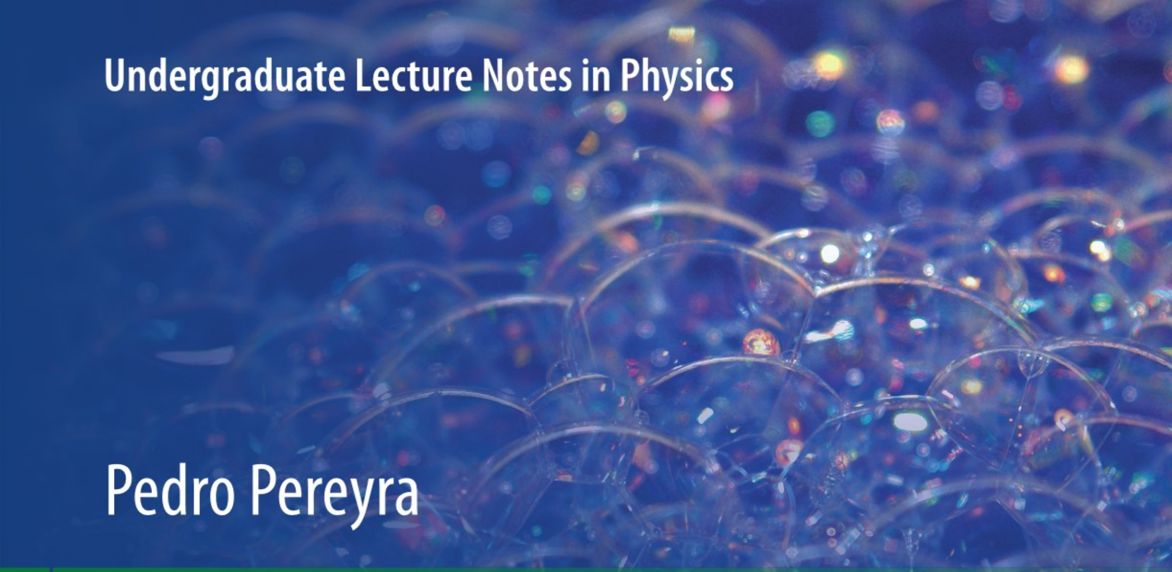


Undergraduate Lecture Notes in Physics

Pedro Pereyra

Fundamentals of Quantum Physics

Textbook for Students of Science and Engineering

Springer

Undergraduate Lecture Notes in Physics

For further volumes:
http://www.springer.com/series/8917

Undergraduate Lecture Notes in Physics (ULNP) publishes authoritative texts covering topics throughout pure and applied physics. Each title in the series is suitable as a basis for undergraduate instruction, typically containing practice problems, worked examples, chapter summaries, and suggestions for further reading.

ULNP titles must provide at least one of the following:

- An exceptionally clear and concise treatment of a standard undergraduate subject.
- A solid undergraduate-level introduction to a graduate, advanced, or non-standard subject.
- A novel perspective or an unusual approach to teaching a subject.

ULNP especially encourages new, original, and idiosyncratic approaches to physics teaching at the undergraduate level.

The purpose of ULNP is to provide intriguing, absorbing books that will continue to be the reader's preferred reference throughout their academic career.

Pedro Pereyra

Fundamentals of Quantum Physics

Textbook for Students of Science and Engineering

Springer

Pedro Pereyra Padilla
Física Teórica y Materia Condensada
Universidad Autónoma Metropolitana
Av. S. Pablo Del Azcapotzalco 180
02200, México D.F.
Mexico

ISSN 2192-4791 ISSN 2192-4805 (electronic)
ISBN 978-3-642-29377-1 ISBN 978-3-642-29378-8 (eBook)
DOI 10.1007/978-3-642-29378-8
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012938636

© Springer-Verlag Berlin Heidelberg 2012
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

To my daughters and son

Naira Citlalli
Jaina Erandi Mariana
Sebastián Fernando

Preface

Various reasons concurred in my decision of writing this book. Some years ago Freeman Dyson, reasoning on the process of learning and teaching quantum theory, came out with the idea that a physics student, after learning the tricks of the quantum formalism and getting right answers, “begins to worry because he does not understand what he is doing”. The student, says Dyson,[1] “has no clear physical picture in his head and tries to arrive at a physical explanation for each of the mathematical tricks”. He gets discouraged and after some months of unpleasant and strenuous time, he suddenly says: “I understand now that there isn’t anything to be understood”. What happens? Dyson suggests that we learn to think in quantum-mechanical language and no longer try to explain in terms of pre-quantum conceptions. As an undergraduate student, facing my first quantum-mechanics textbook, I had similar feelings. I felt lost within a mathematical formalism with abstract objects and concepts. Bras, kets, and operators, came seasoned with counterintuitive statements that shook my own conceptions of nature. Some years later, when I was ready to accept that there was nothing to understand, Thomas A. Brody[2] organized in Mexico a graduate students’ seminar on the interpretations of the mathematical formalism of quantum mechanics; a seminar to show that in the quantum theory, besides the facts and excellent agreement with experimental results, one has to recognize the existence of open epistemological problems. After those years, I taught quantum physics repeatedly, and I have tried to present the fundamentals of the theory as a coherent and comprehensive body of phenomena, using, whenever possible, simple mathematical techniques.

The twentieth century was, to some extent, the century of quantum theory. The fundamentals of the present day nanoscience and nanotechnology are closely linked with the electronic and optoelectronic properties of physical systems in the quantum domain. Some systems like the repulsive square barrier and the square-

---

[1] *Innovation in Physics*, F. J. Dyson, *Scientific American*, **199**, 74 (1958).

[2] *The Philosophy Behind Physics*, T. A. Brody; L. de la Peña and P. E. Hodgson, editors, Springer-Verlag, Berlin-New York, 1993.

well potentials, that were simple academic examples in standard quantum mechanics courses, become systems of current interest in theoretical, experimental, and applied physics. Studying these systems, fruitful approaches, models, and techniques evolved. The scattering approach was one of them, and it has been rather successful for studying transport properties. Since then I have gradually introduced this powerful method in my lectures to discuss some simple problems. Energy–eigenvalue equations and eigenfunctions are straightforwardly derived. The ease with which students learn using this intuitive and algebraic method, enticed me and enabled me to include other systems like the double-barrier potentials, double quantum wells, and superlattices.

The general plan of this book is similar to that of most of the textbooks devoted to a first course of non-relativistic quantum theory. In the first two chapters we will briefly summarize the physical problems that show the limits of the classical theories and the new ideas that explained, permanently or temporarily, those problems. Concepts like energy quantization in the blackbody radiation, wave-particle duality, and the Bohr postulates, are thoroughly discussed. We end up with a simple and intuitive derivation of the Schrödinger equation. On the problem of the interpretation of the wave function, we keep a consistent position, but avoid excessive discussion of some controversial issues, which are studied in other books such as *Introduction to Quantum Mechanics* by L. de la Peña. In the second part of this book (Chaps. 3–5), we study the usual examples: free particle; infinite quantum well; rectangular barrier and the *finite* quantum well. We study also more complex systems like the double barrier, the double quantum well, and the *finite* Kronig-Penney model. Fully solving these examples, we face additional properties like the quantum coherence. We introduce the transfer matrix method and let the students acquaint themselves with systems that are present in industrial opto-electronic devices. Chapter 6 is devoted to the semiclassical method of Wentzel, Kramers, and Brillouin (the WKB approximation). This topic is discussed using also the transfer matrix method. New relations are derived and applied for simple examples. From Chap. 7 onwards, we present most of the standard content in a regular quantum book (operators, expected values, angular momentum and spin, matrix representations, the Pauli equation, etc.), we study the well-known harmonic oscillator, the Hydrogen atom, and the fine structure that spurred the crisis of the old quantum theory, between 1920 and 1925. We end up with a summary of the perturbation theory and a chapter on the distinguishable and indistinguishable identical particles. We analyze the close relation between wave-function symmetries, statistics and spin, and discuss properties like the Bose-Einstein condensation and the Pauli exclusion principle.

Together with the formal presentation of the quantum theory, the main objective of this book is to present a coherent set of concepts and physical phenomena. The most emblematic quantum properties like the tunneling effect through potential barriers and the energy quantization in the confining potentials, appear and reappear in more complex examples, modified or transformed. These phenomena, together with the quantum version of the particle current density, the transmission and reflection coefficients, and the splitting of the energy levels that give rise to the energy bands structure, are some of the physical phenomena studied in this text. One of the aims is to show students that quantum theory, more than a set of axioms and differential equations, is a consistent and intelligible theory that clarifies the quantum behavior present in an uncountable number of realizations of the microscopic and macroscopic systems. We pursue the mathematical rigor and conceptual depth together with simple calculations and enhanced comprehension of the quantum concepts.

The content and depth with which we discuss the fundamental issues of the quantum theory is appropriate for a semester course or a trimester plus a complementary course. In each chapter we have some solved problems, which in certain cases complement the discussion of some topics. The second part of this book can also be used as an introductory course to let electric and electronic engineers get acquainted with semiconductor heterostructures, and the quantum properties of these systems.

I thank Emilio Sordo and Luis Noreña for their friendship and for the facilities that the Universidad Autónoma Metropolitana campus Azcapotzalco provides me; to my students of different generations, especially to Fernando Zubieta, Michael Morales, Maria Fernanda Avila, and Victor Ibarra for their support in an important part of the transcription of my notes into the La-TeX language. I thank my students, colleagues, and collaborators Jose Luis Cardoso and Alejandro Kunold. To Jaime Grabinsky, Juergen Reiter, Arturo Robledo and Herbert Simanjuntak, for their friendship and the useful comments and corrections. To Claus Ascheron, Executive Editor of Springer Verlag, for his continued encouragement and interest in the publication of this book. The friendship and hospitality of Dieter Weiss and the University of Regensburg, where the last steps of this English version were taken. The support of CONACyT Mexico and the Program: Apoyo para Año Sabático is also acknowledged. I also thank the support and love of my wife Liuddys.

