k$ and the mean value of $q(\boldsymbol{R} \cdot \boldsymbol{B})X_k$ in a state with typical atomic size. The ratio of the latter to the former is of order 10^{-5}.

7.16 Evaluate $\frac{\partial}{\partial t}|\psi|^2$ by means of the Schrödinger equation for a particle in an electromagnetic field.

8.2 Write $\psi(\xi) = f(\xi)\exp\{\pm\sqrt{-\varepsilon}\,\xi\}$ and show that the equation for $f(\xi)$ has a solution with leading term ξ^n only if $n = 0$.

8.4
b) Show that $\mathrm{d}\psi/\mathrm{d}\xi$ is an odd function and use (a).

9.2 $\lambda_H/\lambda_D = 1.000\,272$; $\lambda_H/\lambda_T = 1.000\,362$.

9.3 Approximately 4.

9.4
a) The equation separates in cartesian coordinates and the solutions are products of sines.
b) See what the possible values are for n_x, n_y, n_z.

9.6
a) $A = 144$, $n = 3$.

9.7
a)

$$\frac{\varrho}{q_e} = \frac{1}{3\pi^2}\left\{\frac{2Z\Phi}{a_0\beta\xi}\right\}^{3/2}.$$

9.8 The first integral is bounded and the second one vanishes if r is large enough.

10.1 $H_0|\chi_k\rangle + W|\chi_{k-1}\rangle = \sum_{p=0}^{k} \varepsilon_p |\chi_{k-p}\rangle$.

10.2 Project (10.11) on $\langle E_j^\beta|$, where $j \neq n$.

10.3 Corrections of eigenvalues vanish. Corrections of vectors $|E_1\rangle$ and $|E_2\rangle$ are different from zero.

10.4
a) In the degenerate sector one has

$$|\chi_0^1\rangle = \frac{1}{\sqrt{2}}\left\{|E_1\rangle + \frac{a^*}{|a|}|E_2\rangle\right\}, \quad |\chi_1^1\rangle = \frac{1}{E_1 - E_3}\left\{\frac{b^*}{\sqrt{2}}|E_3\rangle + \frac{|b|^2}{4|a|}|\chi_0^2\rangle\right\},$$

$$|\chi_0^2\rangle = \frac{1}{\sqrt{2}}\left\{|E_1\rangle - \frac{a^*}{|a|}|E_2\rangle\right\}, \quad |\chi_1^2\rangle = \frac{1}{E_1 - E_3}\left\{\frac{b^*}{\sqrt{2}}|E_3\rangle - \frac{|b|^2}{4|a|}|\chi_0^1\rangle\right\}.$$

b) Solve the secular equation approximately, keeping second-order terms in a and b. Eigenvalues then coincide with those found by perturbation theory.

10.5
c) 51.

10.7
b)

$$\begin{aligned}
&\langle 300|W|310\rangle = \langle 310|W|300\rangle = a\,;\\
&\langle 310|W|320\rangle = \langle 320|W|310\rangle = b\,;\\
&\langle 311|W|321\rangle = \langle 321|W|311\rangle = c\\
&\qquad = \langle 31{-}1|W|32{-}1\rangle = \langle 32{-}1|W|31{-}1\rangle\,.
\end{aligned}$$

c) Order rows and columns suitably; the matrix breaks down into blocks with dimensions not larger than 3 × 3.
d) Three levels are unchanged; two levels are shifted by $+c$, two more by $-c$; the last two are shifted by $\pm\sqrt{a^2+b^2}$.

10.8 Use the binomial expansion up to second-order terms.

10.9 Pick $\hat{r} = \hat{z}$ and make use of spherical coordinates.

10.11 The value computed up to second order coincides with the exact value.

10.12 Write X^3 and X^4 in terms of creation and annihilation operators.

11.3 Write down matrix elements of equation $(\lambda I - H)G(\lambda) = I$.

11.5

$$F(\theta) = \frac{3Z}{(a|\Delta \boldsymbol{k}|)^3}\left\{\sin(a|\Delta \boldsymbol{k}|) - a|\Delta \boldsymbol{k}|\cos(a|\Delta \boldsymbol{k}|)\right\}.$$

11.6
a) $F(\theta) = Z\{1 + a^2|\Delta \boldsymbol{k}|^2\}^{-2}$.

11.7
a) $f(\theta,\phi) = -(2m/\hbar^2)V_0a^3\{1 + a^2|\Delta \boldsymbol{k}|^2\}^{-1}$.

11.9
a) $\bar{u}_l(r) = rj_v(kr)$ where $v(v+1) = l(l+1) + A$.
b) $\delta_l = (\pi/2)(l - v)$.

11.11 Write (11.77) in terms of h_l^+ and h_l^-.

11.12
b) Write $\Delta \boldsymbol{k} = q_a\boldsymbol{A} + q_b\boldsymbol{B} + q_c\boldsymbol{C} + \boldsymbol{D}$ and show that $\boldsymbol{D} = 0$.

12.1
a) One can write at any time t,

$$\mathrm{i}\hbar\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle - H|\psi(t)\rangle = \sum_i c_i|u_i\rangle\,,$$

where $\{|u_i\rangle\}$ is an orthonormal basis containing $|\psi\rangle$. Show that all coefficients but one vanish.

12.2

$$\frac{\mathrm{d}\varrho}{\mathrm{d}t} = \sum_n \frac{\mathrm{d}w_n}{\mathrm{d}t}\left(|\psi_n(t)\rangle\langle\psi_n(t)|\right) + \frac{1}{\mathrm{i}\hbar}\left[H, \varrho(t)\right].$$

12.4
a) $(1/2)I$.
b) Same result.

12.5
b) All mean values vanish except $\mathrm{Tr}(\varrho_1\sigma_y) = 1$.

12.6 Use the definition of derivative.

12.7 Show that each subsystem obeys the equation given in Exercise (12.1a) and use the result of that exercise.

12.8 Expand $|\chi(1,2)\rangle$ as in (12.26), picking the $|u_i\rangle$ and $|v_k\rangle$ as eigenvectors of $H(1)$ and $H(2)$ respectively.

12.11
a) Evaluate $\mathrm{Tr}\{A\varrho_0A^\dagger\}$ in a basis where ϱ_0 is diagonal.
b) Use properties of the trace and of the evolution operator.

12.13 Note that $\varrho_0 = (I + \boldsymbol{\sigma}\cdot\hat{n}_0)/2$, $\bar{P}_1' = (I + \boldsymbol{\sigma}\cdot\hat{n}_1)/2$, etc.

13.1 Invariance under translations or space inversion is straightforward. To show that the Laplacian is invariant under rotations, find the transformation of $\partial/\partial x_i$. Invariance of the delta function is proved by integrating the product of the right-hand side and an arbitrary function $f(\boldsymbol{r})$.

13.3
b) The representation is reducible.

13.4 Show that if a subspace is invariant under the action of all matrices $M(g)$, so is its orthogonal complement.

13.5 C_{2v}, $\mathrm{D}_{\infty h}$, C_{3v}.

13.6
b) D_4.
c) D_{4h}.

13.8
c) Reckon constraints following from the definition of an orthogonal matrix with unit determinant.

13.10
c) The condition on the determinant reduces the number of parameters by one, in contrast with orthogonal matrices.

13.12
b) Write M as

$$\frac{1}{2}\left\{M + M^T - \frac{2}{3}\mathrm{Tr}(M)I\right\} + \frac{1}{2}\left\{M - M^T\right\} + \frac{1}{3}\mathrm{Tr}(M)I.$$

13.13
b) $[T_1, T_2] = 2\mathrm{i}T_2$, $[T_2, T_3] = \mathrm{i}T_1$, $[T_3, T_1] = 2\mathrm{i}T_3$.

13.18
a) The unit element is (I, O) and the inverse of $(O, \boldsymbol{a})$ is $(O^{-1}, -O^{-1}\boldsymbol{a})$.

14.1 Write the constraints on matrix elements explicitly; solve the resulting equations, keeping in mind the form of $O(\alpha, \beta, \gamma)$.

14.4 Expand binomials and rename summation indices.

14.6
a) Carry out the product of matrices and identify the appropriate element on both sides of the equation.
b) Evaluate $D^l_{00}(\alpha, \beta, \gamma)$.

14.7 Use the Levi-Civita symbol.

14.8 Write $\boldsymbol{S}_1 \cdot \boldsymbol{S}_2$ in terms of $\boldsymbol{S}^2$ and I.

14.12 Write $\boldsymbol{J}_1 \cdot \boldsymbol{J}_2$ in terms of J_{1+}, J_{1-}, etc.

14.15 Note that $4j$ is even and that $j - m$ is an integer.

14.19 The operator of rank 2 is like (14.90) for $\boldsymbol{P}_1 = \boldsymbol{P}_2$.

14.21
b) Equation $L_i\psi = 0$ implies that ψ only depends on r; use this result to solve the other equation.

15.2 Note that if $A = -\mathrm{i}\sum_j a^j \Sigma_{0j}$, then $A^3 = \boldsymbol{a}^2 A$. Moreover, $a^i a^j/(\boldsymbol{a}^2) = v^i v^j/(\boldsymbol{v}^2)$.

15.3
a) Note that $(\sum_j a^j \sigma_{0j})^2 = -\boldsymbol{a}^2 I$.

15.4 Use the identity (15.80).

15.9 Make use of the fact that $\partial^2\chi_1/\partial t^2 \sim (-\mathrm{i}E/\hbar)\partial\chi_1/\partial t$, after justification.

16.1
b)

$$\frac{\mathcal{D}x}{\sqrt{V(x)}} = \lim_{n\to\infty}\left[2\pi\mathrm{i}\hbar\varepsilon V\left(\frac{x+x_n}{2}\right)\right]^{-1/2} \times \prod_{k=1}^{n}\left\{\left[2\pi\mathrm{i}\hbar\varepsilon V\left(\frac{x_k+x_{k-1}}{2}\right)\right]^{-1/2}\mathrm{d}x_k\right\}.$$

16.3 Assume (16.29) is true for $p-1$, and carry out the integral on $\mathrm{d}x_p$.

16.5
b) Convince yourself first that $(\lambda_n^{(1)} - \lambda_n^{(2)})/(\lambda_n^{(1)} + \lambda_n^{(2)}) \to 0$ as $n \to \infty$.

16.6 Show that the last two lines in (16.51) coincide if (16.49) is substituted in them.

16.9 $E/V_0 = -0.8093$; -0.6032; -0.4423; -0.3020; -0.1747; -0.0566.

16.10 There are an infinite number of bound states only if the value of n corresponding to a null value of E is infinite.

16.12
a) Use the fact that $\psi(x) = 0$ for $x < 0$.

16.13

$$E = \left\{\frac{3}{2}\left(n - \frac{1}{4}\right)\pi\hbar\frac{1}{\sqrt{2m}}\frac{V_0}{a}\right\}^{2/3}.$$

17.1 The correction term is $\eta^2\{\langle\phi_1|H|\phi_1\rangle - E\}$.

17.2 The variational method yields $E_0 \leq -V_0/32$.

17.3

a) $E_0 \leq \hbar\omega/\sqrt{2}$. Note that the trial function has a discontinuous derivative at $x = 0$. Its second derivative has therefore a term proportional to $\delta(x)$.

b) $E_0 \leq \hbar\omega/2$.

17.4

$$E_{\text{min}} \sim \frac{3}{2}\frac{\hbar^2}{2m}\left(\frac{3\lambda}{8}\frac{2m}{\hbar^2}\right)^{1/3}.$$

17.9 Examine the equation obtained by multiplying (17.47) and $u_k^*(\boldsymbol{r}_1)$ and by integrating on $\mathrm{d}\boldsymbol{r}_1$.

17.11 Write the atom's energy and the ion's energy by means of (17.41).

17.13 Use (14.82) to compute the product of two spherical harmonics or Legendre polynomials.

18.7 Here the second term in (18.25) is absent.

18.13 One can use either $|\gamma; L, S; M_L, M_S\rangle$ or $|\gamma; L, S; J, M\rangle$ as unperturbed eigenvectors, but the second choice may be simpler.

19.2

a)

$$-\frac{q_e}{m}\boldsymbol{A}\cdot\boldsymbol{P} \to \mathrm{i}q_e E_0 \tau \omega_{km} \mathrm{e}^{-t/\tau} Z.$$

b) Only the probability that the atom is in state $(2p, m = 0)$ does not vanish.

19.6 In the matrix element, introduce the factor $\mathcal{R}^{-1}\mathcal{R}$, where $\mathcal{R}$ is the operator of a rotation by π about $\hat{q}$.

19.7 10^3 W m^{-2}.

19.10 Use the result of Exercise (14.18).

19.11 $(2^{17}\omega_{km}^3 e^2 a_0^2)/(3^{11}\hbar c^3)$.

20.6 Note that i_e transforms $\boldsymbol{r}_a$ into $-\boldsymbol{r}_b$ and $\boldsymbol{r}_b$ into $-\boldsymbol{r}_a$; a reflection in a plane that contains the axis leaves r_a, r_b, θ_a and θ_b invariant.

References

Numbers in square brackets after each entry refer to sections where books and papers are mentioned. An asterisk indicates more than one reference to a book or paper in the same section. P and UC mean "Preface" and "Units and Constants" while 11e, for instance, means "Exercises in Chap. 11." Several monographs not specifically referred to are given as general references.