México Pedro Pereyra

Contents

Units and Useful Constants

Angström 1 Å = 10^{-10} m	Electronvolt 1 eV = 1.60218 10^{-19} J
Micrometer 1 $\mu = 10^{-6}$ m	Nanometer 1 nm = 10^{-9} m
Picosecond 1 ps = 10^{-12} s	Femtosecond 1 fs = 10^{-15} s
GigaHertz 1 GHz = 109 Hz	TeraHertz 1 THz = 10^{12} Hz
Electron charge	$e = -1.6021810^{-19}$ C
Boltzmann constant	$k_B = 1.38066\ 10^{-23}$ J/K
	$= 8.617410^{-5}$ eV/K
Fine structure constant (adimensional)	$\alpha = e^2/\hbar c = 1/137.036$
Planck constant	$h = 6.6261\ 10^{-34}$ Js
	$\hbar = h/2\pi = 1.05457\ 10^{-34}$ Js
	$\hbar = 6.5821\ 10^{-22}$ MeVs
	$2m_e/\hbar^2 = 26.247\ (\mathrm{nm}^2\ \mathrm{eV})^{-1}$
Rydberg constant	$\mathrm{R}_\infty = E_i/hc = 1.09737\ 107\ \mathrm{m}^{-1}$
Electron's Compton wave length	$\lambda_{\mathrm{c}} = h/m_e c = 2.426\ 10^{-12}$ m
Hydrogen ionization energy	$E_i = m_e e^4/2\hbar^2 = \alpha^2 m_e c^2/2$
	$= 13.6057$ eV
Bohr magneton	$\mu_{\mathrm{B}} = q\hbar/2m_e = 9.2740\ 10^{-24}$ J/T
	$= 5.7884\ 10^{-5}$ eV/T
Nuclear magneton	$\mu_{\mathrm{n}} = q\hbar/2m_p = 5.0508\ 10^{-27}$ J/T
	$= 3.1525\ 10^{-8}$ eV/T
Electron mass	$m_e = 9.1093897\ 10^{-31}$ kg
	$m_e c^2 = 0.51099906$ MeV
Proton mass	$m_p = 1.6726231\ 10^{-27}$ kg
	$m_p c^2 = 938.27231$ MeV
Avogadro number	$N_A = 6.0221\ 10^{23}$
Vacuum permeability	$\mu_o = 4\pi\ 10^{-7}\ \mathrm{Hm}^{-1}$
Vacuum permitivity	$\varepsilon_o = 8.854187\ 10^{-12}$ C/Vm
Classical electron radio	$r_e = e^2/m_e c^2 = 2.818\ 10^{-15}$ m
Bohr radio	$a_o = h^2/(m_e e^2) = 0.52918\ 10^{-10}$ m
Light velocity	$c = 299\ 792\ 458$ m/s
	$\hbar c = 197.327$ MeVfm $\simeq$ 1973 eV Å
	$\hbar c = 0.316152\ 10^{-25}$ Jm

Chapter 1
The Origin of Quantum Concepts

At the end of the nineteenth century, just when the classical theories had blossomed into beautiful and elegant formulations, new challenges troubled the scientific community. The spectroscopic methods applied to analyze the atomic and the blackbody radiations, accumulated evidences that could not be explained with the existing theories. The electromagnetic theory, that reached its summit with the Maxwell equations, at the time recognized the ether as the medium of wave propagation, even though the Michelson-Morley experiment denied it. With the discovery of electrons, in 1897, the interest in understanding the atomic structure grew up steadily to become soon a true challenge for experimental and theoretical physicists. These and other problems, underpinned a period of crisis and prolific creativity. Max Planck and Albert Einstein, are emblematic symbols of two new theories of the modern physics that grew out of the crisis: the quantum physics and the relativity theory. Both theories undermined the classical physics and introduced new concepts that not only changed physics but also pervaded and gave shape to the modern culture, dominated by the communications industry and the optoelectronic devices.

To understand the formalism of quantum physics this book deals with, it is instructive to review some major problems that spurred the crisis and the fundamental ideas and concepts that were used to explain them. In this and the following chapter, we will briefly discuss some of these problems. The order of the presentation will not necessarily be in chronological order.

One of the problems that revealed the need of fundamental changes in the classical theories, and the first whose explanation opened the wide world of the quantum physics, was the problem of the *blackbody radiation*. All bodies emit and absorb radiation, and the intensity and frequency distribution of this radiation depends on the body and its temperature. Before Max Planck put forward his theory, the low frequency description of the spectral density was quite satisfactory, but for high frequencies (see Fig. 1.1), it was completely wrong. Max Planck found the correct spectral density assuming oscillators in the walls of the blackbody cavities that could only absorb or emit multiples of discrete amounts of energy.

P. Pereyra, *Fundamentals of Quantum Physics*, Undergraduate Lecture Notes in Physics, DOI: 10.1007/978-3-642-29378-8_1,

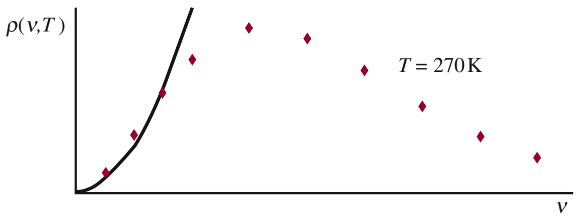


Fig. 1.1 The experimental behavior of the blackbody radiation density, as a function of the frequency ν. The continuous curve, for low frequencies, represents the classical theories description

Another problem that required a new approach, was the *photoelectric effect*, observed when light struck on a metal and is followed by electrons coming out of the metal surface. Prior to the explanation given by Albert Einstein to this problem, it was not clear why the emitted electrons energy depend on the light frequency ν instead of its intensity. Moreover, it was not clear why the electron emission ceased when the light frequency was lower than some critical value ν_c, which depended on the specific metal.

The discovery of the electron by Joseph J. Thompson, in 1897, opened up the atomic structure problem. This problem, and the explanation of the atomic emission lines, remained open for some years. It was clear that, if the electron (the negative electric charge) was part of a neutral atom, one had to admit the existence of positive charges. The problem was, how and where to put, in a stationary configuration, the positive and the negative charges together, and how can one then relate the electronic configuration with the emission lines.

An effect, that appeared some years later, whose explanation was fundamental to the understanding of the quantum phenomenology, was the *Compton effect*. In the Compton effect the light scattered by particles changes its color (i.e. changes its frequency), and the change is a function of the scattering angle, as shown in Fig. 1.4. The explanation of this effect corroborated the quantization concepts as well as some results of the special theory of relativity.

In the following sections we discuss these problems with more detail.

1.1 Blackbody Radiation

All bodies absorb and emit radiation with frequencies that cover the whole spectrum, with intensity distribution that depends on the body itself and its temperature. The problem of the intensity and the color of the radiated light was stated by Kirchhoff in 1859. This problem was also of interest to the electric companies interested in producing light bulbs with maximum efficiency. At the end of the nineteenth century, there were precise measurements of the emitted and absorbed radiation, but without a theoretical explanation. The isolation of the emitting body from other emission sources was an important requirement. One way to achieve this was, for example, to consider a closed box with the inner faces as the emitting surfaces. To observe the radiation inside the box, one makes a small hole. The radiation that comes out from the hole is called the blackbody radiation. Some properties and important results about this radiation were known at the end of the nineteenth century. Among these properties, it was known that the radiation energy density could be obtained from the empirical formula

$$u = \sigma T^4, \quad \text{with } \sigma = 7.56 \times 10^{-16}\,\frac{JK^{-4}}{m^3}. \tag{1.1}$$

Since the energy density can be obtained when the spectral density, $\rho(\omega, T)$, per unit volume and unit frequency is known, it was clear that one had, first, to derive the frequency distribution of the radiation field.

By the end of the nineteenth century, it was usual to assume that the existing physical theories were perfectly competent to explain any experimental result, and could also be used to account for the observed results and to deduce (starting from the implicit first principles) the empirical formulas. It was then reasonable to expect that the radiation density (1.1) and the spectral density $\rho(\omega, T)$, whose behavior was as shown in Fig. 1.1, could be accounted for after an appropriate analysis. However, all theoretical attempts to obtain the spectral density $\rho(\omega, T)$ failed, even though the reasoning lines, as will be seen here, were correct. The failures had a different but subtle cause.

As mentioned before, if the spectral density $\rho(\omega, T)$ would be known, the product

$$du = \rho(\omega, T)d\omega = \rho_\nu(\nu, T)d\nu,$$

integrated over the whole domain of frequencies, should give the energy density sought. In other words, given the spectral density $\rho(\omega, T)$, the energy density would be

$$u = \int_0^\infty \rho(\omega, T)d\omega = \int_0^\infty \rho_\nu(\nu, T)d\nu = \sigma T^4. \tag{1.2}$$

Therefore, the aim was to determine $\rho(\omega, T)$. The theoretical attempts led to a number of basic results and properties. Some of them useful and valid, others not. Let us now mention three of them: the Wien displacement law, as a general condition on the function $\rho(\omega, T)$, and two results that make evident the nature of the problem and the solution offered by Max Planck.