1. E. A. Abers, B. W. Lee: "Gauge Theories," Phys. Rep. **9C**, 1 (1973). [16.1]
2. Y. Aharonov, D. Bohm: "Significance of Electromagnetic Potentials in the Quantum Theory," Phys. Rev. **115**, 485 (1959). [7.8]
3. H. C. Allen, Jr., P. C. Cross: *Molecular Vib-rotors* (Wiley, New York 1963). [20.6]
4. H. L. Anderson, ed.: *Physics Vade Mecum* (American Institute of Physics, New York 1981).
5. R. A. Arndt, L. D. Roper: "Scattering in a Spherical Potential: Motion of Complex-plane Poles and Zeros," Comp. in Phys. **3** (March/April), 65 (1989). [11.8]
6. A. Aspect, J. Dalibard, G. Roger: "Experimental Test of Bell's Inequalities Using Time-varying Analyzers," Phys. Rev. Lett. **49**, 1804 (1982). [21.2, 21.3]
7. A. Aspect, P. Grangier, G. Roger: "Experimental Tests of Realistic Local Theories via Bell's Theorem," Phys. Rev. Lett. **47**, 460 (1981). [21.2]
8. L. E. Ballentine: "The Statistical Interpretation of Quantum Mechanics," Rev. Mod. Phys. **42**, 358 (1970). [6.2]
9. L. E. Ballentine: "Probability Theory in Quantum Mechanics," Am. J. Phys. **54**, 883 (1986). [3.2]
10. L. E. Ballentine: "Resource letter IQM-2: Foundations of Quantum Mechanics Since the Bell Inequalities," Am. J. Phys. **55**, 785 (1987).
11. L. E. Ballentine, ed.: *Foundations of Quantum Mechanics Since the Bell Inequalities, Selected Reprints* (American Association of Physics Teachers, College Park 1988).
12. L. E. Ballentine: "What Do We Learn About Quantum Mechanics from the Theory of Measurement?" Int. J. Theor. Phys. **27**, 211 (1988). [21.5]
13. L. E. Ballentine: *Quantum Mechanics* (Prentice-Hall, Englewood Cliffs 1990). [5.9.2, 6.3, 13.6, 21.2]
14. L. E. Ballentine, J. P. Jarrett: "Bell's Theorem: Does Quantum Mechanics Contradict Relativity?" Am. J. Phys. **55**, 696 (1987). [21.3]
15. C. N. Banwell: *Fundamentals of Molecular Spectroscopy*, 3rd edn. (McGraw-Hill, London 1983). [10.7, 19.6]

16. R. M. Barnett, C. D. Carone, D. E. Groom, T. G. Trippe, C. G. Wohl, B. Armstrong, P. S. Gee, G. S. Wagman, F. James, M. Mangano, K. Mönig, L. Montanet, J. L. Feng, H. Murayama, J. J. Hernández, A. Manohar, M. Aguilar-Benitez, C. Caso, R. L. Crawford, M. Roos, N. A. Törnqvist, K. G. Hayes, K. Hagiwara, K. Nakamura, M. Tanabashi, K. Olive, K. Honscheid, P. R. Burchat, R. E. Shrock, S. Eidelman, R. H. Schindler, A. Gurtu, K. Hikasa, G. Conforto, R. L. Workman, C. Grab, C. Amsler: "Review of Particle Physics," Phys. Rev. D **54**, 2 (1996). [14.6]
17. J. L. Basdevant, J. Dalibard: *The Quantum Mechanics Solver. How to Apply Quantum Theory to Modern Physics* (Springer-Verlag, Berlin 2000).
18. J. L. Basevant, J. Dalibard: *Quantum Mechanics*, including a CD-ROM by M. Joffre (Springer-Verlag, Berlin 2002).
19. A. O. Barut, R. Raczka: *Theory of Group Representations and Applications*, 2nd edn. (Polish Scientific Publishers, Warsaw 1980). [13]
20. J. L. Basdevant: *Mécanique quantique* (Marketing, Paris 1986).
21. W. E. Baylis, G. W. F. Drake: "Units and Constants." In: [71], pp. 1–5. [UC*]
22. J. S. Bell: "On the Einstein Podolsky Rosen Paradox," Physics **1**, 195 (1964). [6.2, 21.2*]
23. J. S. Bell: "On the Problem of Hidden Variables in Quantum Mechanics," Rev. Mod. Phys. **38**, 447 (1966). [6.2]
24. M. V. Berry: "Quantal Phase Factors Accompanying Adiabatic Changes," Proc. R. Soc. London, Ser. A **392**, 45 (1984). [7.8]
25. H. A. Bethe, R. W. Jackiw: *Intermediate Quantum Mechanics*, 2nd edn. (W. A. Benjamin, Reading 1973). [17, 17.5*, 18.5, 19.2]
26. H. A. Bethe, E. E. Salpeter *Quantum Mechanics of One- and Two-Electron Atoms* (Springer-Verlag, Berlin 1957). [19.2]
27. L. C. Biedenharn, J. D. Louck: *Angular Momentum in Quantum Physics. Theory and Applications* (Encyclopedia of Mathematics and its Applications, Vol. 8; Addison-Wesley, Reading 1981). [14]
28. L. C. Biedenharn, J. D. Louck *The Racah-Wigner Algebra in Quantum Theory* (Encyclopedia of Mathematics and its Applications, Vol. 9; Addison-Wesley, Reading 1981). [14]
29. J. D. Bjorken, S. D. Drell: *Relativistic Quantum Mechanics* (McGraw-Hill, New York 1964). [15, 15.5]
30. J. D. Bjorken, S. D. Drell: *Relativistic Quantum Fields* (McGraw-Hill, New York 1965). [17.3]
31. J. M. Blatt: "Practical Points Concerning the Solution of the Schrödinger Equation," J. Comp. Phys. **1**, 382 (1967). [8.2]
32. D. I. Blokhintsev: *Principles of Quantum Mechanics* (Allyn and Bacon, Boston 1964). [10.5, 16.7*]
33. A. Böhm: *Quantum Mechanics. Foundations and Applications*, 3rd edn. (Springer-Verlag, New York 1993).
34. D. Bohm: *Quantum Theory* (Prentice-Hall, Englewood Cliffs 1951). [21.1]
35. D. Bohm: "A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden' Variables" I and II, Phys. Rev. **85**, 166, 180 (1952). [6.2]
36. N. Bohr: "Can Quantum-mechanical Description of Physical Reality be Considered Complete," Phys. Rev. **48**, 696 (1935); [21.1]
37. N. Bohr: *The Philosophical Writings of Niels Bohr* (Papers reproduced by Ox Bow Press, Woodbridge 1987). [6.5]

38. J. S. Bolemon: "Computer Solutions to a Realistic 'One-dimensional' Schrödinger Equation," Am. J. Phys. **40**, 1511 (1972). [8.4, 9.5]
39. J. J. Bollinger, D. J. Heinzen, W. M. Itano, S. L. Gilbert, D. J. Wineland: "Test of the Linearity of Quantum Mechanics by rf Spectroscopy of the $^9\text{Be}^+$ Ground State," Phys. Rev. Lett. **63**, 1031 (1989). [21.5]
40. E. Boridy: *Electromagnétisme. Théorie et applications*, 2nd edn. (Presses de l'Université du Québec, Québec 1990). [7.7, 7.8, 9.6]
41. S. Brandt, H. D. Dahmen: *Quantum Mechanics on the Personal Computer*, 3rd edn. (Springer-Verlag, Berlin 1995). [8]
42. S. Brandt, H. D. Dahmen: *Quantum Mechanics on the Macintosh*, 2nd edn. (Springer-Verlag, New York 1995). [8]
43. B. H. Bransden, C. J. Joachain: *Introduction to Quantum Mechanics* (Longman, Harlow 1989).
44. H. I. Brown: *Perception, Theory and Commitment. The New Philosophy of Science* (Precedent Publishing, Chicago 1977). [6.1]
45. L. M. Brown, A. Pais, B. Pippard, eds.: *Twentieth Century Physics*, Vols. I, II and III (Institute of Physics Publishing, Bristol; American Institute of Physics, Woodbury 1995).
46. M. Brune, E. Hagley, J. Dreyer, X. Maître, A. Maali, C. Wunderlich, J. M. Raimond, S. Haroche: "Observing the Progressive Decoherence of the 'Meter' in a Quantum Measurement," Phys. Rev. Lett. **77**, 4887 (1996). [21.6]
47. M. Bunge: *Foundations of Physics* (Springer-Verlag, New York 1967). [6.1]
48. C. P. Burgess: "Quantum Tunneling and Imaginary-time Paths," Am. J. Phys. **59**, 994 (1991). [16.7]
49. P. Busch, M. Grabowski, P. J. Lahti: *Operational Quantum Physics* (Springer-Verlag, New York 1995).
50. J. F. Clauser, M. A. Horne: "Experimental Consequences of Objective Local Theories," Phys. Rev. D **10**, 526 (1974) [21.3]
51. J. F. Clauser, A. Shimony "Bell's Theorem: Experimental Tests and Implications," Rep. Prog. Phys. **41**, 1881 (1978). [21.2, 21.3]
52. E. R. Cohen, B. N. Taylor: "The Fundamental Physical Constants," Phys. Today **49** (August, Part II), BG9 (1996). [UC]
53. C. Cohen-Tannoudji, B. Diu, F. Laloë: *Quantum mechanics*, Vols. I and II, 2nd edn. (Wiley, New York 1977). [4.1, 5.7, 5.8, 5.9.2, 10, 19.6]
54. S. Coleman: "The uses of instantons." In: *The Whys of Subnuclear Physics*, ed. by A. Zichichi (Plenum Press, New York 1979), pp. 805–941. [16.4]
55. E. U. Condon, H. Odabaşi: *Atomic Structure* (Cambridge University Press, London 1980).
56. E. U. Condon, G. H. Shortley: *The Theory of Atomic Spectra* (Cambridge University Press, London 1951). [18.2*, 18.5, 18.6]
57. J. F. Cornwell: *Group Theory in Physics*, Vols. I and II (Academic Press, London 1984). [13, 13.8]
58. J. P. Costella, B. H. J. McKellar: "The Foldy-Wouthuysen Transformation," Am. J. Phys. **63**, 1119 (1995). [15.5]
59. R. D. Cowan: *The Theory of Atomic Structure and Spectra* (University of California Press, Berkeley 1981). [7.9.1, 9.4, 14.6, 17, 17.3, 17.5*, 18.1, 18.2, 18.4*, 18.5*, 18.7*, 19.5, 19.6]
60. M. Creutz, B. Freedman: "A Statistical Approach to Quantum mechanics," Ann. Phys. (NY) **132**, 427 (1981). [16.1]

61. J. T. Cushing, E. McMullin, eds.: *Philosophical Consequences of Quantum Theory* (University of Notre Dame Press, Notre Dame 1989). [21.2]
62. R. F. Dashen, B. Hasslacher, A. Neveu: "Nonperturbative Methods and Extended-hadron Models in Field Theory. I. Semiclassical Functional Methods," Phys. Rev. D **10**, 4114 (1974). [16.6]
63. A. S. Davydov: *Quantum Mechanics*, 2nd edn. (Pergamon Press, Oxford 1976).
64. W. Demtröder: *Laser Spectroscopy. Basic Concepts and Instrumentation*, 2nd edn. (Springer-Verlag, Berlin 1996). [19.7]
65. P. Dennery, A. Krzywicki: *Mathematics for Physicists* (Harper and Row, New York 1967). [5.9.1, 5.9.2, 7.3, 7.4, 7.9.1*, 7.9.2, 11.3]
66. B. d'Espagnat: "The Quantum Theory and Reality," Sci. Am. **241** (May), 158 (1979).
67. B. d'Espagnat: *Veiled Reality. An Analysis of Present-day Quantum Mechanical Concepts* (Addison-Wesley, Reading 1995). [6.2, 12.4, 21.7*]
68. P. A. M. Dirac: *Lectures on Quantum Mechanics* (Belfer Graduate School of Science, Yeshiva University, New York 1964). [6.6]
69. P. A. M. Dirac: *The Principles of Quantum Mechanics*, 4th edn. (Clarendon Press, Oxford 1967). [6.3, 15]
70. F. Dowker, A. Kent: "On the Consistent Histories Approach to Quantum Mechanics," J. Stat. Phys. **82**, 1575 (1996). [21.7]
71. G. W. F. Drake, ed.: *Atomic, Molecular, and Optical Physics Handbook* (American Institute of Physics, Woodbury 1996). [20]
72. G. W. F. Drake: "High Precision Calculations for Helium." In: [71], pp. 154–171. [17.2]
73. A. R. Edmonds: *Angular Momentum in Quantum Mechanics*, 2nd edn. (Princeton University Press, Princeton 1960). [7.9.1*, 14, 14.1, 14.3, 14.6*, 18.2]
74. A. Einstein: "Remarks Concerning the Essays Brought Together in this Cooperative Volume." In: [201], pp. 665–688. [21.1]
75. A. Einstein, B. Podolsky, N. Rosen: "Can Quantum-mechanical Description of Physical Reality be Considered Complete?" Phys. Rev. **47**, 777 (1935). [21.1*, 21.3]
76. R. M. Eisberg: *Fundamentals of Modern Physics* (Wiley, New York 1961). [1]
77. E. Elbaz: *Mécanique quantique* (Marketing, Paris 1985).
78. H. A. Enge: *Introduction to Nuclear Physics* (Addison-Wesley, Reading 1966). [16.7]
79. M. J. Englefield: *Group Theory and the Coulomb Problem* (Wiley, New York 1972). [14.8]
80. H. Everett III: "'Relative State' Formulation of Quantum Mechanics," Rev. Mod. Phys. **29**, 454 (1957). [21.5]
81. G. Fano: *Mathematical Methods of Quantum Mechanics* (McGraw-Hill, New York, 1971). [2, 5.9.2]
82. J. M. Feagin: *Quantum Methods with Mathematica* (Springer-Verlag, New York 1994). [8]
83. R. P. Feynman: "Space-time Approach to Non-relativistic Quantum Mechanics," Rev. Mod. Phys. **20**, 367 (1948). [16.1]
84. R. P. Feynman, A. R. Hibbs: *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York 1965). [16.1]
85. R. P. Feynman, R. B. Leighton, M. Sands: *The Feynman Lectures on Physics*, Vol. III, *Quantum Mechanics* (Addison-Wesley, Reading 1965). [3.8, 4.1]

86. C. Froese Fischer: *The Hartree–Fock Method for Atoms: A Numerical Approach* (Wiley, New York 1977). [17, 17.5*, 18.7]
87. C. Froese Fischer: "Atomic Structure: Multiconfiguration Hartree–Fock Theories." In: [71], pp. 243–257. [17.5, 18.7, 19.5]
88. M. Gell-Mann, J. B. Hartle: "Quantum Mechanics in the Light of Quantum Cosmology." In: *Complexity, Entropy, and the Physics of Information*, ed. by W. H. Zurek (Addison-Wesley, Reading 1990) pp. 425–458. [21.6]
89. C. C. Gerry, J. Kiefer: "Numerical Path Integration Without Monte Carlo," Am. J. Phys. **56**, 1002 (1988). [16.1]
90. R. Gilmore: *Lie Groups, Lie Algebras, and Some of Their Applications* (Wiley, New York 1974). [13]
91. R. Girard, L. Marchildon: "Are Tachyon Causal Paradoxes Solved?" Found. Phys. **14**, 535 (1984). [21.3]
92. A. Goldberg, H. M. Schey, J. L. Schwartz: "Computer-generated Motion Pictures of One-dimensional Quantum-mechanical Transmission and Reflection Phenomena," Am. J. Phys. **35**, 177 (1967). [8.2]
93. H. Goldstein: *Classical Mechanics*, 2nd edn. (Addison-Wesley, Reading 1980). [6.6, 7.7, 16.1, 20.6*]
94. S. Goldstein: "Quantum Theory Without Observers," Phys. Today **51** (March), 42; (April), 38 (1998). [21.5]
95. I. P. Grant: "Relativistic Atomic Structure." In: [71], pp. 258–286. [18.7]
96. N. I. Greenberg: "Discussion of the Rigid Rotator Problem," Am. J. Phys. **45**, 199 (1977). [20.6]
97. D. M. Greenberger, ed.: *New Techniques and Ideas in Quantum Measurement Theory* (Annals of the New York Academy of Sciences, Vol. 480, New York 1986). [21.4]
98. D. M. Greenberger, M. A. Horne, A. Shimony, A. Zeilinger: "Bell's Theorem Without Inequalities," Am. J. Phys. **58**, 1131 (1990). [21.2]
99. W. Greiner: *Quantum Mechanics: An Introduction*, 4th edn. (Springer-Verlag, Berlin, Heidelberg 2001).
100. W. Greiner, B. Müller: *Quantum Mechanics: Symmetries*, 2nd edn. (Springer-Verlag, New York 1994). [13, 14.8]
101. D. C. Griffin, J. B. McGhie: "General Interactive Computer Program for the Solution of the Schrödinger Equation," Am. J. Phys. **41**, 1149 (1973). [8.4, 9.5]
102. R. B. Griffiths: "Consistent Histories and the Interpretation of Quantum Mechanics," J. Stat. Phys. **36**, 219 (1984). [12.6]
103. R. B. Griffiths: "Correlations in Separated Quantum Systems: A Consistent History Analysis of the EPR Problem," Am. J. Phys. **55**, 11 (1987). [21.7]
104. R. B. Griffiths: "Consistent Histories and Quantum Reasoning," Phys. Rev. A **54**, 2759 (1996). [12.6]
105. M. Hamermesh: *Group Theory and its Application to Physical Problems* (Dover, New York 1989). [13, 13.5, 14.1]
106. N. R. Hanson: *The Concept of the Positron* (Cambridge University Press, London 1963). [15]
107. W. Heisenberg: *Physics and Philosophy. The Revolution in Modern Science* (Harper and Brothers, New York 1958). [6.2]
108. F. Herman, S. Skillman: *Atomic Structure Calculations* (Prentice-Hall, Englewood Cliffs 1963). [17.5]

109. G. Herzberg: *Molecular Spectra and Molecular Structure*, Vol. I, *Spectra of Diatomic Molecules*, 2nd edn. (Van Nostrand Reinhold, New York 1950). [10.7, 20.4, 20.5*]
110. G. Herzberg: *Molecular Spectra and Molecular Structure*, Vol. II, *Infrared and Raman Spectra of Polyatomic Molecules* (Van Nostrand Reinhold, New York 1945).
111. G. Herzberg: *Molecular Spectra and Molecular Structure*, Vol. III, *Electronic Spectra and Electronic Structure of Polyatomic Molecules* (Van Nostrand Reinhold, New York 1966).
112. J. R. Hiller, I. D. Johnston, D. F. Styer: *Quantum Mechanics Simulations – Consortium for Upper-Level Physics Software* (Wiley, New York 1995). [8]
113. J. Hladik: *La théorie des groupes en physique et chimie quantiques* (Masson, Paris 1995). [13]
114. P. R. Holland: *The Quantum Theory of Motion. An Account of the de Broglie–Bohm Causal Interpretation of Quantum Mechanics* (Cambridge University Press, London 1993). [6.2*, 21.3]
115. B. R. Holstein: "The Linear Potential Propagator," Am. J. Phys. **65**, 414 (1997). [16e]
116. B. R. Holstein, A. R. Swift: "Path Integrals and the WKB Approximation," Am. J. Phys. **50**, 829 (1982). [16.4]
117. B. R. Holstein, A. R. Swift: "Barrier Penetration via Path Integrals," Am. J. Phys. **50**, 833 (1982). [16.7]
118. M. Horbatsch: *Quantum Mechanics Using Maple* (Springer-Verlag, New York 1995). [8]
119. J. O. Hornkohl, C. Parriger: "Angular Momentum States of the Diatomic Molecule," Am. J. Phys. **64**, 623 (1996). [20.4]
120. S. Howard, S. K. Roy: "Coherent States of a Harmonic Oscillator," Am. J. Phys. **55**, 1109 (1987). [5.8]
121. D. L. Huestis: "Radiative Transition Probabilities." In: [71], pp. 394–410. [20.6]
122. C. Itzykson, J. B. Zuber: *Quantum Field Theory* (McGraw-Hill, New York 1980). [17.3]
123. J. D. Jackson: *Classical Electrodynamics*, 3rd edn. (Wiley, New York 1999). [UC, 7.7*, 9.6]
124. M. Jammer: *The Philosophy of Quantum Mechanics* (Wiley, New York 1974). [P, 6.2, 6.4, 6.5, 12.2, 21.1, 21.4*]
125. M. Jammer: *The Conceptual Development of Quantum Mechanics*, 2nd edn. (American Institute of Physics, Woodbury, 1989). [P, 1, 1.5, 4.1]
126. J. P. Jarrett: "On the Physical Significance of the Locality Conditions in the Bell Arguments," Noûs **18**, 569 (1984). [21.3*]
127. J. M. Jauch: *Foundations of Quantum Mechanics* (Addison-Wesley, Reading 1968). [5.9.2]
128. B. R. Johnson: "New Numerical Methods Applied to Solving the One-dimensional Eigenvalue Problem," J. Chem. Phys. **67**, 4086 (1977). [8.2]
129. B. R. Johnson, J. O. Hirschfelder, K. H. Yang "Interaction of Atoms, Molecules, and Ions with Constant Electric and Magnetic Fields," Rev. Mod. Phys. **55**, 109 (1983). [19.2]
130. E. Joos, H. D. Zeh: "The Emergence of Classical Properties Through Interaction with the Environment," Z. Phys. B **59**, 223 (1985). [21.6]
131. T. Kato: *Perturbation Theory for Linear Operators*, 2nd edn. (Springer-Verlag, Berlin 1980). [10.1]

132. Y. S. Kim, E. P. Wigner: "Canonical Transformation in Quantum Mechanics," Am. J. Phys. **58**, 439 (1990). [5.8, 16.8]
133. C. Kittel: *Introduction to Solid State Physics*, 7th edn. (Wiley, New York 1996). [11.9]
134. D. H. Kobe, K. H. Yang: "Gauge Invariance in Quantum Mechanics: Zero Electromagnetic Field," Am. J. Phys. **51**, 163 (1983). [7.8]
135. T. S. Kuhn: *The Structure of Scientific Revolutions*, 2nd edn. (The University of Chicago Press, Chicago 1970). [1.5]
136. L. D. Landau, E. M. Lifshitz: *Mechanics*, 3rd edn. (Pergamon, Oxford 1976). [14.8, 16.1, 20.6*]
137. L. D. Landau, E. M. Lifshitz: *The Classical Theory of Fields*, 4th edn. (Pergamon, Oxford 1975). [6.4, 7.7*]
138. L. D. Landau, E. M. Lifshitz: *Quantum Mechanics (Non-relativistic Theory)*, 3rd edn. (Butterworth-Heinemann, Oxford 1977). [7.5*, 11e, 13.5*, 16.4, 20, 20.3, 20.4, 20.6*]
139. L. D. Landau, E. M. Lifshitz, V. B. Berestetskii, L. P. Pitaevskii: *Relativistic Quantum Theory*, Part I (Pergamon, Oxford 1971). [15, 15.6]
140. L. D. Landau, E. M. Lifshitz, L. P. Pitaevskii: *Statistical Physics*, Part I, 3rd edn. (Butterworth-Heinemann, Oxford 1980). [12.5*, 13.8]
141. L. D. Landau, E. M. Lifshitz: *Electrodynamics of Continuous Media*, (Pergamon, Oxford 1960). [10.6]
142. R. Latter: "Atomic Energy Levels for the Thomas–Fermi and Thomas–Fermi–Dirac Potential," Phys. Rev. **99**, 510 (1955). [9.4, 9.5]
143. N. N. Lebedev: *Special Functions and their Applications* (Dover, New York 1972). [5.9.3, 7.9.1, 7.9.2, 7.9.3]
144. A. J. Leggett: "Quantum Mechanics and Realism at the Macroscopic Level: Is an Experimental Discrimination Feasible?" In: [97], pp. 21–24. [21.6]
145. A. J. Leggett: "Reflections on the Quantum Measurement Paradox." In: *Quantum Implications: Essays in Honour of David Bohm*, ed. by B. J. Hiley, F. David Peat (Routledge and Kegan Paul, London 1987) pp. 85–104. [21.4]
146. J. Leite Lopes, B. Escoubès: *Sources et évolution de la physique quantique. Textes fondateurs* (Masson, Paris 1995) [P]
147. W. Lichten, D. Shiner, Z. X. Zhou: "Measurement of the Lamb Shifts in Singlet Levels of Atomic Helium," Phys. Rev. A **43**, 1663 (1991). [17.2]
148. F. London, E. Bauer: "The Theory of Observation in Quantum Mechanics." In: [235], pp. 217–259. [21.4, 21.5]
149. J. D. Louck: "Angular Momentum Theory." In: [71], pp. 6–55. [14, 14.6]
150. P. K. MacKeown: "Evaluation of Feynman Path Integrals by Monte Carlo Methods," Am. J. Phys. **53**, 880 (1985). [16.1]
151. H. Margenau: *The Nature of Physical Reality* (Ox Bow Press, Woodbridge 1950). [6.1]
152. H. Margenau: "Philosophical Problems Concerning the Meaning of Measurement in Physics," Philos. Sci. **25**, 23 (1958). [6.3*]
153. H. Margenau: "Measurements in Quantum Mechanics," Ann. Phys. (NY) **23**, 469 (1963). [6.3*]
154. W. C. Martin, W. L. Wiese: "Atomic Spectroscopy." In: [71], pp. 135–153. [9.7, 19.5]
155. J. P. Mathieu, ed.: *Histoire de la physique*, t. 2, *La physique au XXe siècle* (Lavoisier, Paris 1991). [1]

156. R. McDougall: *Numérisation des équations de Hartree–Fock.* MSc Thesis, Université du Québec à Trois-Rivières (2001). [9.7]
157. J. R. Merrill: "Introducing Scattering Theory with a Computer," Am. J. Phys. **41**, 1156 (1973). [8.5]
158. J. R. Merrill: *Using Computers in Physics* (University Press of America, Lanham 1980). [8e*]
159. E. Merzbacher: *Quantum Mechanics*, 3rd edn. (Wiley, New York 1998). [16.4]
160. A. Messiah: *Quantum Mechanics*, Vols. I and II (North-Holland, Amsterdam 1961). [5.9.1, 6.6, 7.5, 13.6]
161. S. A. Meyer, M. A. Morgenstern, S. K. Knudson, D. W. Noid: "Novel Method for WKB Analysis of Multidimensional Systems," Am. J. Phys. **59**, 145 (1991). [16.6]
162. P. J. Mohr: "Tests of Fundamental Physics." In: [71], pp. 341–351. [15.5, 15.6]
163. G. S. Monk: *Light: Principles and Experiments*, 2nd edn. (Dover, New York 1963). [6.4]
164. C. Monroe, D. M. Meekhof, B. E. King, D. J. Wineland: "A 'Schrödinger Cat' Superposition State of an Atom," Science **272**, 1131 (1996). [21.6]
165. C. E. Moore: *Atomic Energy Levels, as Derived from the Analyses of Optical Spectra*, Circular No. 467 of the U.S. National Bureau of Standards, Vols. I–III (U.S. Government Printing Office, Washington 1949–58; reissued as NSRDS-NBS 35, Vols. I–III, 1971). [18.4]
166. C. E. Moore: *Selected Tables of Atomic Spectra*, NSRDS-NBS 3, Sects. 1–8 (U.S. Government Printing Office, Washington 1965–79). [18.4]
167. A. W. Naylor, G. R. Sell: *Linear Operator Theory in Engineering and Science* (Springer-Verlag, New York 1982). [5.9.2]
168. I. Newton: *Opticks*, 4th edn. (Dover, New York 1952). [1.1]
169. R. G. Newton: *Scattering Theory of Waves and Particles* (McGraw-Hill, New York 1966). [11, 11.4]
170. C. Ngô, H. Ngô: *Physique quantique. Introduction* (Masson, Paris 1991). [19.3, 19.6]
171. D. R. Nilson: "Bibliography on the History and Philosophy of Quantum Physics." In: [222], pp. 457–520.
172. R. Omnès: "Consistent Interpretations of Quantum Mechanics," Rev. Mod. Phys. **64**, 339 (1992). [16.8]
173. R. Omnès: *The Interpretation of Quantum Mechanics* (Princeton University Press, Princeton 1994). [12.2, 12.6*, 16.8, 21.6, 21.7*]
174. A. Pais: *'Subtle is the Lord... ' The Science and the Life of Albert Einstein* (Oxford University Press, Oxford 1982). [1]
175. A. Pais: *Inward Bound. Of Matter and Forces in the Physical World* (Clarendon Press, Oxford 1986). [1]
176. D. Park: *Introduction to the Quantum Theory*, 2nd edn. (McGraw-Hill, New York 1974). [16.7]
177. J. L. Park: "Nature of Quantum States," Am. J. Phys. **36**, 211 (1968) [6.2]
178. L. Pauling, E. B. Wilson, Jr.: *Introduction to Quantum Mechanics* (McGraw-Hill, New York 1935). [10.6]
179. M. G. Payne, M. Edwards: "Approximate Atomic Structure Calculations Using a Microcomputer," Comp. in Phys. **7**, 465 (1993). [8.4, 9.5]
180. L. L. Pennisi: *Elements of Complex Variables* (Holt, Rinehart and Winston, New York 1963). [11.3]

181. A. Peres: "What is a State Vector?" Am. J. Phys. **52**, 644 (1984) [6.2, 6.3]
182. A. Peres: "When is a Quantum Measurement?" Am. J. Phys. **54**, 688 (1986) [6.2]
183. M. I. Petrashen, E. D. Trifonov: *Applications of Group Theory in Quantum Mechanics* (MIT Press, Cambridge 1969). [13, 13.5*, 13.8*, 14.8, 15.1]
184. K. R. Popper: "Quantum Mechanics Without 'the Observer'." In: *Quantum Theory and Reality*, ed. by M. Bunge (Springer-Verlag, New York 1967), pp. 7–44. [6.2, 6.4]
185. K. R. Popper: *Quantum Theory and the Schism in Physics: From the Postscript to the Logic of Scientific Discovery* (Rowman and Littlefield, Totowa 1982).
186. J. Preskill: "Battling Decoherence: The Fault-tolerant Quantum Computer," Phys. Today **52** (June), 24 (1999). [21.6]
187. W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling: *Numerical Recipes. The Art of Scientific Computing* (Cambridge University Press, London 1986). [8.2*]
188. G. Racah: "Theory of Complex Spectra. I," Phys. Rev. **61**, 186 (1942). [18.1]
189. G. Racah: "Theory of Complex Spectra. II," Phys. Rev. **62**, 438 (1942). [18.1, 18.3]
190. G. Racah: "Theory of complex spectra. III," Phys. Rev. **63**, 367 (1943). [18.1]
191. R. Rajaraman: "Some Non-perturbative Semi-classical Methods in Quantum Field Theory (a Pedagogical Review)," Phys. Rep. **21C**, 227 (1975). [16.6]
192. L. R. Ram-Mohan, S. Saigal, D. Dossa, J. Shertzer: "The Finite-element Method for Energy Eigenvalues of Quantum Mechanical Systems," Comp. in Phys. **4**, 50 (1990). [8.4]
193. H. Rechenberg: "Quanta and Quantum Mechanics." In: [45], Vol. I, pp. 143–248. [1]
194. C. Reischer, A. Paradis, R. Leblanc: *Eléments d'algèbre linéaire* (Presses de l'Université du Québec, Québec 1992). [2]
195. G. Roepstorff: *Path Integral Approach to Quantum Physics* (Springer-Verlag, New York 1994). [16.1]
196. C. C. J. Roothaan: "New Developments in Molecular Orbital Theory," Rev. Mod. Phys. **23**, 69 (1951). [20.2]
197. J. Rosmorduc, ed.: *Histoire de la physique*, t. 1, *La formation de la physique classique* (Lavoisier, Paris 1987). [1]
198. J. J. Sakurai: *Modern Quantum Mechanics*, revised by S. F. Tuan (Addison-Wesley, Reading 1994). [13.6]
199. M. Sargent III, M. O. Scully, W. E. Lamb, Jr.: *Laser Physics* (Addison-Wesley, Reading 1974). [19.3]
200. L. I. Schiff: *Quantum Mechanics*, 3rd edn. (McGraw-Hill, New York 1968). [P, 7.5*, 7.9.1, 10, 10.6, 14.8, 16.4, 19.4]
201. P. A. Schilpp, ed.: *Albert Einstein: Philosopher-Scientist* (Open Court Publishing Company, La Salle 1970).
202. E. W. Schmid, G. Spitz, W. Lösch: *Theoretical Physics on the Personal Computer*, 2nd edn. (Springer-Verlag, Berlin 1990). [8.5, 8e, 17.2]
203. F. Schwabl: *Advanced Quantum Mechanics* (Springer-Verlag, Berlin 1999).
204. F. Schwabl: *Quantum Mechanics*, 3rd edn. (Springer-Verlag, Berlin 2002).
205. M. Sebhatu, W. E. Gettys: "A Least-squares Method for Extracting Phase Shifts," Comp. in Phys. **3** (September/October), 65 (1989). [8.5]