(i) *Wien's displacement law*, refers to the displacement of the maximum of $\rho_\nu(\nu, T)$ with the temperature, i.e. the change of color of the emitted radiation as the body is heated up or cooled down. This phenomenon implies a necessary condition on the spectral density $\rho_\nu(\nu, T)$. To obtain the empirical law (1.1), the spectral density should be a function like

$$\rho_\nu(\nu, T) = \nu^3 f(\nu/T), \tag{1.3}$$

with $f(\nu/T)$ a function to be determined. It is easy to verify that a spectral density like this will certainly produce a result that goes as T^4.

(ii) An important result, that was shown experimentally, was the independence of the blackbody radiation from the specific material of the emitting walls. Taking into account this result, it was clear that one could model the radiating walls as if they were made up of independent harmonic oscillators. If each oscillator has a characteristic oscillation frequency ν, and the characteristic frequencies have a distribution function $\rho_\nu(\nu)$, one will obtain, using the classical physics laws, the average oscillator energy (per degree of freedom[1]).

$$\bar{E} = \frac{\pi^2 c^3}{\omega^2}\rho(\omega) = \frac{c^3}{8\pi\nu^2}\rho_\nu(\nu). \tag{1.4}$$

(iii) On the other hand, from the classical statistical physics it was known that the average energy per degree of freedom in thermal equilibrium at temperature T, is given by

$$\bar{E} = \frac{1}{2}k_B T, \tag{1.5}$$

with k_B the Boltzmann constant.[2]

Combining these results, and taking into account that waves polarize in two perpendicular planes, the classical physics analysis led to the spectral density

$$\rho_\nu(\nu, T) = \frac{8\pi\nu^2}{c^3}k_B T, \tag{1.6}$$

known as the Rayleigh-Jeans formula. This expression fulfills the Wien law when

$$f(\nu/T) = \frac{8\pi k_B}{c^3}\frac{1}{\nu/T}, \tag{1.7}$$

and predicts a spectral distribution that grows quadratically with the frequency ν. As can be seen in Fig. 1.1, the Rayleigh-Jeans formula plotted as a continuous curve, describes well the experimental curve only in the low frequency region.[3] Although the Rayleigh-Jeans distribution is compatible with the Wien displacement law, it diverges in the high frequency region. Thus, the integral (1.2) also diverges. This behavior was known as the "ultraviolet catastrophe".[4]

---

[1] For derivations of the average energy and some other important expressions, see *Lectures on Physics* by Richard P. Feynman, Robert Leighton and Matthew Sands (Addison-Wesley, 1964) and *Introducción a la mecánica cuántica* by L. de la Peña (Fondo de Cultura Económica and UNAM, México, 1991).

[2] $k_B = 1.38065810^{-23}$J/K^{-1}.

[3] J. W. S. Rayleigh, *Philosophical Magazine*, series 5, 49 (301): 539 (1900); J. H. Jeans, *Phil. Trans. R. Soc. A*, **196** 274: 397 (1901).

[4] Apparently this term was coined by Paul Ehrenfest, some years later.

In the twilight of the nineteenth century, on December 14th of 1900, Max Planck presented to the German Society of Physics and published in the *Annalen der Physik*,[5] the first unconventional solution to this problem. To explain the blackbody spectrum, Planck introduced the formal assumption that the electromagnetic energy could be absorbed and emitted only in a quantized form. He found that he could account for the observed spectra, and could prevent the divergence of the energy density, if the oscillators in the walls of the radiating and absorbing cavity, oscillating with a frequency ν, lose or gain energy in multiples of a characteristic energy E_ν, called a quantum of energy.

If the energy that is absorbed or emitted by an oscillator with frequency ν, is

$$E = nE_\nu \quad n = 1, 2, 3, \ldots, \tag{1.8}$$

a multiple of a characteristic energy E_ν, the probability of finding the oscillator in a state of energy E, which in the continuous energy description is given by the Boltzmann distribution function

$$p(E, T) = \frac{e^{-E/k_BT}}{\int e^{-E/k_BT} dE}, \tag{1.9}$$

will change to

$$p(E_\nu, T) = \frac{e^{-nE_\nu/k_BT}}{\sum_n e^{-nE_\nu/k_BT}}. \tag{1.10}$$

With energy discretization the integrals have to be replaced by sums. This seemingly inconsequential change, was at the end a fundamental one. With a discretized distribution function the average energy becomes

$$\bar{E} = \frac{\sum_n nE_\nu e^{-nE_\nu/k_BT}}{\sum_n e^{-nE_\nu/k_BT}}. \tag{1.11}$$

It is not difficult to see that the sums in the numerator and denominator can easily be evaluated. Indeed if we remember that

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots = \sum_{n=0} x^n, \tag{1.12}$$

for absolute value of x less than 1, and also that

$$\frac{d}{dx}\sum_{n=0} x^n = \frac{1}{x}\sum_{n=0} nx^n = \frac{1}{(1-x)^2}, \tag{1.13}$$

we can show that the energy average is given by

---

[5] M. Planck, *Ann. Phys.*, **4**, 553 (1901).

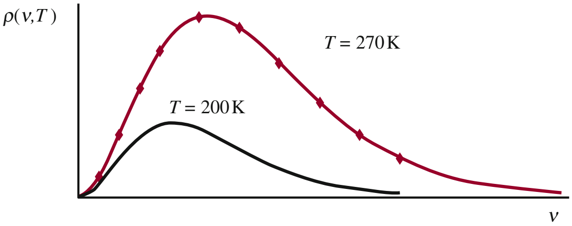


Fig. 1.2 The experimental spectral distribution and the Planck distribution for different temperatures. The description is perfect from low to high frequencies

$$\bar{E} = \frac{E_\nu}{e^{E_\nu/k_B T} - 1}. \tag{1.14}$$

Using $e^x \simeq 1 + x$ for small x, this expression reduces to $k_B T$ in the high temperatures limit. Combining with the average energy of equation (1.4), we have

$$\rho_\nu(\nu, T) = \frac{8\pi\nu^2}{c^3} \frac{E_\nu}{e^{E_\nu/k_B T} - 1}. \tag{1.15}$$

For this distribution to satisfy the Wien displacement law, the characteristic energy E_ν, the minimum energy, absorbed or emitted, must be proportional to the frequency ν. Writing the characteristic energy as

$$E_\nu = h\nu, \tag{1.16}$$

with h the famous *Planck constant*,[6] Max Planck found the spectral density

$$\rho_\nu(\nu, T) = \frac{8\pi\nu^2}{c^3} \frac{h\nu}{e^{h\nu/k_B T} - 1}, \tag{1.17}$$

that is known as *the Planck spectral density*. This density describes the experimental results all the way from the low to the high frequencies (see Fig. 1.2). It is easy to verify that at high temperatures, for which $k_B T \gg h\nu$, the Planck spectral density reduces to the spectral density of Rayleigh and Jeans.

An important test for the Planck spectral density is the calculation of the energy density of the radiation field. Using Planck's spectral density, we have

$$u = \frac{8\pi h}{c^3} \int_0^\infty \frac{\nu^3 d\nu}{e^{h\nu/k_B T} - 1} = \frac{8\pi k_B^4}{c^3 h^3} T^4 \int_0^\infty \frac{x^3 dx}{e^x - 1}, \tag{1.18}$$

where $x = h\nu/k_B T$. As the integral on the right-hand side of (1.18) is a finite number (in fact equal to $\pi^4/15$), this energy has not only the correct temperature dependence, it can be used, based on the empirical energy density of (1.1), to obtain the Planck constant

---

[6] This is one of the fundamental constants in physics and of the laws of nature.

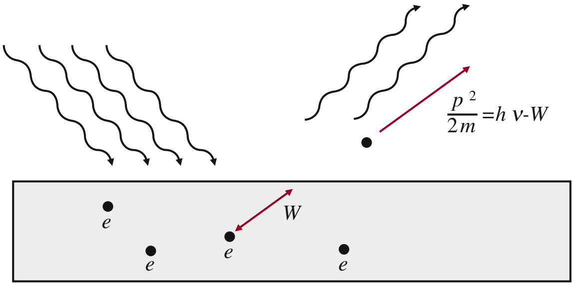


Fig. 1.3 The photoelectric effect. When an electromagnetic radiation, with energy $h\nu$, strikes a metal surface, the electrons absorb the whole energy and pay part for releasing the electrons from the metal (known as the work function W) to become free particles. The remaining energy $h\nu - W$, if some is left, transforms into the electron's kinetic energy $p^2/2m$

$$h = 6.6260755 \times 10^{-34}\text{Js} = 4.1356692 \times 10^{-15}\text{eVs}.$$

It is common to express $h\nu$ as $\hbar\omega$ with $\hbar = h/2\pi = 1.054572 \times 10^{-34}$Js and $\omega = 2\pi\nu$. Note that the units of h are energy×time. It is not difficult to show that the very small magnitude of h hides the quantum phenomena. Indeed, if the frequencies of oscillations were, say of the order of 10 Hz, the absorbed or emitted energies would be, as mentioned earlier, multiples of $h\nu \approx 10^{-32}$ J; a very small amount of energy. We can then ask if changes of this magnitude can be observed or not in the macroscopic physical systems. To answer this question let us suppose that we have a classical oscillator, which is a particle of mass $m = 1$g attached to a spring of constant $k = 10$ N/m. If the particle's oscillations amplitude is, say $x_o = 1$ cm, its energy and oscillations frequency would be, respectively, of the order of $E \simeq 5 \times 10^{-4}$J and $\nu \simeq 10$ Hz. Furthermore, if the precision measuring the energy is, say $\Delta E \simeq 10^{-6}E$, the number of *quanta* of energy contained in ΔE would be $n = \Delta E/h\nu$, of the order of 10^{22}; a very large number! Hence, the contribution of one quantum of energy is rather negligible. We will see later that the number of quanta will be considerably less, of the order of one, when the energies and particles are of atomic dimensions (Figs. 1.3, 1.4).

1.2 The Photoelectric Effect

In 1887, Heinrich Hertz noticed that when a metal surface was illuminated with UV light, electrons were ejected from the surface - provided the light's frequency was above a certain, metal-dependent, threshold ν_c. The main properties related with these phenomena are:

(i) *the speed of the ejected electrons depends only on the frequency of the incident light*;

(ii) *the number of emitted electrons depends on the intensity of the incident radiation*;
(iii) *for each metal there is a frequency ν_c, called the critical frequency, below which no photoelectric phenomenon is observed.*

In 1902 Philipp Lenard noticed that "the usual conception that the light energy is continuously distributed over the space through which it spreads, encounters serious difficulties when trying to explain the photoelectric phenomenon'.[7] In 1905 applying Planck's idea of quantization of the absorbed and emitted radiation energy, Albert Einstein was able to explain the photoelectric effect by assuming a "corpuscular" nature for the quanta of light, i.e. a dual nature where a wave and a particle property coexist, and are part of the fundamental characteristics of the same object.

If the quantum of radiation, the light corpuscle, has an energy $h\nu$, and this energy is transmitted to an electron in the corpuscle-electron interaction, part of the energy is used to expel the electron from the metal [8] and the remaining is transformed into its kinetic energy. This means that:

$$\frac{1}{2}m_e v^2 = h\nu - W. \tag{1.19}$$

When the frequency ν of the incident radiation is such that $h\nu$ coincides with W, the kinetic energy is zero, but when $h\nu < W$ the electron will not be able to get out of the metal. Thus $h\nu = W$ is a particular condition that defines the critical frequency $\nu_c = W/h$. Each metal has its own critical frequency. When $\nu \geq \nu_c$, each photon, absorbed in the electron-photon interaction, makes possible the release of one free electron, the number of free electrons depends then on the number of the photons, i.e. on the radiation intensity. In this way, all three characteristics of the photoelectric effect got a rather simple explanation.

Furthermore, according to the special theory of relativity, a massless particle, like the radiation field corpuscle,[9] has the linear momentum

$$p = \frac{E_\nu}{c} = \frac{h\nu}{c}. \tag{1.20}$$

Since the wave velocity c is equal to the product $\nu\lambda$, where λ is the wave length, we can write this relation as

$$p = \frac{h}{\lambda} = \hbar k, \tag{1.21}$$

where $k = 2\pi/\lambda = w/c$ is the wavenumber. Since both, p and k, are vector quantities we must write in general as

---

[7] In A. Einstein, *Ann. Phys.* **17**,132 (1905).

[8] This is known as the work function W, and it is related to the electron's binding energy.

[9] In the especial theory of relativity we have the relation $E^2 = p^2c^2 + (m_oc^2)^2$ between energy E, momentum p and the rest energy m_oc^2. It is clear that for $m_o = 0$, we are left with $E = pc$. If m is the mass of the particle when it is moving and $E = mc^2$, the rest mass m_o and the moving particle mass m are related by $m^2(1 - v^2/c^2) = m_o^2$.

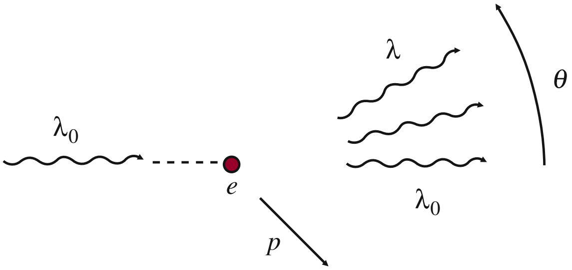


Fig. 1.4 The Compton effect. The change of the scattered radiation wave length $\Delta\lambda = \lambda - \lambda_0$ (or change of frequency $\Delta\omega = \omega - \omega_0$), depends on the scattering angle θ. This effect can be explained when the radiation-electron interaction is modeled as a collision of particles

$$\mathbf{p} = \hbar\mathbf{k}. \tag{1.22}$$

This is a very important and recurrent relation in quantum theory, and will be useful to explain the Compton effect, that we will discuss now, and to derive the Schrödinger equation later.

1.3 The Compton Effect

In 1923, A. H. Compton noticed that the wavelength of X-rays scattered by electrons in graphite increases as a function of the scattering angle θ. It turns out that these phenomena can be understood by making use of the dual particle-wave nature for the quantum particles involved in this process, together with the basic relations of the special theory of relativity, that was already well established in those years. Indeed, when the electron-photon interaction is analyzed as a collision of two particles, the conservation laws of energy and momentum lead to

$$\hbar\omega_0 + E_0 = \hbar\omega + E, \tag{1.23}$$

and

$$\hbar\mathbf{k}_0 + \mathbf{p}_0 = \hbar\mathbf{k} + \mathbf{p}. \tag{1.24}$$

Here $E_0 = m_0c^2$, $E = mc^2 = m_0c^2/\sqrt{1 - v^2/c^2}$, $\mathbf{p}_0 = 0$ and $\mathbf{p} = m\mathbf{v}$ are the electron energies and momenta, before and after the collision, respectively. Taking the squares of these equations, written as

$$m^2c^4 = m_0^2c^4 + 2m_0c^2\hbar(\omega_0 - \omega) + \hbar^2(\omega_0 - \omega)^2 \tag{1.25}$$

and

$$\hbar^2(k^2 + k_0^2 - 2kk_0\cos\theta) = m^2v^2, \tag{1.26}$$

and using the relations $k = \omega/c$, $k_0 = \omega_0/c$ and $m^2c^4 = m_0^2c^4 + m^2v^2c^2$, we can easily obtain the following equation

$$\omega\omega_0(1 - \cos\theta) = \frac{m_0c^2}{\hbar}(\omega_0 - \omega), \tag{1.27}$$

that can be written also as

$$\lambda - \lambda_0 = \frac{h}{m_0c}(1 - \cos\theta) = 2\lambda_c \sin^2\frac{\theta}{2}. \tag{1.28}$$

In the last equation the Compton wavelength $\lambda_c = h/m_0c$ was introduced. A definition that ascribes an undulatory property, the wavelength, to particles with mass. A quantity that tends to zero as the mass increases. For electrons $\lambda_c \approx 2.4 \times 10^{-3}$nm. It is evident from (1.28) that $\lambda \geq \lambda_0$. It shows also that the difference $\Delta\lambda = \lambda - \lambda_0$ reaches its maximum value when the dispersion is backwards ("backscattering"), i.e. when $\theta = \pi$. The relative change of the wavelength $\Delta\lambda/\lambda_0$, is of the order of λ_c/λ_0. This ratio allows one to establish whether the Compton effect will be observed or not. If the incident radiation is a visible light, λ_0 from 400 to 750 nm, the relative change will be of the order of 10^{-5}. In contrast, if the incident radiation has a wavelength shorter than those of the visible light, the relative change will be greater, and the Compton effect might be seen. For instance, for X-ray with $\lambda_0 \approx 10^{-2}$nm the relative change is of the order of 10^{-1}, which means 10% of the incident wavelength. This effect can easily be observed and was observed indeed!

1.4 Rutherford's Atom and Bohr's Postulates

Searching for the atomic structure, Rutherford studied the dispersion of α particles by thin films of gold. To explain the high amount of α particles at high dispersion angles, Rutherford suggested that atoms have a charge $+Ne$ at the center, surrounded by N electrons which, following the "saturnian" atom hypothesis proposed by Hantaro Nagaoka, rotate in saturnian rings of radii R. With this model of atoms, Rutherford deduced the angular distribution of the scattered particles. The results agreed substantially with those obtained by Geiger in 1910. Although these experiments and the theory did not reveal the sign of the charge, Rutherford assumed always that the positive charge was at the center.

According to the classical theory, an electron that rotates in a circular orbit around a positive charge Ze is subject to a "centrifugal" force and to an attractive Coulomb force. When the magnitudes of these forces are equal, it is possible to express the electron energy as

$$E = -\frac{1}{2}m\omega^2\left(\frac{Ze^2}{2E}\right)^2, \tag{1.29}$$

and the angular frequency by

$$\omega = \frac{2}{e^2 Z}\sqrt{\frac{2\,|E|^3}{m}}. \tag{1.30}$$

Here m and e are the mass and charge of the electron, while E is its energy. If the absolute electron energy E takes any real value, the frequency ω will take also any real value. But the experimental results showed that the emitted frequencies were discrete.[10] In 1885, J. J. Balmer found that the wavelengths of the four visible lines of the Hydrogen spectrum can be obtained from

$$\frac{1}{\lambda} = \kappa\left(\frac{1}{4} - \frac{1}{n^2}\right), \quad \text{with} \quad n = 3, 4, 5, 6, \tag{1.31}$$

and $\kappa = 17465\text{cm}^{-1}$. Shortly after J. R. Rydberg showed that all known series can be obtained from

$$k = \frac{2\pi}{\lambda} = R\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right), \qquad \text{with} \qquad n_1 < n_2, \tag{1.32}$$

and $R = 109735.83\text{cm}^{-1}$. This constant is known as Rydberg's constant. Some time later, Niels Bohr, using Rutherford's model and the fundamental ideas introduced by Planck and Einstein, on the quantization of energy, proposed the following postulates on the stationary states of atoms and on the emitted and absorbed radiation frequencies:

I. *an atomic system can only exist in a number of discrete states;*
II. *the absorbed or emitted radiation during a transition between two stationary states has a frequency* ν *given by* $h\nu = E_i - E_f$.