206. E. Segrè: *From Falling Bodies to Radio Waves: Classical Physicists and Their Discoveries* (W. H. Freeman, New York 1984). [1]
207. E. Segrè: *From X-rays to Quarks: Modern Physicists and Their Discoveries* (W. H. Freeman, San Francisco, 1980). [1]
208. R. Shankar: *Principles of Quantum Mechanics*, 2nd edn. (Plenum Press, New York 1994). [6.6]
209. A. Shapere, F. Wilczek, eds.: *Geometric Phases in Physics* (World Scientific, Singapore 1989).
210. J. C. Slater: "Physical Meaning of Wave Mechanics," J. Franklin Inst. **207**, 449 (1929). [6.2]
211. J. C. Slater: *Quantum Theory of Atomic Structure*, Vols. I and II (McGraw-Hill, New York 1960). [17, 18.2*]
212. J. C. Slater: *Quantum Theory of Molecules and Solids*, Vol. I, *Electronic Structure of Molecules* (McGraw-Hill, New York 1963). [20, 20.2, 20.3]
213. J. C. Slater: *Quantum Theory of Molecules and Solids*, Vol. II, *Symmetry and Energy Bands in Cristals* (McGraw-Hill, New York 1965).
214. J. C. Slater: *Quantum Theory of Matter*, 2nd edn. (R. F. Krieger Publishing Company, Huntington 1977). [9.4, 17, 17.4, 17.5, 20, 20.3*]
215. I. H. Sloan: "Errors in the Numerov and Runge–Kutta Methods," J. Comput. Phys. **2**, 414 (1968). [8.2]
216. J. J. Slosser, P. Meystre: "Resource Letter CQO-1: Coherence in Quantum Optics," Am. J. Phys. **65**, 275 (1997). [5.8]
217. H. P. Stapp: "The Copenhagen Interpretation," Am. J. Phys. **40**, 1098 (1972). [6.5]
218. S. Sternberg: *Group Theory and Physics* (Cambridge University Press, London 1994). [13.1, 14.1]
219. C. H. Storry, N. E. Rothery, E. A. Hessels: "Precision Separated-oscillatory-field Measurement of the $n = 10$ $^{+}F_3$-$^{+}G_4$ Interval in Helium: A Precision Test of Long-range, Radiative, and Retardation Effects," Phys. Rev. Lett. **75**, 3249 (1995). [17.2]
220. R. F. Streater, A. S. Wightman: *PCT, Spin and Statistics, and All That* (Benjamin/Cummings, Reading 1978). [17.3]
221. F. Suppe, ed.: *The Structure of Scientific Theories*, 2nd edn. (University of Illinois Press, Urbana 1977). [6.1]
222. P. Suppes, ed.: *Logic and Probability in Quantum Mechanics* (Synthese Library, Vol. 78; Reidel, Dordrecht 1976).
223. J. D. Talman: *Special Functions. A Group Theoretical Approach* (W. A. Benjamin, New York 1968). [7.9.1, 7.9.3, 14.3*]
224. J. R. Taylor: *Scattering Theory* (Wiley, New York 1972). [11, 11.2, 11.4, 11.8]
225. C. D. Tesche: "Schrödinger's Cat: A Realization in Superconducting Devices." In: [97], pp. 36–50. [21.6]
226. M. Tinkham: *Group Theory and Quantum Mechanics* (McGraw-Hill, New York 1964). [13, 13.8]
227. B. L. van der Waerden, ed.: *Sources of Quantum Mechanics* (Dover, New York 1968). [P]
228. B. C. van Fraassen: *Quantum Mechanics: An Empiricist View* (Clarendon Press, Oxford 1991). [21.3]
229. P. B. Visscher: "A Fast Explicit Algorithm for the Time-dependent Schrödinger Equation," Comp. in Phys. **5**, 596 (1991). [8.2]

230. J. von Neumann: *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton 1955). [6.3, 21.5]
231. S. Weinberg: "Precision Tests of Quantum Mechanics," Phys. Rev. Lett. **62**, 485 (1989) [21.5]
232. S. Weinberg: *The Quantum Theory of Fields*, Vol. I (Cambridge University Press, London 1995). [14.1]
233. M. Weissbluth: *Atoms and Molecules* (Academic Press, New York 1978). [13.5, 17, 18.2*, 18.4, 18.5*, 20, 20.2*, 20.3, 20.5]
234. H. Weyl: *The Theory of Groups and Quantum Mechanics*, 2nd edn. (Dover, New York 1950). [13]
235. J. A. Wheeler, W. H. Zurek, eds.: *Quantum Theory and Measurement* (Princeton University Press, Princeton 1983). [21.4]
236. E. T. Whittaker, G. N. Watson: *A Course of Modern Analysis*, 4th edn. (Cambridge University Press, London 1927). [7.9.1, 7.9.3, 11.3]
237. G. C. Wick, A. S. Wightman, E. P. Wigner: "The Intrinsic Parity of Elementary Particles," Phys. Rev. **88**, 101 (1952) [3.2]
238. E. P. Wigner: *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra* (Academic Press, New York 1959). [13, 13.1, 14.1]
239. E. P. Wigner: "Remarks on the Mind Body Question." In: *The Scientist Speculates*, ed. by I. J. Good (Heinemann, London 1961) pp. 284–302. [21.5]
240. E. P. Wigner, "The Problem of Measurement," Am. J. Phys. **31**, 6 (1963). [21.5]
241. B. G. Wybourne: *Classical Groups for Physicists* (Wiley, New York 1974). [14.8]
242. P. Yam: "Bringing Schrödinger's Cat to Life," Sci. Am. **276** (June), 124 (1997). [21.6]
243. D. R. Yarkony: "Molecular Structure." In: [71], pp. 357–377. [20.1]
244. W. Yourgrau, A. van der Merwe, eds.: *Perspectives in Quantum Theory* (Dover, New York 1979).
245. R. N. Zare: *Angular Momentum. Understanding Spatial Aspects in Chemistry and Physics* (Wiley, New York 1988). [14]
246. W. H. Zurek: "Pointer Basis of Quantum Apparatus: Into What Mixture Does the Wave Packet Collapse?" Phys. Rev. D **24**, 1516 (1981) [21.6]
247. W. H. Zurek: "Environment-induced Superselection Rules," Phys. Rev. D **26**, 1862 (1982). [21.6]
248. W. H. Zurek: "Decoherence and the Transition from Quantum to Classical," Phys. Today **44** (October), 36 (1991). [21.6]

Index

The symbol "n" following a page number refers to a footnote, while the symbol "e" refers to an exercise.

21 Long-Distance Correlations and Measurement

Ever since its formulation in 1925–26, quantum mechanics has explained a very large number of phenomena. Examples have been given throughout this book. Properties of atoms, molecules, nuclei, solids, superconductors and superfluids, among others, cannot be understood without the systematic use of quantum mechanics. No major discrepancies are known between its predictions and experiment. Quantum mechanics is no doubt one of the best corroborated physical theories ever.

In spite of its triumphs, quantum mechanics is subject to persistent criticism. It is taken to task for its statistical character: it cannot, in general, predict the result of individual measurements. One is led to wonder whether quantum mechanics provides a complete description of phenomena within its scope. We shall see how A. Einstein, B. Podolsky and N. Rosen have formulated this question, and what direction J. S. Bell subsequently gave it. Quantum mechanics is also alleged to be incapable of properly describing the measurement process. We will analyze this problem in detail and investigate some of the solutions that have been proposed.

21.1 Einstein, Podolsky and Rosen

Quantum mechanics correctly predicts the half-life of radioactive nuclei. It does not, however, predict the time when a specific nucleus disintegrates. Thus one could think that the quantum-mechanical formalism does not provide a complete description of physical reality. Einstein, Podolsky and Rosen [75] formulated the problem in a particularly penetrating way.[1]

Einstein, Podolsky and Rosen first state that a theory is *complete* only if "every element of the physical reality [has] a counterpart in the physical theory." But what is an element of *physical reality*? Einstein and his coworkers do not answer this question exhaustively, but they propose a sufficient criterion for the existence of an element of physical reality: "If, without in any way disturbing a system, we can predict with certainly [...] the value of a phy-

---

[1] Jammer ([124], Chap. 6) examines the argument of Einstein, Podolsky and Rosen and the elaborate discussion that followed.

L. Marchildon, *Quantum Mechanics*

sical quantity, then there exists an element of physical reality corresponding to this physical quantity."

Einstein, Podolsky and Rosen then consider two physical system [labeled (1) and (2)] which interact for some time and then separate. Let $|\chi(1,2)\rangle$ be the total system's state vector at a given time. Let $A(1)$ be a dynamical variable pertaining to system (1) and $\{|a_i(1)\rangle\}$ a basis of orthonormal eigenvectors of $A(1)$. One can write

$$|\chi(1,2)\rangle = \sum_i |a_i(1)\rangle \otimes |\psi_i(2)\rangle , \tag{21.1}$$

where the $|\psi_i(2)\rangle$ are vectors in the state space of system (2), not necessarily orthogonal or normalized. The index i can be discrete or continuous and its range finite or infinite, depending on $A(1)$.

Now let $B(1)$ be another dynamical variable pertaining to system (1) and $\{|b_s(1)\rangle\}$ a basis of orthonormal eigenvectors of $B(1)$. One can also write

$$|\chi(1,2)\rangle = \sum_s |b_s(1)\rangle \otimes |\phi_s(2)\rangle , \tag{21.2}$$

where, again, the $|\phi_s(2)\rangle$ are vectors in the state space of system (2).

We suppose that $A(1)$ and $B(1)$ each have a nondegenerate spectrum. Suppose also that a measurement is made on $A(1)$ and that value a_j is obtained. No matter what the subsequent evolution of system (1) (be it absorbed, destroyed or left in state $|a_j(1)\rangle$), theorem 2 on p. 264 entails that the second system's state vector is then proportional to $|\psi_j(2)\rangle$.

Had we decided, on the other hand, to measure dynamical variable $B(1)$, the result obtained would have been one of the values b_s. Thus the second system's state vector would have been proportional to one of the $|\phi_s(2)\rangle$. But in general the $|\psi_i(2)\rangle$ are different from the $|\phi_s(2)\rangle$. Hence depending on the dynamical variable pertaining to system (1) that one decides to measure, system (2) subsequently finds itself in states described by two different state vectors. And this, so it seems, without any physical change occurring in the second system due to the measurement made on the first one.

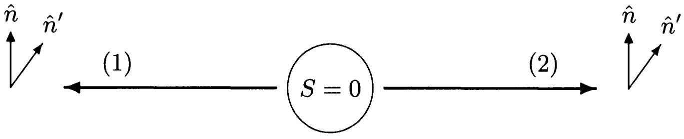

Fig. 21.1. Two spin 1/2 particles prepared in a singlet state and then separating

An example ([34], Chap. 22) can help make these general considerations more concrete (Fig. 21.1). For systems (1) and (2) take two spin 1/2 particles. These particles are first prepared in the singlet state and they eventually come apart. Neglecting spatial wave functions for simplicity, one can write the total system's state vector as

$$|\chi(1,2)\rangle = \frac{1}{\sqrt{2}} \{|+;\hat{n}\rangle|-;\hat{n}\rangle - |-;\hat{n}\rangle|+;\hat{n}\rangle\} . \tag{21.3}$$

In the right-hand side the first factor in the tensor product of two kets pertains to system (1), and the second factor to system (2). Kets $|\pm;\hat{n}\rangle$ represent states wherein the particle has a spin projection on unit vector $\hat{n}$ equal to $\pm\hbar/2$. Note that the singlet state representation does not depend on the choice of vector $\hat{n}$. This means that for any unit vector $\hat{n}'$,

$$|\chi(1,2) = \frac{1}{\sqrt{2}} \{|+;\hat{n}'\rangle|-;\hat{n}'\rangle - |-;\hat{n}'\rangle|+;\hat{n}'\rangle\} . \tag{21.4}$$

Denote by $S_n(1)$ and $S_{n'}(1)$ the projections of the first particle's spin on the $\hat{n}$ and $\hat{n}'$ axes.[2] Suppose one measures $S_n(1)$. If $+\hbar/2$ is obtained, then the second particle's state vector is given by $|-;\hat{n}\rangle$. Therefore one can predict with certainty that a measurement of $S_n(2)$, the spin projection of the second particle on $\hat{n}$, will yield result $-\hbar/2$. Likewise if $-\hbar/2$ is obtained upon measuring $S_n(1)$, then surely a measurement of $S_n(2)$ will yield result $+\hbar/2$.

At the time when $S_n(1)$ is measured, the two particles can be arbitrarily far from each other. It looks as if a measurement of $S_n(1)$ does not in any way disturb the second particle. Without disturbing the second particle, it is thus possible to predict with certainty the value of its spin projection on the $\hat{n}$ axis. Hence according to the criterion of Einstein, Podolsky and Rosen, there exists an element of physical reality corresponding to the projection on the $\hat{n}$ axis of the second particle's spin. This conclusion is true even if the first particle's spin is not actually measured.

By a completely similar argument, one concludes that there exists an element of physical reality corresponding to the spin projection of the second particle on the $\hat{n}'$ axis. In the quantum-mechanical formalism, however, nothing corresponds to simultaneous values of a particle's spin projections on two different axes. In the sense of Einstein, Podolsky and Rosen, quantum mechanics is therefore an incomplete theory.

In the first paragraphs of their paper, Einstein and coworkers insisted on the fact that "[the] elements of the physical reality cannot be determined by a priori philosophical considerations, but must be found by an appeal to results of experiments and measurements." In his answer Bohr [36] in fact pointed out the necessity to take into account the whole measurement setup. According to Bohr the reality criterion of Einstein, Podolsky and Rosen is ambiguous when using the phrase "without in any way disturbing a system."

---

[2] One has $S_n(1) = (\hbar/2)\boldsymbol{\sigma}_1 \cdot \hat{n}$ and $S_{n'}(1) = (\hbar/2)\boldsymbol{\sigma}_1 \cdot \hat{n}'$. Operators $S_n(1)$ and $S_{n'}(1)$ correspond to dynamical variables $A(1)$ and $B(1)$ introduced above.

Properties and dynamical variables have meaning and are defined only in relation with the apparatus used to measure them. If one introduces a particle's spin components on the $\hat{n}$ or $\hat{n}'$ axes, the word "system" must in part refer to the experimental setup used to measure one or the other.

A measurement of $S_n(1)$ cannot, on Bohr's interpretation, mean anything else than the establishment of a correlation between the behavior of the first particle and an instrument rigidly attached to objects defining the reference frame where angular momentum is measured. By a measurement of $S_n(1)$ these objects acquire an undetermined angular momentum, making their use for the determination of $S_{n'}(2)$ impossible. As soon as $S_n(1)$ is measured, experimental conditions for the determination of $S_{n'}(2)$, through correlations between the first and second particles, no longer exist. Hence one cannot say that $S_{n'}(2)$ then has physical reality.

Einstein admitted that Bohr's position is consistent. However, he never abandoned the search for a more complete approach. Later he would summarize the situation by saying ([74], p. 682) that the following two statements are incompatible:

i) "The description by means of the ψ-function is *complete*."
ii) "The real states of spatially separated objects are independent of each other."

Clearly the first statement represents Bohr's position, and the second one Einstein's.

21.2 The Bell Inequality

Suppose that following Einstein, statement (i) in Sect. 21.1 is rejected. Then naturally one is led to the hypothesis of hidden variables. They afford a more complete description than quantum mechanics, which allows to predict with certainty the second particle's spin projection on any axis.

The full consequences of this hypothesis were brought to light by Bell [22]. Essentially Bell showed that the conjunction of statement (ii) and the negation of statement (i) entails something incompatible with quantum mechanics. If quantum mechanics is true, the statement "(ii) and not-(i)" is therefore false. But we already know that the statement "(ii) and (i)" is false. Hence statement (ii) is false. Thus Einstein's solution to the dilemma is insufficient. If quantum mechanics is true, one cannot only reject (i). One must, no matter what, reject (ii). This means that the real states of spatially separated objects are not independent of each other.

Let us examine more closely what Bell showed. Bell considers the situation described in Fig. 21.1, where two spin 1/2 particles interact and eventually come apart. The total system is in the singlet state $|\chi(1,2)\rangle$. Consider the dynamical variable $\boldsymbol{\sigma}_1 \cdot \hat{n}$, proportional to the projection of the first particle's spin on the $\hat{n}$ axis. Its value can be predicted with certainty without the first particle being disturbed. Bell therefore assumes it exists for every $\hat{n}$ and is given by $\xi_1(\hat{n};\lambda)$. Here λ is a hidden variable. Of course $\xi_1(\hat{n};\lambda) = \pm 1$. Similarly, Bell assumes that the value of dynamical variable $\boldsymbol{\sigma}_2 \cdot \hat{n}'$, proportional to the projection of the second particle's spin on the $\hat{n}'$ axis, exists and is given by $\xi_2(\hat{n}';\lambda)$. For the rest of the argument, it does not matter whether λ is a discrete or a continuous variable, or whether it stands for one or several variables. It is crucial, however, that ξ_1 does not depend on $\hat{n}'$ and that ξ_2 does not depend on $\hat{n}$. Empirically, this means that the value of the first particle's spin projection does not depend on the axis along which one chooses to measure the second particle's spin. This is a locality condition, corresponding to statement (ii) of Sect. 21.1.

Let $p(\lambda)$ be the probability density of the hidden variable λ. Bell shows there exists no $p(\lambda)$ that exactly (or even approximately) reproduces the statistical predictions of quantum mechanics.

The proof is very simple. Denote by $P(\hat{n},\hat{n}')$ the mean value of the product of the first particle's spin projection on $\hat{n}$ and the second particle's spin projection on $\hat{n}'$. According to quantum mechanics

$$P_{\mathrm{qm}}(\hat{n},\hat{n}') = \langle\chi(1,2)|\boldsymbol{\sigma}_1\cdot\hat{n}\otimes\boldsymbol{\sigma}_2\cdot\hat{n}'|\chi(1,2)\rangle = -\hat{n}\cdot\hat{n}'. \tag{21.5}$$

The mean value of the product of spin projections (or, since mean values of projections vanish, the *correlation* of spin projections) can also be written in terms of the hidden variable as

$$P_{\mathrm{hv}}(\hat{n},\hat{n}') = \int \mathrm{d}\lambda\, p(\lambda)\xi_1(\hat{n};\lambda)\xi_2(\hat{n}';\lambda), \tag{21.6}$$

where the integral covers all values of λ.

In the specific case where $\hat{n}' = \hat{n}$, (21.5) implies that

$$P_{\mathrm{qm}}(\hat{n},\hat{n}) = -1. \tag{21.7}$$

Experimentally, this corresponds to the observation that spin projections on a given axis for two particles in the same pair always have opposite values. The hidden variable must reproduce that observation. We know that $p(\lambda)$ is a probability density. This means that

$$p(\lambda) \geq 0, \qquad \int \mathrm{d}\lambda\, p(\lambda) = 1. \tag{21.8}$$

Values of ξ_1 and ξ_2 are always equal to $+1$ or -1. Hence if $\hat{n} = \hat{n}'$, the right-hand side of (21.6) can be equal to -1 only if

$$\xi_1(\hat{n};\lambda) = -\xi_2(\hat{n};\lambda) \tag{21.9}$$

for every λ, except perhaps on a set of zero probability. At any rate one can write

$$P_{\mathrm{hv}}(\hat{n},\hat{n}') = -\int \mathrm{d}\lambda\, p(\lambda)\xi_1(\hat{n};\lambda)\xi_1(\hat{n}';\lambda). \tag{21.10}$$

Now let $\hat{n}''$ be another unit vector. Since $[\xi_1(\hat{n}';\lambda)]^2 = 1$,

$$\begin{aligned}
&P_{\mathrm{hv}}(\hat{n},\hat{n}') - P_{\mathrm{hv}}(\hat{n},\hat{n}'') \\
&\quad = -\int \mathrm{d}\lambda\, p(\lambda)\left\{\xi_1(\hat{n};\lambda)\xi_1(\hat{n}';\lambda) - \xi_1(\hat{n};\lambda)\xi_1(\hat{n}'';\lambda)\right\} \\
&\quad = \int \mathrm{d}\lambda\, p(\lambda)\xi_1(\hat{n};\lambda)\xi_1(\hat{n}';\lambda)\left\{-1 + \xi_1(\hat{n}';\lambda)\xi_1(\hat{n}'';\lambda)\right\}.
\end{aligned}$$

Using the fact that the absolute value of an integral is never greater than the integral of the absolute value, one easily finds that

$$\begin{aligned}
|P_{\mathrm{hv}}(\hat{n},\hat{n}') - P_{\mathrm{hv}}(\hat{n},\hat{n}'')| &\leq \int \mathrm{d}\lambda\, p(\lambda)\left\{1 - \xi_1(\hat{n}';\lambda)\xi_1(\hat{n}'';\lambda)\right\} \\
&\leq 1 + P_{\mathrm{hv}}(\hat{n}',\hat{n}'').
\end{aligned} \tag{21.11}$$

This relation, which must be satisfied by spin projection correlations in every local hidden-variable theory, is called the *Bell inequality.* Note that unit vectors $\hat{n}$, $\hat{n}'$ and $\hat{n}''$ are completely arbitrary.

Let us now show that no local hidden-variable theory can reproduce quantum-mechanical correlations. If there was such a theory, we would have $P_{\mathrm{hv}}(\hat{n},\hat{n}') = P_{\mathrm{qm}}(\hat{n},\hat{n}')$. Using (21.5) and Bell's inequality, we could write that

$$|\hat{n}\cdot\hat{n}' - \hat{n}\cdot\hat{n}''| \leq 1 - \hat{n}'\cdot\hat{n}''. \tag{21.12}$$

It is easy to see that this relation is not always satisfied, in particular for

$$\hat{n} = \hat{x}, \qquad \hat{n}' = \frac{1}{\sqrt{2}}(\hat{x}+\hat{y}), \qquad \hat{n}'' = \hat{y}. \tag{21.13}$$

Hence no local hidden-variable theory can reproduce exactly the quantum-mechanical statistical predictions. This is *Bell's theorem.* Bell also showed that local hidden-variable theories cannot reproduce quantum-mechanical predictions arbitrarily well.[3]

One sees at once that Bell's proof no longer holds if the locality hypothesis is not satisfied, i.e. if $\xi_1 = \xi_1(\hat{n},\hat{n}';\lambda)$ and $\xi_2 = \xi_2(\hat{n}',\hat{n};\lambda)$. Indeed (21.7) no longer allows to write ξ_2 in terms of ξ_1.

Let us reformulate slightly differently the result just obtained. When two spin 1/2 particles are in state $|\chi(1,2)\rangle$ and are arbitrarily far from one another, one can predict with certainty the spin projection of one of them on an arbitrary axis, without disturbing anything around. In the present context,

---

[3] An analysis of Bell's theorem and its consequences can be found in [61]. See also [13], Chap. 20 and [97].

a theory is called *local* and *deterministic* if at any time it attributes a well-defined value to the spin projection on any axis $\hat{n}$, independently of what happens far from the particle.[4] Bell's conclusion is that quantum-mechanical predictions are incompatible with every local deterministic theory.

But who in the end is right: quantum mechanics or local deterministic theories? To answer his question, one should measure the correlations $P(\hat{n}, \hat{n}')$ and see whether they obey quantum-mechanical laws or Bell's inequality. The $P(\hat{n}, \hat{n}')$ are experimentally determined by a large number of measurements made on similarly prepared systems. Actually, experimentalists use photons rather than spin 1/2 particles. From local deterministic theories, an inequality incompatible with quantum mechanics can also be proved in this case [51]. Experiments made ([6], [7]) leave little doubt. Results contradict local deterministic theories and agree with quantum-mechanical predictions.

21.3 Quantum Mechanics and Relativistic Locality

Bell's theorem and the experimental measurement of spin correlations strongly suggest that, in contrast with Einstein's intuition, the real states of spatially separated objects are not independent. But to what extent are they dependent of each other? We saw, for instance, that the Bell inequality no longer holds if the value of the spin projection on some axis $\hat{n}$ depends, in addition to the hidden variable, on the axis $\hat{n}'$ along which the other spin is measured. Metaphorically, one would say that the first particle "knows" in which direction the second particle's spin is measured. In some experiments though [7], the second spin's measurement axis was picked so late that even a light signal could not have transmitted this information so that it reached the first particle before its spin was measured. Does this mean that information is transmitted faster than the speed of light?

The question is of cardinal importance. In any theory that obeys Einstein's principle of relativity, faster-than-light information transfer entails insoluble paradoxes [91]. But Einstein's principle of relativity is experimentally corroborated to a very high degree. Thus we must closely examine whether quantum mechanics implies faster-than-light information transfer.

At this stage it is useful to enlarge the discussion. In the proof of Bell's inequality, we used a hypothesis about locality and one about determinism. But the latter is not really necessary. Inequalities like Bell's can be proved even if the specification of the system's state does not categorically determine measurement results, but only determines probabilities [50], [51].

To understand this better, let us introduce additional notation. Let ξ_1 and ξ_2 be the values of two spin 1/2 particles' spin projections on the $\hat{n}_1$ and $\hat{n}_2$ axes. Denote by μ_1 and μ_2 any local information, other than direction of

---

[4] To avoid confusion we should point out that the word "deterministic" does not refer here to the time evolution.

$\hat{n}_1$ and $\hat{n}_2$ axes, necessary to specify the spin-measuring devices. Furthermore let Λ denote the state of the two-particle total system. Here Λ can be a state vector, a hidden variable or something else. We let $P(\xi_1, \xi_2 | \hat{n}_1, \hat{n}_2, \mu_1, \mu_2, \Lambda)$ stand for the probability that a measurement of both spins yields ξ_1 and ξ_2, conditional on the variables shown. We also let $W_1(\xi_1 | \hat{n}_1, \mu_1, \mu_2^0, \Lambda)$ stand for the probability that measurement of the first particle's spin yields ξ_1, conditional on the variables shown. Here μ_2^0 means that no measurement is made on the second spin. The symbol $W_2(\xi_2 | \hat{n}_2, \mu_2, \mu_1^0, \Lambda)$ is introduced with similar meaning. In a deterministic theory values of probabilities P, W_1 and W_2 are restricted to 0 and 1, since Λ determines the result of every measurement. In general P, W_1 and W_2 can take any value in the interval $[0, 1]$.

We say that a theory satisfies the hypothesis of *strong locality* if for any value of the arguments, probabilities obey the equation

$$P(\xi_1, \xi_2 | \hat{n}_1, \hat{n}_2, \mu_1, \mu_2, \Lambda) = W_1(\xi_1 | \hat{n}_1, \mu_1, \mu_2^0, \Lambda) W_2(\xi_2 | \hat{n}_2, \mu_2, \mu_1^0, \Lambda). \tag{21.14}$$

Strong locality implies that the probability that a measurement of the first spin's projection yields value ξ_1, for instance, does not depend on the second spin's measurement result, or even on the axis along which the second spin is measured. Strong locality, or similar essentially equivalent hypotheses, are sufficient to prove inequalities which, like Bell's, are incompatible with quantum mechanics. Anyway, in the quantum-mechanical singlet state $W_1 = 1/2 = W_2$, whereas P depends on the relative orientation of $\hat{n}_1$ and $\hat{n}_2$. Hence obviously strong locality and quantum mechanics are incompatible.

It is interesting to sum both sides of (21.14) on ξ_1 or on ξ_2. Since probabilities W_1 and W_2 are normalized one gets

$$\sum_{\xi_1} P(\xi_1, \xi_2 | \hat{n}_1, \hat{n}_2, \mu_1, \mu_2, \Lambda) = W_2(\xi_2 | \hat{n}_2, \mu_2, \mu_1^0, \Lambda) , \tag{21.15}$$

$$\sum_{\xi_2} P(\xi_1, \xi_2 | \hat{n}_1, \hat{n}_2, \mu_1, \mu_2, \Lambda) = W_1(\xi_1 | \hat{n}_1, \mu_1, \mu_2^0, \Lambda) . \tag{21.16}$$

A theory satisfies the hypothesis of *simple locality* if for any value of the arguments, probabilities satisfy (21.15) and (21.16).