With these postulates he showed that it is possible to explain the separation and regularity of the spectral lines, and the Rydberg formula for the Hydrogen spectrum. Indeed, if one assumes that the energy is quantized as[11]

$$E_n = nh\nu/2, \tag{1.33}$$

one can easily obtain the energy

$$E_n = -\frac{2\pi^2 m e^4}{h^2 n^2}, \quad \text{with} \quad n = 1, 2, 3, \ldots \tag{1.34}$$

---

[10] At that time it was common to assume that the emitted radiation frequency was related with the electron's oscillation frequency.

[11] N. Bohr, *Philosophical Magazine*, ser. 6 vol. 26, 1 (1913). Notice the factor 1/2.

Given this energy it is possible to evaluate the difference $E_{n_2} - E_{n_1}$ (associated with the transition $E_{n_1} \to E_{n_2}$) and to obtain Rydberg's formula.

$$E_{n_2} - E_{n_1} = \frac{2\pi^2 me^4}{h^2}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right), \quad \text{with} \quad n = 1, 2, 3, \ldots \tag{1.35}$$

In the early years of the twentieth century, the electromagnetic theory was considered one of the most firmly established theories. It was known that accelerated charges radiate energy (Bremsstrahlung). In Rutherford's model the orbiting electrons are accelerated charges, therefore they must lose energy and eventually collapse into the nucleus. Nonetheless, that did not seem to occur. Given the coincidence with the Rydberg formula, and the difficulty to explain these basic contradictions, Bohr's postulates were accepted, for some years, as factual statements that reflect the behavior of nature at the microscopic level:

$$h\nu = E_{n_2} - E_{n_1} = \frac{2\pi^2 me^4}{h^2}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right). \tag{1.36}$$

Bohr's model accounts for the observed results but does not explain why an accelerated electron remains in a stable orbit, neither the emission mechanism, nor the laws that determine the transition probabilities. Nevertheless, these postulates and the correspondence principle with the classical description in the limit of large quantum numbers n (also proposed by Bohr), were held for many years, and constituted what later was called the old quantum theory. Meanwhile, there were many attempts to explain them on a firmer basis, as well as to give them experimental support or to question their general validity. In these attempts two fundamental schools became pre-eminent in the future development of the quantum theory: the school of Arnold Sommerfeld in Munich and the work of Max Born and collaborators in Göttingen. It is beyond the purpose of this book to analyze in detail the work of Sommerfeld, Born, Van Vleck, Heisenberg, Jordan, Pauli, Dirac etc.. We just recall and recognize that the cumulative work of all of them reached, in the joint work of Born, Heisenberg and Jordan, a zenithal point with the matrix version of the quantum theory, the matrix mechanics. A few months later, following a different line of thought, closer to the particle diffraction experiments and to the wave-particle duality proposed by Louis de Broglie, Erwin Schrödinger introduced the wave version of the quantum theory, the quantum wave mechanics. In the following chapter we discuss with more detail this alternative approach to quantum theory. To conclude this chapter we will summarize the ideas of Einstein published[12] in 1917, with the title "About the quantum theory of radiation" where, among other results, Einstein deduced the Planck distribution and the second postulate of Niels Bohr.

---

[12] A. Einstein, *Mitteilungen der Physikalischen Gesellschaft Zürich* **18** , 47 (1916) and Phys. Zs. **18**, 121 (1917).

1.5 On Einstein's Radiation Theory

As an extension of the ideas that were used to explain the photoelectric effect, where the quantization concept was not restricted to the emission and absorption mechanism but taken also as a characteristic property of the radiation field, Einstein derived the Planck distribution and the Bohr postulate, assuming that the emitting and absorbing molecules in the walls, in thermal equilibrium with the blackbody radiation, are themselves allowed to exist only in a discrete set of states.

1.5.1 Planck's Distribution and the Second Postulate of Bohr

Einstein assumed that if a molecule exists only in a discrete set of states with energies E_1, E_2, ... the relative frequency of finding the molecule in the state n, in analogy with the Boltzmann-Gibbs distribution, should be given by

$$f_n = c_n e^{-E_n/k_B T}, \tag{1.37}$$

where c_n is a normalization constant. A molecule in the state of energy E_n, in thermal equilibrium with an electromagnetic field characterized by a spectral density ρ, absorbs or emits energy and changes to the state of energy E_m, with $E_m < E_n$ in the absorption process and $E_m > E_n$ in the emission process. The probabilities for these processes to occur during a time interval dt are

$$dW_n^m = B_n^m \rho dt \quad \text{and} \quad dW_m^n = B_m^n \rho dt. \tag{1.38}$$

Here, the coefficient B_n^m represents the transition probability per unit of time, from the state with energy E_n to the state with energy E_m. These are transitions induced by the molecule-field interaction. Since the state n occurs with the frequency f_n given in (1.37), the number of transitions per unit time, from n to m, can be written as

$$c_n e^{-E_n/k_b T} B_n^m \rho dt. \tag{1.39}$$

For the excited molecules, Einstein envisaged also the possibility of making transitions to lower states by spontaneous emission, independent of the field. Therefore the probability that a molecule emits during a time dt is

$$dW = (A_m^n + B_m^n \rho) dt, \tag{1.40}$$

with A_m^n the probability of spontaneous emission per unit of time. To preserve the equilibrium of these processes, one needs the balance condition

$$e^{-E_n/k_b T} B_n^m \rho = e^{-E_m/k_b T} (A_m^n + B_m^n \rho). \tag{1.41}$$

It is easy to verify that this condition, with $B_n^m = B_m^n$, leads to the spectral density

$$\rho = \frac{A_m^n}{B_m^n} \frac{1}{e^{\left(E_m - E_n\right)/k_b T} - 1}. \tag{1.42}$$

From Wien's displacement law, it follows that

$$\frac{A_m^n}{B_m^n} = \alpha \nu_{nm}^3 \quad \text{and} \quad E_m - E_n = h\nu_{nm}, \tag{1.43}$$

with α and h constants to be determined, for example, by comparing with the Rydberg and the Rayleigh-Jeans formulas at high temperatures. It is worth noticing that Einstein deduced, at the same time, Planck's distribution and the second postulate of Niels Bohr. We shall now briefly refer to Einstein's specific heat model.

1.5.2 Einstein's Specific Heat Model

In an exercise of congruence, Einstein suggested in 1907 that the quantization hypothesis should explain also other physical problems where the classical description was in contradiction with the experimental observations. One of these was the specific heat of solids that, according to classical theory, must be constant, but experimentally tends to zero as the temperature goes to zero. If atoms in a solid are represented by oscillators, the average energy of one atom, per degree of freedom, will be given by

$$\bar{E} = \frac{h\nu}{e^{h\nu/k_B T} - 1}, \tag{1.44}$$

with $\nu = \omega/2\pi$ the average frequency of oscillations. If all atoms, in the Einstein model vibrate with the average frequency ν, the internal energy of a solid containing N atoms, with 3 degrees of freedom each, is given by

$$U = \frac{3N\hbar\omega}{e^{\hbar\omega/k_B T} - 1}. \tag{1.45}$$

When the temperatures are high, the internal energy of the solid takes the form

$$U = 3Nk_B T, \tag{1.46}$$

and the specific heat will be

$$C_V = \left(\frac{\partial U}{\partial T}\right)_V = 3Nk_B. \tag{1.47}$$

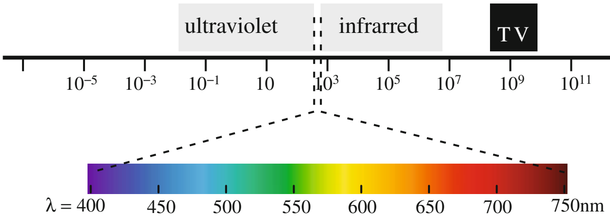


Fig. 1.5 The visible electromagnetic spectrum

This expression coincides with the classical results. If temperatures are low, we have

$$C_V = 3Nk_B \left(\frac{\hbar\omega}{kT}\right)^2 e^{-\hbar\omega/kT}, \tag{1.48}$$

an expression that tends to zero exponentially when $T \to 0$. This function agrees qualitatively well with the experimental results at low temperatures, a result that was well known at the beginning of the 20th century. A more quantitative approach and in actual agreement with C_V that really behaves as T^3 was derived by Debye.[13]

1.6 Solved Problems

Exercise 1 A quantum of electromagnetic radiation has an energy of 1.77 eV. What is the associated wavelength? To what color does this radiation correspond? (Fig. 1.5)

Solution We will use Planck's relation $E = h\nu = hc/\lambda$. The constant hc in this and other problems is

$$\begin{aligned} hc &= 6.63\ 10^{-34}\,\text{Js}\ 3.0\ 10^{8}\frac{m}{s} \\ &= 19.89\ 10^{-26}\ \text{Jm}\ \frac{1\,\text{eV}}{1.6\,10^{-19}\text{J}}\,\frac{1\text{nm}}{10^{-9}\text{m}} \\ &= 1243\ \text{eV nm}. \end{aligned} \tag{1.49}$$

With this result we get

$$\lambda = \frac{hc}{E} = \frac{1243}{1.77}\ \text{nm} = 702.25\ \text{nm}. \tag{1.50}$$

---

[13] By adopting the idea of Einstein and limiting the frequency ω to a maximum frequency ω_D.

This wavelength corresponds to red light, which includes wavelengths between 640 and 750 nm, approximately.