Theorem 1 on p. 264 entails that quantum mechanics, where Λ coincides with the state vector, satisfies simple locality. One can see that any deterministic theory that satisfies simple locality also satisfies strong locality. Indeed if values of P, W_1, and W_2 are restricted to 0 and 1, (21.15) and (21.16) imply (21.14).

In addition to strong and simple locality, we need to introduce the notion of *relativistic locality.* This is the hypothesis that any faster-than-light information transfer is impossible. We then have the following theorem.

Theorem In the context of spin experiments and insofar as state Λ can be prepared at will, simple and relativistic locality are equivalent.

Proof We will show that if simple locality is violated, so is relativistic locality. From this one concludes that relativistic locality implies simple locality. The other half of the proof (i.e. that simple locality implies relativistic locality) can be found in [126].

Suppose that simple locality is violated. For some values of arguments one then has

$$\sum_{\xi_1} P(\xi_1, \xi_2 | \hat{n}_1, \hat{n}_2, \mu_1, \mu_2, \Lambda) \neq W_2(\xi_2 | \hat{n}_2, \mu_2, \mu_1^0, \Lambda). \tag{21.17}$$

[The case where (21.16) does not hold can be analyzed similarly.] Consider a device that produces a large number of pairs of spin 1/2 particles prepared in state Λ. Suppose two observers far from one another each receive one particle of each pair and are equipped with spin measurement apparatus. The second observer is instructed to measure the projection along $\hat{n}_2$ of the spin of every particle reaching him. The first observer is instructed to take one of the following decisions, just before the particles reach him: (i) either measure the spin projection along $\hat{n}_1$ of every particle, (ii) or do no measurement. In case (i) the probability that the second observer finds value ξ_2 is given by the left-hand side of (21.17). In case (ii) the probability is given by the right-hand side. As the probabilities are different, the second observer can, with enough measurements, know with arbitrarily large confidence the first observer's choice. Since there is no restriction on the distance between observers, this device allows information transfer at arbitrarily large speed. ♣

The interpretation of the theorem is straightforward. Because quantum mechanics satisfies simple locality, it is impossible to use spin correlations obeying quantum-mechanical laws to send information faster than the speed of light.

Experimentally, spin correlations appear to obey quantum-mechanical laws and to violate inequalities following from strong locality. We know that strong locality entails simple locality. But what is the precise relation between the two? To answer this question we introduce the following definition [126]. A theory of spin correlations is *complete* if for all values of arguments, probabilities satisfy the equation[5]

$$\begin{aligned} &P(\xi_1, \xi_2 | \hat{n}_1, \hat{n}_2, \mu_1, \mu_2, \Lambda) \\ &= \left\{ \sum_{\xi_2'} P(\xi_1, \xi_2' | \hat{n}_1, \hat{n}_2, \mu_1, \mu_2, \Lambda) \right\} \left\{ \sum_{\xi_1'} P(\xi_1', \xi_2 | \hat{n}_1, \hat{n}_2, \mu_1, \mu_2, \Lambda) \right\}. \end{aligned} \tag{21.18}$$

Completeness means that the probability to obtain a given result upon a spin measurement does not depend on the result found upon measurement

---

[5] It is shown in [14] that the notion of completeness introduced by Einstein, Podolsky and Rosen [75] entails the present one.

of the other spin. No information about a spin measurement can be drawn from a measurement of the other one, other than what is already contained in the state description. Note however that the result of a spin measurement can depend on the direction along which the other spin is measured. Any deterministic theory, for which values of P are restricted to 0 and 1, satisfies completeness.

It is easy to show that strong locality is logically equivalent to the conjunction of simple locality and completeness. Substantial evidence points to the fact that strong locality is violated in nature, while relativistic locality is obeyed. Insofar as state Λ can be prepared at will this means that simple locality is satisfied, hence that completeness is violated. One can show ([228], Chaps. 4 and 10) that violation of (21.18) means that spin correlations cannot be explained by the assumption of a common cause.

It is interesting to note that some hidden-variable theories [114] solve the dilemma differently. These theories, being deterministic, automatically satisfy completeness. They violate simple locality without necessarily violating relativistic locality. This comes from the fact that they do not allow enough control for the preparation of the state Λ.

21.4 The Measurement Problem

As with every physical theory, quantum-mechanical predictions are confirmed or falsified by measurements. In general measurement involves interaction between an object and an apparatus, after which the pointer of the apparatus shows a value. This value is related, in some way, to the state of the object. It is well-defined, in the sense that there is consensus among observers registering it.

In quantum mechanics objects on which measurements are performed are microobjects, whereas apparatus are macroscopic. An interaction process between a microobject and a macroscopic apparatus is a measurement of a dynamical variable A associated with the microobject if the following conditions are met (Sect. 6.3):

i) To each eigenvalue a_i of operator A there correspond one or many values $\alpha(a_i)$ of the apparatus's pointer, so that $\alpha(a_i) \neq \alpha(a_j)$ if $a_i \neq a_j$.
ii) No matter what the microobject's initial state, the pointer shows one of the $\alpha(a_j)$ once the interaction is over.
iii) Whenever the microobject's initial state coincides with an eigenvector $|a_i\rangle$ of A, the pointer, following the interaction, necessarily shows $\alpha(a_i)$.

The value shown by the apparatus is well-defined provided that if $a_i \neq a_j$, values $\alpha(a_i)$ and $\alpha(a_j)$ are easily distinguished.

Quantum mechanics is a theory of microobjects. In a *reductionist* worldview all sense objects are, ultimately, aggregates of microobjects. The laws of microobjects solely determine, at least in principle, the behavior of sense objects.[6]

Let us for the moment adopt this point of view and assume that quantum mechanics is exact. Then it should account for the whole measurement process, in particular the behavior of the macroscopic apparatus. Obviously we have in mind here a possibility in principle. No description of the measurement process can actually pinpoint each of the 10^{25} or so microobjects which together make up the macroscopic apparatus. So let us see if a simplified, and nonetheless convincing, theory of the measurement process characterized by (i), (ii) and (iii) can be given.

In a purely quantum perspective the microobject and measurement apparatus are two quantum systems that can be labeled (1) and (2). The total system's state vector $|\Psi(1,2;t)\rangle$ [or, more generally, the density operator $\varrho(1,2;t)$] satisfies the Schrödinger equation (or the appropriate time-evolution equation). To properly understand the nature of the measurement problem we will first set the discussion in a rather simple framework, by making the following hypotheses:

a) The total system made up of microobject and apparatus is perfectly isolated from the rest of the universe and is described at any time t by a state vector $|\Psi(1,2;t)\rangle$.
b) The state vector represents the state of an individual quantum system.
c) The apparatus is devised to measure a dynamical variable A associated with the microobject, with a discrete and nondegenerate spectrum.
d) To every eigenvalue a_i of the operator A there corresponds one and only one value $\alpha_i = \alpha(a_i)$ of the apparatus's pointer, and just one state vector $|\alpha_i\rangle$ of the apparatus; moreover $\alpha_i \neq \alpha_j$ if $a_i \neq a_j$.
e) The interaction Hamiltonian $H_I(1,2)$ vanishes for $t < -T$ or $t > T$, that is, the interaction between microobject and apparatus happens in the time interval $[-T,T]$.
f) Before interaction, the total system's state vector has the form

$$|\Psi(1,2)\rangle = |\phi_0(1)\rangle \otimes |\alpha_0(2)\rangle\,, \tag{21.19}$$

where $|\phi_0(1)\rangle$ is an arbitrary vector and $|\alpha_0(2)\rangle$ is a vector representing the specific state in which the apparatus is prepared before measurement. $|\phi_0(1)\rangle$ and $|\alpha_0(2)\rangle$ are normalized.

In Sect. 21.5 we will assess the relevance and necessity of these hypotheses and will be able to weaken them considerably. Let us see here whether we can satisfy with them the three characteristics of the measurement process. From hypothesis (d) one sees immediately that (i) is satisfied. On the other hand, hypothesis (a) implies that for all t and Δt there exists a unitary operator $U(t+\Delta t, t)$ that relates $|\Psi(1,2;t+\Delta t)\rangle$ to $|\Psi(1,2;t)\rangle$. In particular

---

[6] For a critique of the reductionist worldview see [145], where a good synthesis of the measurement problem can be found.

$$|\Psi(1,2;T)\rangle = U(T,-T)|\Psi(1,2;-T)\rangle . \qquad (21.20)$$

Let $|a_i\rangle$ be the microobject's initial state. To satisfy (iii) the final state of the apparatus, by virtue of hypothesis (d), must be $|\alpha_i\rangle$. Hence for every i,

$$U(T,-T)\left\{|a_i\rangle \otimes |\alpha_0\rangle\right\} = |\phi_i\rangle \otimes |\alpha_i\rangle . \qquad (21.21)$$

Here there are no restrictions on the microobject's final state $|\phi_i\rangle$. It may be, however, that $|\phi_i\rangle = |a_i\rangle$, which corresponds to a situation where the interaction process does not change the microobject's state. Then we say that we have a *type I* measurement.[7]

Suppose now that the microobject's initial state does not coincide with an eigenvector of A, but is rather a linear combination of vectors $|a_i\rangle$. This means that

$$|\phi_0\rangle = \sum_i c_i |a_i\rangle . \qquad (21.22)$$

Due to the linearity of the evolution operator one must have

$$\begin{aligned} |\Psi(1,2;T)\rangle &= U(T,-T)\left\{\sum_i c_i|a_i\rangle \otimes |\alpha_0\rangle\right\} \\ &= \sum_i c_i\, U(T,-T)\left\{|a_i\rangle \otimes |\alpha_0\rangle\right\} \\ &= \sum_i c_i |\phi_i\rangle \otimes |\alpha_i\rangle . \qquad (21.23) \end{aligned}$$

According to hypothesis (b), the vector $|\Psi(1,2;t)\rangle$ represents the state of an individual quantum system. Thus the total system is in a superposition of states where the pointer of the apparatus shows distinct values. This means that in general, it is impossible to satisfy property (ii) of the measurement process, according to which the apparatus must show a precise α_j following interaction. This, under specific hypotheses, is the quantum-mechanical measurement problem.[8]

In Sects. 21.5 and 21.6 we will analyse several proposals to either solve or eliminate the measurement problem. It is useful, however, first to examine a specific process. The example is due to London and Bauer [148]. Although it does not satisfy every hypothesis given above, it does correspond rather closely to the general scheme.

---

[7] A type I measurement is usually defined as one whose immediate repetition reproduces the original result ([124], Sect. 11.4). In the case where eigenvalue a_i is degenerate, this only implies that $|\phi_i\rangle$ is one of the eigenvectors associated with a_i, and not necessarily the original vector. Our definition is more restrictive.

[8] The bibliography on the measurement problem is very large. The reader is more specifically referred to [97], [124], Chap. 11 and [235].

Consider a measurement of an atom's magnetic moment by means of a Stern–Gerlach device. Let $\boldsymbol{R}$ be the coordinates of the atom's center of mass, in the apparatus's reference frame. Denote collectively by ϱ the atom's internal (in general multidimensional) coordinates, and by π the corresponding momentum operators. Let M be the atom's mass, $\boldsymbol{\mu}$ the total magnetic moment operator, $H_0(\varrho, \pi)$ the internal Hamiltonian and $\boldsymbol{B}$ an external magnetic field. The Schrödinger equation is given by

$$\mathrm{i}\hbar\frac{\partial}{\partial t}\Psi(\varrho, \boldsymbol{R}, t) = \left\{-\frac{\hbar^2}{2M}\nabla_R^2 + H_0(\varrho, \pi) - \boldsymbol{\mu}\cdot\boldsymbol{B}\right\}\Psi(\varrho, \boldsymbol{R}, t). \tag{21.24}$$

In general $\boldsymbol{\mu}$ involves orbital angular momentum and spin. Likewise Ψ includes orbital as well as spin functions. There will be no need, however, to specify them in detail.

A Stern–Gerlach device works through a magnetic field gradient. Thus $\boldsymbol{B} = \boldsymbol{B}(\boldsymbol{R})$. We assume $\boldsymbol{B}$ is time independent and essentially constant on atomic sizes. Moreover we assume (Sect. 4.1) that for all practical purposes the field direction coincides with the one of its gradient. Hence

$$\boldsymbol{B}(\boldsymbol{R}) = B(\boldsymbol{R})\hat{n}, \qquad \boldsymbol{\nabla} B(\boldsymbol{R}) = |\boldsymbol{\nabla} B(\boldsymbol{R})|\,\hat{n}, \tag{21.25}$$

where $\hat{n}$ is a constant unit vector.

If $\boldsymbol{B}(\boldsymbol{R})$ were strictly spatially constant, it would be possible to separate Schrödinger's equation by letting

$$\Psi(\varrho, \boldsymbol{R}, t) = \phi(\varrho)v(\boldsymbol{R}, t). \tag{21.26}$$

We would then obtain the following equation for $\phi(\varrho)$:

$$\{H_0(\varrho, \pi) - \boldsymbol{\mu}\cdot\boldsymbol{B}\}\,\phi(\varrho) = E\phi(\varrho). \tag{21.27}$$

In principle this equation can be solved. In practice it can be solved approximately by the methods developed in Chaps. 17 and 18. Let $\bar{E}$ be the lowest eigenvalue of (21.27) in the absence of a magnetic field. That eigenvalue is degenerate owing to the $2j+1$ values of projection of total angular momentum. In the presence of a magnetic field the degeneracy is removed. Eigenvalues are given by (18.72), that is,

$$E_m = \bar{E} + m\gamma B, \tag{21.28}$$

where $-j \le m \le j$ and $\gamma = g\mu_B$. Eigenfunctions corresponding to different values of m can be denoted by $\phi_m(\varrho)$.