Exercise 2 A mass of 2kg attached to a spring, whose constant is $k = 200$ N/m, is moving harmonically with an amplitude of 10 cm on a smooth surface without friction. If we assume that its energy can be written as a multiple of the quantum of energy $\hbar\omega$ i.e. as $E = n\hbar\omega$, determine the number of quanta n. If the energy is determined with an error ΔE of one part in a million, i.e. $\Delta = \pm E/1000000$, how many quanta of energy are contained in ΔE?

Solution We will calculate first the energy of the system. This is a classical system and its energy at the maximum elongation point (when the kinetic energy is zero) is

$$E = \frac{1}{2}kA^2 = \frac{1}{2}200\ \frac{\mathrm{N}}{\mathrm{m}}(0.1)^2\,[\mathrm{m}^2] = 1\,\mathrm{J}. \tag{1.51}$$

The error will be then

$$\Delta E = \pm\frac{1}{1000000}\,E = \pm 0.000001\ E. \tag{1.52}$$

To use the energy $nh\nu$ and get the number n, we need the frequency

$$\begin{aligned}\nu &= \frac{1}{2\pi}\sqrt{\frac{k}{m}} = \frac{1}{2\pi}\sqrt{\frac{200}{2}}\\ &= \frac{10}{2\pi}\,\mathrm{Hz} = 1.5915\,\mathrm{Hz}.\end{aligned} \tag{1.53}$$

Therefore

$$n = \frac{E}{h\nu} = \frac{1\,\mathrm{J}}{6.63\ 10^{-34}\,\mathrm{J\,s}\ 1.5915\,[\mathrm{Hz}]} = 9.4769\ 10^{32}. \tag{1.54}$$

This is a very big number. The number of quanta of energy to which the error ΔE corresponds is

$$\Delta n = \frac{\Delta E}{h\nu} = \pm\frac{0.000001\,\mathrm{J}}{6.63\ 10^{-34}\,\mathrm{J\,s}\ 1.5915\,\mathrm{Hz}} = \pm 9.4769\ 10^{26}. \tag{1.55}$$

Exercise 3 Let us assume that X-rays are produced in a collision of electrons with a target. If electrons deliver all their energy in this process, what minimum-acceleration voltage is required to produce X-rays with a wavelength of 0.05nm?

Solution When an electron is accelerated by a potential difference ΔV, it acquires a potential energy $e\Delta V$. If this energy is transformed into the energy $h\nu$ of a photon of wavelength $\lambda = c/\nu$, we have

$$e\Delta V = hc/\lambda. \tag{1.56}$$

Hence

$$\Delta V = \frac{\hbar c}{e\lambda} = \frac{1243 \text{ eV nm}}{1.6 \ 10^{-19}\text{C}\, 0.05\,\text{nm}} \frac{1.6 \ 10^{-19}\text{J}}{1\,\text{eV}}$$
$$= 24860\frac{\text{J}}{\text{C}}. \tag{1.57}$$

Thus the required accelerating potential is $\Delta V = 24860$V.

1.7 Problems

1. Find the average energy in equation (1.14).
2. Show and discuss Wien's displacement law.
3. Show that Planck's spectral density can also be written as

$$\rho(\omega, T) = \frac{\omega^2}{\pi^2 c^3} \frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1}. \tag{1.58}$$

4. Consider the equation (1.14) and show that to fulfil the Wien displacement law, the minimum absorbed or emitted energy E_ν must be proportional to the frequency ν.
5. Show that at high temperatures, the Planck distribution reduces to that of Rayleigh and Jeans.
6. Plot Planck's spectral density as a function of the frequency ν for three different temperatures: T = 50K, 200K and 300K. Determine the frequency where the spectral density reaches its maximum value. Determine in which direction the maximum of the spectral density moves as the temperature increases.
7. Prove that Planck's constant can be expressed as

$$\hbar = \frac{k_B}{c} \left(\frac{\pi^2 k_B}{15a} \right)^{1/3}, \tag{1.59}$$

and obtain its numerical value in Js and eVs.
8. If the escape energy of electrons from a metal surface is 1.0 eV when the metal is irradiated with green light, what is the work function for that metal?
9. If a potassium photocathode is irradiated with photons of wavelength $\lambda = 253.7$ nm (corresponding to the resonant line of mercury), the maximum energy of the emitted electrons is 3.14 eV. If a visible radiation with $\lambda = 589$ nm (resonance line of sodium) is used, the maximum energy of emitted electrons is 0.36 eV.

    a. Calculate the Planck constant.
    b. Calculate the work function for the extraction of electrons in potassium.

  c. What is the maximum radiation wavelength to produce the photoelectric effect in potassium?

10. a) An antenna radiates with a frequency of 1 MHz and an output power of 1 kW. How many photons are emitted per second? b) A first-magnitude star emits a light flux of $\sim 1.6\,10^{-10}$ W/m^2, measured at the earth surface, with an average wavelength of 556 nm. How many photons per second pass through the pupil of an eye?
11. Derive equations (1.27) and (1.28).
12. A 10MeV photon hits an electron at rest. Determine the maximum loss of photon energy. Would this loss change if the collision is with a proton at rest?
13. Determine the relative change in wavelength when the incident radiation in a Compton experiment is of X-rays (with wavelength $\lambda_0 \sim 10^{-9}$cm) on graphite.
14. Determine the wavelengths of the Balmer series for Hydrogen.
15. What is the energy difference $E_m - E_n$ if the emitted light is: a) yellow b) red and c) blue?

Chapter 2
Diffraction, Duality and the Schrödinger Equation

The *ad hoc* postulates introduced to explain the Hydrogen emission lines, assuming that the atomic system exists only in a set of energy states, were taken with reticence, and the resistance grew when new experimental results, for atoms in magnetic fields, could not be explained based on those postulates. Physicists of the stature of Sommerfeld, Kramers, Heisenberg and Born were involved in different attempts to formalize the quantization phenomenon. Throughout the 1910s and still in the 1920s, many problems were approached using the old quantum theory. The rotational and vibrational spectra of molecules were studied and the electron spin was envisioned. Arnold Sommerfeld, using his semiclassical quantization rules, that will be considered below, studied the relativistic Hydrogen atom and introduced the fine structure constant.[1]

At the end of the last chapter, we mentioned the matrix formulation of the quantum theory by Heisenberg, Born and Jordan. In this theory, closely related to Bohr's model, the physical quantities are represented by a collection of Fourier coefficients like

$$X_{n_2,n_1}(t) \equiv e^{i2\pi(E_{n_2}-E_{n_1})t/h} X_{n_2,n_1}(0), \tag{2.1}$$

with two indices corresponding to the initial and final states. This phenomenological theory was the first to appear;[2] another, likewise phenomenological but more intuitive, led to the wave mechanics formulation by Erwin Schrödinger. In this book, we will study and we will solve a number of examples using the Schrödinger equation. Important precedents for the Schrödinger theory were the particle-wave duality, ex-

---

[1] Arnold Sommerfeld introduced the fine-structure constant in 1916, in his relativistic deviations of the atomic spectral lines in the Bohr model. The first physical interpretation of the fine-structure constant, $\alpha = e^2/(4\pi\epsilon_o\hbar c) = 7.297352569810^{-3}$, is the ratio of the velocity of the electron in the first circular orbit of the relativistic Bohr atom to the speed of light in vacuum. It appears naturally in Sommerfeld's analysis, and determines the size of the splitting or fine-structure of the hydrogenic spectral lines. The same constant appears in other fields of modern physics.

[2] M. Born, W. Heisenberg, and P. Jordan, *Zeitschrift für Physik*, **35**, 557(1925] (received November 16, 1925). [English translation in: B. L. van der Waerden, editor, *Sources of Quantum Mechanics* (Dover Publications, 1968) ISBN 0-486-61881-1].

P. Pereyra, *Fundamentals of Quantum Physics*, Undergraduate Lecture Notes in Physics,
DOI: 10.1007/978-3-642-29378-8_2, © Springer-Verlag Berlin Heidelberg 2012

tended for massive particles by Louis de Broglie. We will review now these topics, the Sommerfeld-Wilson-Ishiwara's quantization rules, and we will end up with a simple derivation of the Schrödinger equation based on the wave equation and the fundamental duality relation $\lambda = \hbar/p$ for the particle momentum and the wavelength.

2.1 Sommerfeld-Wilson-Ishiwara's Quantization Rule

Arnold Sommerfeld, William Wilson and Jun Ishiwara, independently and almost simultaneously,[3] showed that, to study the behavior of quantum systems with periodic movements, as electrons in the Rutherford model, it is convenient to introduce the action variables, defined as

$$J_i = \oint p_i dq_i, \tag{2.2}$$

where p_i are the momenta and q_i the corresponding coordinates, and to postulate the action quantization in the form

$$J_i = n_i h = 2\pi\hbar n_i, \tag{2.3}$$

with n_i an integer. The integral $\oint dq_i$ means one period of motion. Since the action has the same units as the Planck constant, it is often called the quantum of action.

2.1.1 The Quantization Rules and the Energies of the Hydrogen Atom

We will see now how the quantization rule (2.3) applied to the Hydrogen atom action variable J_φ leads to the quantized Hydrogen-atom energies.[4] We start with the Hydrogen atom Hamiltonian written as

$$H = \frac{p_r^2}{2m} + \frac{p_\varphi^2}{2mr^2} - \frac{e^2}{r}. \tag{2.4}$$

Here the variable φ is a *cyclic variable*,[5] thus its canonical conjugate, the angular momentum p_φ, is a constant of the motion. Indeed

$$\dot{p}_\varphi = -\frac{\partial H}{\partial \varphi} = 0. \tag{2.5}$$

---

[3] Einstein also proposed similar rules of quantization.

[4] We assume here that the Hydrogen nucleus is at rest.

[5] A cyclic or ignorable variable does not appear explicitly in the Hamiltonian.

Using the canonic Hamilton equation we have also

$$\dot{\varphi} = \frac{\partial H}{\partial p_\varphi} = \frac{p_\varphi}{mr^2}. \tag{2.6}$$

This leads us to express the angular momentum p_φ as

$$p_\varphi = mr^2 \dot{\varphi}. \tag{2.7}$$

Since the angular momentum p_φ is a constant of the motion, we obtain

$$J_\varphi = \oint p_\varphi d\varphi = 2\pi p_\varphi, \tag{2.8}$$

and applying the Sommerfeld-Wilson-Ishiwara's quantization rule, we have

$$p_\varphi = mr^2 \dot{\varphi} = n\hbar. \tag{2.9}$$

This shows that *the angular momentum* of the Hydrogen-atom electron is quantized. To obtain the energy, we also need the momentum p_r. For circular electron orbits, r is constant. Thus, we must have

$$\dot{r} = \frac{\partial H}{\partial p_r} = \frac{p_r}{m} = 0. \tag{2.10}$$

In a circular motion the radial component of the linear momentum is, of course, zero, i.e. $p_r = 0$. From the canonical Hamilton equation we have

$$\dot{p}_r = -\frac{\partial H}{\partial r} = \frac{p_\varphi^2}{mr^3} - \frac{e^2}{r^2}, \tag{2.11}$$

and we conclude that

$$\frac{p_\varphi^2}{mr^3} - \frac{e^2}{r^2} = 0. \tag{2.12}$$

Substituting this equation, as well as (2.7) and (2.9) for the corresponding variables in the Hamiltonian function of equation (2.4), we have

$$H \rightarrow E_n = -\frac{2\pi^2 me^4}{h^2 n^2}. \tag{2.13}$$

This energy coincides with that in (1.34). Sommerfeld and his collaborators, among whom was Werner Heisenberg, applied these postulates to systems whose orbits are not necessarily circular, with interesting results related to the spectral lines intensities and the hyperfine structures.

2.1.2 The Rotator

A simple system where the quantization rule (2.3) can be used to obtain the quantization of the angular momentum is the rotator. A mass M attached to the end of a massless rigid rod of length R. In the space of dimension two, this system is described by the Lagrangian

$$L = \frac{MR^2}{2}\dot{\theta}^2, \tag{2.14}$$

and the momentum p_θ, conjugate to the polar angle θ, is a constant of the motion given by

$$p_\theta = MR^2\dot{\theta}. \tag{2.15}$$

The quantization of the polar angle θ action

$$J_\theta = \oint p_\theta d\theta = 2\pi p_\theta, \tag{2.16}$$

leads to

$$p_\theta = MR^2\dot{\theta} = n\hbar. \tag{2.17}$$

In the Bohr model, this restriction imposed on the circular orbits was enough to determine the energy levels.

In the space of dimension three, a rigid rotator can be described by two angles: the inclination relative to a z-axis, θ, and the rotator angle in the x-y plane, φ. The Lagrangian of this system is also purely kinetic

$$L = \frac{MR^2}{2}\dot{\theta}^2 + \frac{MR^2}{2}\sin^2\theta\dot{\varphi}^2, \tag{2.18}$$

and the conjugate momenta are p_θ and $p_\varphi = MR^2 \sin\theta\dot{\varphi}$. Since the variable φ is cyclic, the z-component of the angular momentum, p_φ, is a constant of the motion. Thus

$$p_\varphi = m_\varphi\hbar. \tag{2.19}$$

The integer m_φ, is called the magnetic quantum number, and corresponds to the z component of the angular momentum that will be studied later. If a charged particle is at the end of the rotator, its magnetic moment is along the z axis and proportional to p_φ. This phenomenon, the quantization of the angular momentum along a z-axis, was called space quantization.[6]

---

[6] In the modern quantum mechanics, the angular momentum is quantized in the same way, but the process of quantization does not pick out a preferred axis. For this reason, the name 'space quantization' fell out of favor, and the same phenomenon is now called the quantization of the angular momentum.

2.2 The De Broglie Wave-Particle Duality

In 1923, Louis de Broglie proposed an extension of the wave-particle dualism, introduced by Einstein for the electromagnetic radiation, to particles with mass. In a brief and "ingenious article", according to Schrödinger, de Broglie showed that for a particle with periodic motion, frequency ν_0 and rest mass m_o, one can define two frequencies.[7] The first is the frequency

$$\nu_1 = \nu_0\sqrt{1-\beta^2}, \tag{2.20}$$

that an observer at rest measures, with $\beta = v/c$. This relation is a direct consequence of the time dilation in the special theory of relativity. The second frequency is a consequence of the inertia principle of the special theory of relativity ($E = mc^2$) and of Planck's postulate, $E = h\nu$, that in the particle's system of reference can be written as $h\nu_0 = m_0c^2$. For the observer at rest, it will be $h\nu = mc^2$. Therefore

$$\nu = \frac{\nu_0}{\sqrt{1-\beta^2}}. \tag{2.21}$$

In the same article, de Broglie showed that the relation between these frequencies

$$\nu_1 = \nu\big(1-\beta^2\big), \tag{2.22}$$

is consistent with the following two solutions: the function

$$\sin(2\pi\nu_1 t), \tag{2.23}$$

describing the particle's orbital motion, and the function

$$\sin\left(2\pi\nu t - kx\right), \tag{2.24}$$

describing the wave motion of the particle. If we follow the phase $\theta = 2\pi\nu t - kx$, the propagation velocity of a fixed or constant phase point (determined by $d\theta/dt = 0$) gives us the phase velocity $v_p = dx/dt = w/k$, that can be written as $v_p = \hbar\omega/\hbar k = c/\beta$. The phases $2\pi\nu_1 t$ and $2\pi\nu t\left(1 - x/tv_p\right)$ must be in harmony and resonate as the electron moves in its orbit. Indeed, if we take $2\pi\nu_1 t = 2\pi\nu t\left(1 - x/tv_p\right)$ and use the phase velocity $v_p = c/\beta$, we also obtain (2.22). De Broglie also suggested, that a necessary condition for the stability of the orbiting electrons is that the wavelength λ will fit an integer number of times in the orbit length $2\pi R$, i.e. that

$$2\pi R = n\lambda. \tag{2.25}$$

---

[7] L. de Broglie, *Comptes Rendus*, **177**, 507 (1923).

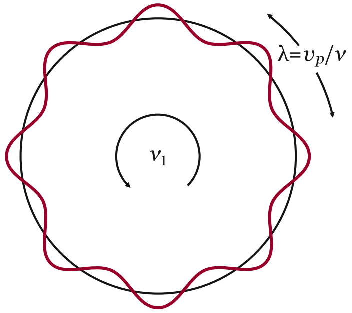


Fig. 2.1 In the de Broglie quantization condition the electron orbit length equals an integer number of the wavelength λ

If T is the period of the orbiting motion, it is easy to show, by simple substitution, that this stability condition can be written as

$$\frac{m_0\beta^2c^2T}{\sqrt{1-\beta^2}} = nh. \tag{2.26}$$

Taking into account that $m = m_0/\sqrt{1-\beta^2}$ and $\beta = v/c$, the last equation can be written in the form

$$mv^2 = nh\nu, \tag{2.27}$$

that is compatible with the quantization *assumption* used by N. Bohr for his phenomenological model, and with the quantization rule of Sommerfeld-Wilson-Ishiwara. From the stability condition (2.25) and the last equation, we obtain the important relation

$$mv\lambda = h, \tag{2.28}$$

that summarizes the de Broglie duality concept, generally expressed in the form

$$\lambda = \frac{h}{p}. \tag{2.29}$$

This λ is known as the *de Broglie wavelength*, for particles with mass. This result shows that (2.26) and the quantization conditions are equivalent to assuming the stability condition of the orbital motion, that fixes the length of the electron orbit as a multiple of its de Broglie's wavelength.

In his doctoral thesis, de Broglie expressed his hope that the wave behavior of particles would be observed experimentally. Indeed, electron-diffraction experiments by metal surfaces (C. Davisson and LH Germer 1927) and thin films (GP Thomson 1928) were performed later and showed that electron's wave properties, with wavelength $\lambda = h/mv$, manifest as predicted by de Broglie.