In the case where $\boldsymbol{B}$, although not constant, does not change much on atomic lengths, the functions $\phi_m(\varrho)$ still represent, for all practical purposes, the dependence of Ψ on relative coordinates. Furthermore, the atom's energy is approximately given by

$$E_m = \bar{E} + m\gamma B(\boldsymbol{R}), \tag{21.29}$$

where $\boldsymbol{R}$ is a parameter. To a good approximation, the Schrödinger equation is then solved by a wave function like $\Psi_m(\varrho, \boldsymbol{R}, t) = \phi_m(\varrho) v_m(\boldsymbol{R}, t)$, where $v_m(\boldsymbol{R}, t)$ satisfies

$$\mathrm{i}\hbar \frac{\partial}{\partial t} v_m(\boldsymbol{R}, t) = \left\{ -\frac{\hbar^2}{2M} \nabla_R^2 + \bar{E} + m\gamma B(\boldsymbol{R}) \right\} v_m(\boldsymbol{R}, t). \tag{21.30}$$

Take the origin of coordinates $\boldsymbol{R}$ somewhere near the Stern–Gerlach device's center. One can write

$$B(\boldsymbol{R}) \approx B_0 + |\boldsymbol{\nabla} B|_0 \, (\hat{n} \cdot \boldsymbol{R}). \tag{21.31}$$

Equation (21.30) becomes

$$\begin{aligned} &\mathrm{i}\hbar \frac{\partial}{\partial t} v_m(\boldsymbol{R}, t) \\ &\quad = \left\{ -\frac{\hbar^2}{2M} \nabla_R^2 + \left(\bar{E} + m\gamma B_0 \right) + \left(m\gamma \, |\boldsymbol{\nabla} B|_0 \right) \hat{n} \cdot \boldsymbol{R} \right\} v_m(\boldsymbol{R}, t). \end{aligned} \tag{21.32}$$

This is the Schrödinger equation for a particle acted upon by a spatially constant and time-independent force, with intensity proportional to m. Solutions of that equation are wave packets moving away from a straight line. For each value of m the trajectory of the wave packet's center follows from Ehrenfest's equations and corresponds to the trajectory of a classical particle. This analysis thus reproduces the behavior of an atom carrying magnetic moment which goes through a Stern–Gerlach device.

Let us now bring the analysis together with the general discussion of the measurement process. The dynamical variable to be measured is the projection along $\hat{n}$ of the total magnetic moment $\boldsymbol{\mu}$. Eigenfunctions of this operator coincide with the $\phi_m(\varrho)$. The $v_m(\boldsymbol{R}, t)$ correspond to apparatus vectors $|\alpha_i\rangle$, since the pointer is associated with the mean value of the coordinate $\boldsymbol{R}$. Indeed if $\phi_m(\varrho) v_{\text{in}}(\boldsymbol{R}, t)$ is the atom's initial wave function, then the final wave function is given by $\phi_m(\varrho) v_m(\boldsymbol{R}, t)$. To different values of m there correspond wave packets centered on different mean values of $\boldsymbol{R}$.

But suppose the atom's initial wavefunction involves a superposition of several values of m, so that

$$\Psi_{\text{in}}(\varrho, \boldsymbol{R}, t) = \sum_m c_m \phi_m(\varrho) v_{\text{in}}(\boldsymbol{R}, t). \tag{21.33}$$

Then the final wave function will be given by

$$\Psi(\varrho, \boldsymbol{R}, t) = \sum_m c_m \phi_m(\varrho) v_m(\boldsymbol{R}, t). \tag{21.34}$$

The final state thus represents a superposition of states wherein the atom goes in macroscopically distinct directions. Therefore the apparatus does not show a precise value after interaction. The quantum description of measurement does not seem capable of satisfying property (ii).

21.5 Search for a Solution

The crux of the measurement problem lies in the linearity of the evolution operator. In a nutshell, if microobject's state $|a_j\rangle$ entails apparatus's state $|\alpha_j\rangle$, then a linear combination of microobject states $|a_i\rangle$ entails a linear combination of apparatus states $|\alpha_i\rangle$. The latter each correspond to specific and mutually exclusive values of the pointer. A superposition of states wherein the pointer of the apparatus has distinct values at once appears to contradict the requirement to obtain a well-defined value for a measurement result and, more simply, to contradict everyday experience.

The measurement problem was recognized early, by von Neumann [230] among others. He realized that unitary evolution leads to superposition of macroscopically distinct states. Furthermore, he saw that there is no use to introduce a second apparatus to measure the value of the first's pointer. Indeed inasmuch as the evolution of the total system (microobject, first and second apparatus) is unitary, the second apparatus would also end up in a superposition of macroscopically distinct states. The solution proposed by von Neumann essentially consists in postulating that the Schrödinger equation no longer holds at the time of measurement. But what is this precisely? It is the instant when an observer becomes conscious of the pointer's value. The state vector of the total system made up of microobject and apparatus is correctly given by (21.23) up to the time when a conscious subject observes the pointer. At this instant, the state vector undergoes the following abrupt change:

$$\sum_i c_i |\phi_i\rangle \otimes |\alpha_i\rangle \rightarrow |\phi_j\rangle \otimes |\alpha_j\rangle \,. \tag{21.35}$$

The final state corresponds to a well-defined value α_j of the pointer. The probability of obtaining α_j is equal to $|c_j|^2$, the absolute square of the coefficient associated with $|a_j\rangle$ in the initial state. The hypothesis represented by (21.35) is called the *projection postulate.* The abrupt transition from a linear combination to one of its components is known as the *collapse of the state vector.*

Von Neumann's hypothesis is ingenious. Its success is largely independent of where the border between microobject and measurement apparatus, or the border between apparatus and conscious subject, lies. The process represented by (21.35), however, seems closer to a requirement of perception than to a physical mechanism. It thus appears to reinstate the mind-body dualism that natural sciences had largely eliminated. In their lucid presentation of von Neumann's theory, London and Bauer ([148], p. 252) explain the conscious subject's action in the following way:

> The observer has a completely different impression. For him it is only the object and the apparatus that belong to the external world, to what he calls "objectivity." By contrast he has *with himself* relations

> of a very special character. He possesses a characteristic and quite familiar faculty which we can call the "faculty of introspection." He can keep track from moment to moment of his own state. By virtue of this "immanent knowledge" he attributes to himself the right to create his own objectivity – that is, to cut the chain of statistical correlations summarized in $\sum_i c_i|\phi_i\rangle \otimes |\alpha_i\rangle \otimes |O_i\rangle$[9] by declaring "I am in the state $|O_j\rangle$" or more simply, "I see α_j," or even directly, "$A = a_j$."

The breakdown of the Schrödinger equation and unitary evolution of the state vector occurs, according to von Neumann, upon intervention of the conscious subject. In a similar analysis, Wigner [239] associates this discontinuity more generally with all living processes. He believes that living processes should be described by nonlinear equations, which entail a nonunitary evolution of the state vector. We should point out that Wigner's proposal is more and more difficult to maintain, since the linear character of quantum systems' evolution is confirmed better and better [39], [231].

Von Neumann's hypothesis may be bold, but not so much as the solution proposed by Everett [80]. In its usual formulation, the measurement problem consists in explaining how the superposition $\sum_i c_i|a_i\rangle \otimes |\alpha_i\rangle$ ends up in only one component $|a_j\rangle \otimes |\alpha_j\rangle$. Everett postulates that all components simultaneously exist, but in different universes. Thus each measurement gives rise to a subdivision of the universe in many different ones, at least one universe corresponding to each measurement result. This interpretation appears to be consistent. It is difficult, however, to truly believe it. Note that it predicts that the state vector reduction happens at a specific moment, after which it is irreversible.

Von Neumann's theory rejects the universal validity of Schrödinger's equation. There are physical systems (conscious subjects or, in Wigner's approach, living beings) that are not quantum systems.

In a very different way, the Copenhagen interpretation also rejects the universal validity of quantum mechanics. According to Bohr, the concept of macroscopic apparatus obeying the laws of classical mechanics is necessary for the formulation of quantum mechanics. If a macroscopic apparatus is viewed as an aggregate of microobjects, that aggregate no longer works as an apparatus. The aim of quantum mechanics is to predict the behavior of macroscopic apparatus interacting with microobjects. In the Copenhagen interpretation the measurement problem as put forth by von Neumann simply does not arise.

In contrast with Bohr et von Neumann, many attempt to solve the measurement problem while maintaining the universal validity of quantum me-

---

[9] $|O_i\rangle$ represents the observer's state. The original notation is adapted to ours.

chanics and the Schrödinger equation. We now examine the question in this angle.[10]

The formulation of the measurement problem given in Sect. 21.4 makes use of the six hypotheses (a)–(f). But none looks really necessary to the definition of a measurement process. Could it be that the problem is simply the consequence of too restrictive hypotheses?

Let us first look at hypotheses (c) and (e). They relate to the operator's spectrum and the interaction time interval. It is not difficult to see that they have no other purpose than to simplify the discussion. The measurement problem can be formulated even if the interaction Hamiltonian is more general than in (e), or for dynamical variables with continuous or degenerate spectra. Besides, the problem would remain even if hypothesis (c) could not be weakened, as there are in nature dynamical variables with discrete and nondegenerate spectra.

Although this may not be so obvious, one can also do away with hypotheses (d) and (f). From (d) one should keep that $\alpha(a_i) \neq \alpha(a_j)$ for $a_i \neq a_j$, which is part of the definition of the measurement process. But the rest of the hypothesis is unrealistic. A macroscopic apparatus approximately involves 10^{25} degrees of freedom. To a macroscopic value of the pointer (specified up to a small macroscopic interval) there correspond a huge number of quantum states (16.90) and a huge number of eigenvalues of the quantum dynamical variable associated with the pointer. Hypothesis (f) is no more realistic. It is much more reasonable to initially associate a statistical mixture with the apparatus than a pure state. Thus the total system's initial state should be described by a density operator rather than by a state vector, so that

$$\varrho_{\text{in}}(1,2) = \sum_k w_k \left(|\phi_0(1)\rangle\langle\phi_0(1)|\right) \otimes \left(|\alpha_{0k}(2)\rangle\langle\alpha_{0k}(2)|\right) . \tag{21.36}$$

Vectors $|\alpha_{0k}(2)\rangle$ represent quantum states compatible with the initial macroscopic configuration of the apparatus.

We shall thus replace hypothesis (f) by a hypothesis (f′) according to which the total system's initial state is given by (21.36). Likewise hypothesis (d) can be replaced by a hypothesis (d′) which reads as:

d′) To each eigenvalue a_i of operator A there correspond a large number of values α_{ik} of the pointer, and a large number of state vectors $|\alpha_{ik}\rangle$ of the apparatus; moreover α_{ik} and $\alpha_{jk'}$ are macroscopically distinct if $a_i \neq a_j$.

Note that the range of index k is the same on $|\alpha_{0k}\rangle$ and $|\alpha_{ik}\rangle$. Indeed the two vectors are related by the evolution operator as

$$U(T,-T)\left\{|a_i\rangle \otimes |\alpha_{0k}\rangle\right\} = |\phi_{ik}\rangle \otimes |\alpha_{ik}\rangle . \tag{21.37}$$

---

[10] See also [94] which, in addition to the avenue we are investigating, surveys two other solutions in terms of hidden variables and "spontaneous localization."

Hypotheses (d$'$) and (f$'$) are natural and much less restrictive than (d) and (f). Nonetheless one can show ([12], [240]) that even under (d$'$) and (f$'$), a microobject's initial state like $|\phi_0(1)\rangle = \sum_i c_i |a_i\rangle$ inevitably leads to a final state wherein the apparatus is in a superposition of macroscopically distinct states. Substituting (d$'$) and (f$'$) to (d) and (f) therefore does not solve the measurement problem.

For all practical purposes, all investigators agree on the fact that weakening hypotheses (c), (d), (e) and (f) does not solve the measurement problem. This is not so, however, with hypotheses (a) and (b). Some believe the measurement problem disappears if (a) or (b) is eliminated or weakened.

Let us first look at (b). To the individual interpretation of the state vector one can substitute the statistical or the epistemic interpretation. In the statistical interpretation vector $\sum_i c_i |a_i\rangle \otimes |\alpha_i\rangle$, associated with microobject and apparatus after interaction,[11] represents an ensemble of similarly prepared systems. To correctly analyze this ensemble it is useful to compare it to another one, described by the density operator

$$\varrho(1,2;T) = \sum_i |c_i|^2 \left(|a_i\rangle\langle a_i|\right) \otimes \left(|\alpha_i\rangle\langle\alpha_i|\right). \tag{21.38}$$

Equation (21.38) represents an ensemble where each apparatus shows a well-defined value, a fraction $|c_j|^2$ of all apparatus showing value α_j. The measurement problem would be solved if one could show that the initial state $\sum_i c_i |a_i\rangle \otimes |\alpha_0\rangle$ transforms into a proper mixture represented by $\varrho(1,2;T)$. Unfortunately no unitary operator can transform a pure state into a statistical mixture.

Can one claim that vector $\sum_i c_i |a_i\rangle \otimes |\alpha_i\rangle$ also represents an ensemble in which a fraction $|c_j|^2$ of all apparatus show value α_j? Maybe, but a positive answer entails the existence of hidden variables. If values of dynamical variables coincide with eigenvalues of the corresponding operators, a complete theory of the measurement process should explain why a measurement of the $\hat{z}$ component of spin, for instance, induces the parameters which determine that component (rather than another one) to previously adjust.

To the individual interpretation of the state vector one can also substitute the epistemic interpretation. The state vector then represents the knowledge we have about the quantum system. Before observing the apparatus, our knowledge of the system is determined by the vector $|\Psi(1,2;T)\rangle$. After we have seen the pointer, it is represented by one of the $|a_j\rangle \otimes |\alpha_j\rangle$, associated with a probability equal to $|c_j|^2$. Does this explain the transition between the pure state $|\Psi(1,2;T)\rangle$ and the statistical mixture $\varrho(1,2;T)$? Even in the epistemic representation, the quantum-mechanical formalism should correctly describe the evolution of the state vector or the density operator. Stating that our knowledge of the system's state changes abruptly at the time of measurement sets the problem, but does not solve it.

---

[11] See (21.23). We assume for simplicity that the measurement is of type I.

21.6 Decoherence

There remains to investigate hypothesis (a). Like (d) and (f) it is unrealistic. One cannot completely isolate a quantum system from the rest of the universe. But can we isolate it well enough? Before answering this question let us go back to the measurement of an atom's spin by a Stern–Gerlach device. In a 0.1 m long apparatus the magnetic field gradient should be about 10^2 T m^{-1} to produce a beam deviation of 0.01 radian. On the other hand, a variation of 10^{-6} T m^{-1} in the field gradient will in 0.1 m produce a change of about 1 radian in the phase of the wave packet $v_m(\boldsymbol{R}, t)$. For the phase of the wave packet and, consequently, the coefficients c_i of the final wave function to be well determined, it is thus necessary to control the field with extraordinary precision (better than one part in 10^8). Hence even in a situation where the quantum system is completely isolated it is very difficult, actually, to tell the pure state $|\Psi(1,2;T)\rangle$ from the statistical mixture $\varrho(1,2;T)$.

Let us now weaken hypothesis (a) and try to take the environment into account. For this purpose we consider a total system made of three subsystems, i.e. microobject, measurement apparatus and environment which we label (1), (2) and (3) respectively. We assume the total system is isolated and described at any time by a state vector $|\Psi(1,2,3;t)\rangle$. As a matter of fact, we apply hypothesis (a) to the system (1) + (2) + (3) rather than to the system (1) + (2).

The measurement apparatus is a macroscopic system. Many of its parts, for example its base, lid, handles, etc., are not essential to its proper working. This means there is considerable freedom where to set the dividing line between apparatus and environment, i.e. between subsystems (2) and (3). Besides, recall that the characterization of a macroscopic object requires the identification of a certain number of collective coordinates, introduced in Sect. 16.8. We assume the state space associated with subsystem (2) corresponds to the space where the apparatus's collective coordinates and conjugate momenta act. Thus all microscopic coordinates of the apparatus belong to subsystem (3). Of course the pointer's position is one of the collective coordinates.

During the measurement process the microobject and the apparatus interact by means of a term $H_I(1,2)$ similar to the one we considered above. The apparatus and the environment interact by means of a term $H_I(2,3)$. At the end of the process the total system's state vector is given by

$$|\Psi(1,2,3;T)\rangle = U(T,-T)|\Psi(1,2,3;-T)\rangle . \tag{21.39}$$

If the microobject's initial state is represented by (21.22) one gets, by analogy with (21.23), that the total system's final state is represented by

$$|\Psi(1,2,3;T)\rangle = \sum_i c_i|a_i\rangle \otimes |\alpha_i\rangle \otimes |e_i\rangle , \tag{21.40}$$

where $|\alpha_i\rangle$ takes all collective coordinates into account and the normalized vector $|e_i\rangle$ denotes a state of the environment.[12]

The microobject and measurement apparatus, taken together, can be considered as a subsystem of the total system (1) + (2) + (3). What is the state of this subsystem? Following Sect. 12.4, this state is represented by a density operator obtained by computing the partial trace of the total system's density operator. Hence,

$$\begin{aligned}\varrho(1,2;T) &= \mathrm{Tr}_3\left\{|\Psi(1,2,3;T)\rangle\langle\Psi(1,2,3;T)|\right\} \\ &= \mathrm{Tr}_3\left\{\sum_{i,j} c_i c_j^* \left(|a_i\rangle\langle a_j|\right)\otimes\left(|\alpha_i\rangle\langle\alpha_j|\right)\otimes\left(|e_i\rangle\langle e_j|\right)\right\} \\ &= \sum_{i,j} c_i c_j^* \langle e_j|e_i\rangle \left(|a_i\rangle\langle a_j|\right)\otimes\left(|\alpha_i\rangle\langle\alpha_j|\right). \qquad (21.41)\end{aligned}$$

Assume that $|e_i\rangle$ and $|e_j\rangle$, i.e. state vectors of the environment associated with distinct values of the pointer, are orthogonal. Then $\langle e_j|e_i\rangle = \delta_{ij}$ and one gets

$$\varrho(1,2;T) = \sum_i |c_i|^2 \left(|a_i\rangle\langle a_i|\right)\otimes\left(|\alpha_i\rangle\langle\alpha_i|\right), \qquad (21.42)$$

which coincides with (21.38). Thus interaction with the environment has, so it seems, explained the transition between a pure state and a statistical mixture.

We will soon examine the hypothesis on vectors $|e_i\rangle$. But first we must face several objections. The density operator (21.42) represents an improper mixture. So it is not clear that one can associate with it precise values of the pointer. Moreover, the total system's state vector (21.40) still involves a coherent superposition of macroscopically distinct states. To the last remark one can reply that system (3) is not isolated. Thus correlations rapidly vanish in a wider and wider environment. The argument going from (21.40) to (21.42) can be made not only for a specific dividing line between systems (2) and (3), but also for any reasonable one, and any definition of the environment. As a matter of fact, there is no system (1) + (2) + (3) that is in a pure state.[13] Furthermore, it is possible that interaction with the environment singles out the decomposition (21.42) of the density operator instead of

---

[12] Obviously (21.40) is a highly simplified representation of the final state. A given value of the pointer corresponds to a large number of values of the collective coordinates, and these states are associated with numerous states of the environment.

[13] This statement should be qualified if the environment includes the whole universe. It is hardly conceivable, however, that the solution of the measurement problem hinges on whether the whole universe is in a pure state or a statistical mixture. See [88] for a discussion of the quantum state of the universe.

any other representation, so that for all practical purposes we have a proper mixture.

Let us now come back to the orthogonality of vectors $|e_i\rangle$. Several considerations suggest that it is a phenomenon happening in all dissipative systems. Specifically, consider the state space associated with the collective coordinates of a macroscopic object, and suppose there is dissipative interaction between collective and microscopic coordinates. Suppose also that at time t_0 the state vector in the space of collective coordinates consists in a superposition of macroscopically distinct states. Then it appears that the total system's state evolves so that very quickly, the nondiagonal terms of the reduced density operator vanish. This phenomenon is called *decoherence* ([173], Chap. 7; [246]; [247]; [248]).

A simple example will make the argument more concrete. Consider a macroscopic bob (of a few hundred grams perhaps) suspended by a string of negligible mass. We assume the pendulum oscillates in a plane, and in a high vacuum. Damping is thus weak and due to the string's elasticity and friction. The angle between the string and the vertical axis is an appropriate collective coordinate. We denote by $|\phi_\theta\rangle$ the quantum state representing an oscillating motion of amplitude θ.

Suppose that at $t = 0$, the pendulum is prepared in state $|\phi_{\theta(0)}\rangle$. If microscopic coordinates are taken into account the total system's state is given by

$$|\chi(0)\rangle = |\phi_{\theta(0)}\rangle \otimes |v\rangle\,, \tag{21.43}$$

where $|v\rangle$ refers to microscopic coordinates. Vector $|v\rangle$ represents a state where the internal energy of the bob and string is well-defined (up to small fluctuations) and equal to E.

As time goes on, friction transforms the pendulum's kinetic energy into internal energy. At time t the total system's state is therefore represented by

$$|\chi(t)\rangle = \sum_i c_i(t) |\phi_{\theta_i(t)}\rangle \otimes |v_i\rangle, \qquad \sum_i |c_i(t)|^2 = 1. \tag{21.44}$$

Values $\theta_i(t)$ here are all close to the value $\theta(t)$ obtained by solving the classical damped-oscillator problem. The normalized vectors $|v_i\rangle$ correspond to energies close to $E + \Delta E(t)$, where $\Delta E(t)$ is the total energy dissipated by the oscillator up to t. It is important to realize that these energy levels are extremely close. Their number can be estimated by a formula like (16.90).

Let us now turn to a different situation. Suppose that at $t = 0$ the pendulum is prepared in a superposition of two macroscopically distinct states $|\phi_{\theta(0)}\rangle$ and $|\phi_{\eta(0)}\rangle$. The total system's state is then given as

$$|\chi(0)\rangle = \left\{\alpha|\phi_{\theta(0)}\rangle + \beta|\phi_{\eta(0)}\rangle\right\} \otimes |v\rangle\,. \tag{21.45}$$

At time t,

$$|\chi(t)\rangle = \alpha \sum_i c_i(t)|\phi_{\theta_i(t)}\rangle \otimes |v_i\rangle + \beta \sum_j d_j(t)|\phi_{\eta_j(t)}\rangle \otimes |u_j\rangle . \tag{21.46}$$

Vectors $|v_i\rangle$ correspond to energies close to $E + \Delta E(t)$, while vectors $|u_j\rangle$ correspond to energies close to $E + \Delta \bar{E}(t)$. These energies are different. Indeed the energy $\Delta E(t)$ dissipated by a pendulum of initial amplitude $\theta(0)$ is different from the energy $\Delta \bar{E}(t)$ dissipated by a similar pendulum of initial amplitude $\eta(0)$.

Let us compute, at time t, the density operator of the subsystem associated with collective coordinates. Insofar as the difference between ΔE and $\Delta \bar{E}$ is much larger than energy fluctuations, one has $\langle v_i|u_j\rangle = 0$ and one finds that

$$\begin{aligned} \varrho(t) &= \mathrm{Tr}_{\mathrm{micro}} \{ |\chi(t)\rangle\langle\chi(t)| \} \\ &= |\alpha|^2 \sum_{i,k} c_i(t) c_k^*(t) \langle v_k|v_i\rangle |\phi_{\theta_i(t)}\rangle\langle\phi_{\theta_k(t)}| \\ &\quad + |\beta|^2 \sum_{j,l} d_j(t) d_l^*(t) \langle u_l|u_j\rangle |\phi_{\eta_j(t)}\rangle\langle\phi_{\eta_l(t)}| \, . \end{aligned} \tag{21.47}$$

Energy fluctuations in the string and bob are extremely small. Inasmuch as $\theta(0)$ and $\eta(0)$ differ appreciably, decoherence follows very quickly. For all practical purposes it is irreversible.

Decoherence seems to be a property of all dissipative systems. It matters little whether dissipation occurs in a internal environment (as in the example of the pendulum) or in an external one. An instance of decoherence by external environment is analyzed by Joos and Zeh [130]. They examine the density operator of a particle in the basis of eigenvectors of the position operator. For a particle of radius more than 10^{-7} m, Joos and Zeh show that collisions with air molecules even at very low pressure, or merely interaction with sunlight, is enough to make the density operator diagonal in less than 10^{-6} s.

Decoherence is a consequence of dissipation. This suggests the possibility that in nondissipative systems decoherence might not occur, or be considerably slowed down. In other words, some of these systems might be observed in a superposition of macroscopically (or mesoscopically) distinct states. In this connection a beryllium atom was prepared in a superposition of two distinct quantum states, and these two states were then separated by a distance of 800 Å [164]. Decoherence was observed associated with a field of a few photons only, inside a cavity [46]. This field of investigation began just recently but could soon undergo rapid progress [144], [186], [225], [242].

21.7 True Propositions and Reliable Propositions

The phenomenon of decoherence entails that in a measurement process, the density operator of microobject and apparatus soon becomes diagonal in the basis determined by the pointer of the apparatus. Does this mean that quantum mechanics predicts that the pointer shows a precise value at the end of the process? There is no consensus on that among investigators.

It is clear, however, that quantum mechanics does not predict the specific value that an apparatus will show. It describes the final state of microobject and apparatus by means of a density operator associated with a statistical mixture. This suggests that the statistical interpretation of quantum mechanics may be the most natural one. Quantum mechanics does not directly refer to individual systems but to statistical ensembles of systems.[14]

The notion of consistent histories, introduced in Sect. 12.6, provides a different formulation of the main results of measurement theory. In this context the family of histories is built as follows. Time t_0 corresponds to the moment when microobject and apparatus are prepared in initial state $\varrho(t_0)$. Instants t_1 and t_2 correspond to the beginning and end of the measurement process, respectively. One associates with t_1 the projectors P_{a_i} that correspond to different eigenvalues of the microobject's dynamical variable A. One associates with t_2 the projectors P_{α_i} that correspond to different values of the pointer. If the first measurement is followed by a second one, projectors P_{b_k} and P_{β_k} are similarly introduced at two instants t_3 and t_4 connected with the beginning and end of the second measurement. One can show that in every case one gets a complete family of consistent histories.

With this family of histories and the ensuing logic, Omnès ([173], Chap. 8) obtains the following results.

Theorem 1 The proposition that the pointer shows value α_i and the proposition that A has value a_i are logically equivalent.

Theorem 2 The probability that the pointer shows value α_i is equal to the probability that A has value a_i.

Theorem 3 If the first measurement is of type I, an immediately repeated measurement of the same dynamical variable yields the same result.

Theorem 4 If the first measurement is of type I and yields value a_i, then the probability of results of the second one can be calculated by means of the density operator

$$\varrho_{\text{fi}}(t_2) = \frac{P_{a_i}\varrho_{\text{in}}(t_1)P_{a_i}}{\text{Tr}\left\{P_{a_i}\varrho_{\text{in}}(t_1)P_{a_i}\right\}}, \tag{21.48}$$

where $\varrho_{\text{in}}(t_1)$ represents the microobject's initial state at the beginning of the first measurement.

---

[14] Here we have in mind those versions of the statistical interpretation that do not involve hidden variables. For a different analysis see Omnès ([173], Chap. 8) who rejects the statistical interpretation and introduces a postulate on the uniqueness of reality.

These results express the content of (21.42) in a formal way. Note that ϱ_{fi} and ϱ_{in} act in the microobject's state space. Theorem 4 embodies state vector reduction, here obtained as a consequence of measurement theory and the process of decoherence. We point out, however, that these results would not all be valid if the family of histories was built differently.

The logical structure based on consistent families of histories leads to a precise definition of the truth of propositions about quantum properties. First define a *fact* as a proposition asserting the real value of a collective coordinate of a classical object. Omnès ([173], Chap. 8) states that all present facts are true, as well as all their past consequences. A consistent logic is *sensible* if it includes all present facts and their past consequences, together with their negation. A proposition about a quantum property is then *true* if it meets the following two conditions:

i) It can be added to every sensible logic without breaking its consistency.
ii) In such an extended logic it is equivalent to a fact.

One can show that if at the end of a measurement the pointer shows value α_i, then it is true that dynamical variable A has value a_i at the time of measurement.

It is very important to note that the proposition that A has value a_i is not true, in the above sense, before the beginning of measurement. It is true at the time of measurement only. On the other hand, a proposition is *reliable* if it meets the following conditions:

i) It can be added to some sensible logics without breaking their consistency; there exist, however, sensible logics to which it cannot be so added.
ii) In sensible logics where it enters consistently, the proposition follows from a fact.

One can show that the proposition that A has value a_i is reliable before the measurement process.

The concept of reliable proposition sheds new light on the Einstein–Podolsky–Rosen argument presented in Sect. 21.1.[15] Suppose a system of two spin 1/2 particles is prepared in the singlet state (21.3). At time t_1 the component $S_n(1)$ is measured and the value $+\hbar/2$ is obtained. At time t_3 the component $S_{n'}(2)$ is measured and the value $+\hbar/2$ is obtained. One can show that at any time t_2 such that $t_1 < t_2 < t_3$, the following two propositions are reliable: (i) the value of $S_{n'}(2)$ is equal to $+\hbar/2$; (ii) the value of $S_n(2)$ is equal to $-\hbar/2$. Of course these two propositions belong to different logics. Thus Einstein, Podolsky and Rosen's elements of physical reality do not correspond to true propositions, but only to reliable ones.

In spite of its interest, Omnès's notion of truth has been criticized ([67], Sect. 11.4; [70]). Let A be an operator that commutes with the Hamiltonian of a quantum system. Suppose the system is prepared so that its state vector

---

[15] See also [103].

is an eigenvector of A with eigenvalue a_i. Intuitively one would think that it is true that the value of A is equal to a_i, since any subsequent measurement will yield this result. But such is not the case according to Omnès's criterion. The proposition is merely reliable. Suppose then that the criterion of truth is changed so that true propositions coincide with consequences of state preparations. The argument of Einstein, Podolsky and Rosen then shows that the truth of a proposition can depend on a preparation made instantaneously at an arbitrary large distance. This does not really correspond to our notion of truth.

Intuitive notions and common sense are extracted from daily experience and its classical-mechanical formalization. It should come as no surprise that they can never be reconciled with quantum mechanics. Nonetheless the quantum-mechanical formalism is precise and consistent, and it is one of the best confirmed of physical theories. Let us conclude with d'Espagnat ([67], p. xvi) that "whoever tries to form an idea of the world – and of man's position within the world – has to take the findings of quantum physics most seriously into account."