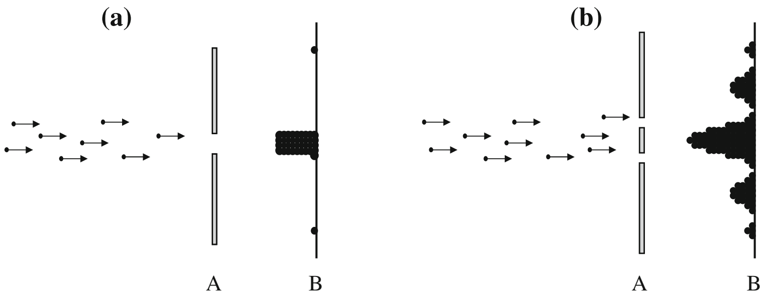


Fig. 2.2 The distribution of electrons passing through a single slit is almost rectangular, while that of electrons passing through a double slit resembles the constructive-destructive interference, characteristic of waves

2.3 Diffraction Experiments and the Wave-Like Nature of Particles

In addition to the electron diffraction experiments by Davisson and Germer and by Thomson, showing that electrons under certain circumstances manifest wave properties, the diffraction experiments by one and two slits, made also evident this kind of properties for quantum particles. In fact, if a beam of particles moves towards a screen (B), and in the way stands another screen with one slit (the screen A), the electrons distribution in B will be almost rectangular, as shown in Fig. 2.2a). However, when a second slit exists on the screen A, the distribution of particles in B is as shown in Fig. 2.2b), which is not the simple addition of two rectangular distributions, as one could expect for "microscopic" particles like sand grains, but a distribution that resembles the behavior of waves. This type of distribution contains clear features of constructive and destructive interference observed when a beam of light passes through the slits. It is worth noticing that each electron contributes with one point to the diffraction pattern, therefore the wave interference pattern emerges only when the number of points is large. This means, on one side, that the undulating characteristics of the diffraction pattern manifests as a collective behavior. However, electrons may arrive on the screen at different times.[8] In that case the common interpretation is that each quantum particle interferes with itself and, besides this, that the corpuscular or wave characteristics manifest themselves depending on whether the quantum particles are subject to a single or two slits diffraction experiment. What this means is still an open question. In any case, if electrons passing through the double slit are represented by the superposition of two spherical wave functions emerging from two slits, we can describe them with

---

[8] Recent experiments claim that similar distributions come out when particles are sent one by one.

$$\begin{aligned}\phi(\mathbf{r},t) &= a_1\phi_1 + a_2\phi_2 \\ &= a_1\left(e^{i\mathbf{k}\cdot\mathbf{r}_1} + c_1\frac{e^{ikr_1}}{r_1}\right)e^{-iwt} + a_2\left(e^{i\mathbf{k}\cdot\mathbf{r}_2} + c_2\frac{e^{ikr_2}}{r_2}\right)e^{-iwt},\end{aligned} \tag{2.30}$$

where $\mathbf{r}_1 = \mathbf{r} - \mathbf{d}/2$, $\mathbf{r}_2 = \mathbf{r} + \mathbf{d}/2$, $\mathbf{k} = k\hat{\mathbf{x}}$ and the origin in the line of length d that joins the two slits. The electrons distribution on the screen (at $x = L$), for the special case where $a_1 = a_2 = c_1 = c_2 = 1$, will be given by

$$|\phi(\mathbf{r},t)|^2_{x=L} = 4 + \frac{1}{r_1^2} + \frac{1}{r_2^2} + 4\frac{\cos k(L-r_1)}{r_1} + 4\frac{\cos k(L-r_2)}{r_2} + 2\frac{\cos k(r_1-r_2)}{r_1 r_2}.$$

This function describes the interference of waves passing through the slits. It has the highest maximum in the middle and secondary maxima when the condition $d\sin(\tan^{-1}y/L) = n\lambda$ is fulfilled. If the wave functions are described only by spherical waves, i.e., if

$$\phi(\mathbf{r},t) = a_1\phi_1 + a_2\phi_2 = a_1\frac{e^{ikr_1-iwt}}{r_1} + a_2\frac{e^{ikr_2-iwt}}{r_2}, \tag{2.31}$$

and $a_1 = a_2 = 1$, we will have

$$|\phi(\mathbf{r},t)|^2 = \frac{1}{r_1^2} + \frac{1}{r_2^2} + 2\,\frac{\cos k(r_1-r_2)}{r_1 r_2}. \tag{2.32}$$

This function is plotted in Fig. 2.3 for two values of t: for t_1 corresponding to $x = x_1 \simeq L/3$ and for t_2 corresponding to $x = 2x_1 \simeq 2L/3$. In this figure, the spherical waves moving away from the slits are also drawn. The geometric place of the interference points are indicated by blue lines. These lines coincide with the principal and the secondary wave maxima.

2.4 Schrödinger's Wave Mechanics

Motivated by the de Broglie theory, the Austrian physicist Erwin Schrödinger, who was not pleased with the quantization a la Bohr, had in mind that the quantization phenomenon could rather be related to an eigenvalue problem. In fact, in 1926 he published a series of papers in the "Annalen der Physik" with title: "Quantisierung und Eigenwertprobleme" (quantization and eigenvalue problems) where he proposed and developed the foundations of the Schrödinger wave mechanics. On January 26, 1926, Schrödinger sent the first of these papers[9] where he considers a variational problem and derives, among others, the equation

---

[9] E. Schrödinger, *Ann. Phys.* **79**, 361 (1926).

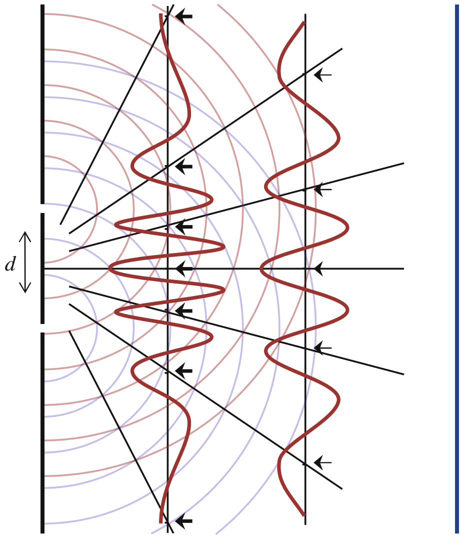


Fig. 2.3 The blue lines indicate the locus of the interference of spherical waves coming from two slits. The snapshots show the wave function $|\phi(\mathbf{r}, t)|^2$ in the equation (2.32), for times t_1 and t_2 such that $x = x_1 \simeq L/3$ and $x = 2x_1 \simeq 2L/3$. The position of the maxima (indicated by arrows) coincides with the blue lines

$$\nabla^2\psi + \frac{2m(E-V)}{\hbar^2}\psi = 0, \tag{2.33}$$

with ψ a real function and E the energy. As will be seen below, this equation is closely related with the wave equation

$$\nabla^2\psi - \frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2} = 0, \tag{2.34}$$

and with the Hamilton equation for particles with mass and total energy

$$E = \frac{p^2}{2m} + V(\mathbf{r}). \tag{2.35}$$

Schrödinger's equation is, in some ways, a hybrid in which the wavelike and the corpuscular descriptions merge. To make this fact evident let us now derive the stationary Schrödinger equation (2.33), starting from the wave equation.

Suppose we have the wave equation (2.34). A solution $\psi(\mathbf{r}, t)$, factored in the form

$$\psi(\mathbf{r}, t) = e^{-i\omega t}\varphi(\mathbf{r}), \tag{2.36}$$

and replaced in the wave equation, leaves us with the differential equation

$$\nabla^2\varphi(\mathbf{r}) + \frac{\omega^2}{v^2}\varphi(\mathbf{r}) = 0. \tag{2.37}$$

On the other side, let us consider the de Broglie relation $\lambda = h/p$ (that synthesizes the duality concept that matured during the first quarter of the twentieth century), as a bridge to make contact with the massive particles dynamics, which together with some identities introduced before leads to

$$\frac{\omega^2}{v^2} = \left(\frac{2\pi}{\lambda}\right)^2 = \frac{p^2}{\hbar^2}. \tag{2.38}$$

If we replace this relation and make use of the Hamilton equation (2.35) in (2.37), we have the celebrated Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\varphi(\mathbf{r}) + V(\mathbf{r})\varphi(\mathbf{r}) = E\varphi(\mathbf{r}), \tag{2.39}$$

known as *the stationary Schrödinger equation* that was quoted in (2.33). The potential function $V(\mathbf{r})$ is a characteristic property of the physical system. The difference between one system and another, resides basically in their potential functions and the boundary conditions. As in many problems of classical physics, the differential equation and the boundary conditions define not only the possible solutions φ_i but also the characteristic energy values E_i. The characteristic values are the parameter values for which the physical problem admits solutions. This is how the energy discretization phenomenon emerges, in a natural way. Solving the Schrödinger equation is, essentially, equivalent to solving an eigenvalue problem, as Schrödinger envisioned. In the subsequent chapters, we will discuss this fundamental problem, by solving specific examples.

In 1927, P.A.M. Dirac showed that quantum wave mechanics provides the same description as the matrix formalism of Heisenberg, Born and Jordan. They represent two faces of the same quantum theory. Other significant quantum equations are: the Pauli equation and the relativistic equations of Dirac and Klein-Gordon. In Chap. 11, we will briefly mention the Pauli equation. The relativistic equations are beyond the scope of this book.

Since the most general description depends on the spatial and time coordinates, it is worth showing one way to incorporate time in the Schrödinger formalism. To obtain the Schrödinger equation (2.39), it was crucial to write the wave equation in its steady-state form, which was possible once the wave function $\psi(\mathbf{r}, t)$ was factored as $\varphi(\mathbf{r})e^{-i\omega t}$ and the de Broglie relation $\mathbf{p} = \hbar\mathbf{k}$ was taken into account. If we derive $\psi(\mathbf{r}, t)$ with respect to time, we have

$$\frac{\partial\psi}{\partial t} = -i\omega\psi = -i\frac{e^{-i\omega t}}{\hbar}E\varphi(\mathbf{r}). \tag{2.40}$$

Combining this equation with the stationary Schrödinger equation we obtain the time dependent equation

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V(\mathbf{r})\psi = i\hbar\frac{\partial\psi}{\partial t}, \tag{2.41}$$

known as the complete Schrödinger equation. Its solutions depend on position $\mathbf{r}$ and time t. Unlike the classical physics where the physical variables, like position, velocity, etc., can be straightforwardly obtained when the equations of motion are solved, in the quantum theory these variables have to be obtained, like pulled out, from the solution $\varphi(\mathbf{r})e^{-i\omega t}$. We will discuss later, with more detail, the formal representation of the quantum theory, and the expected values of basic variables such as position, momentum, energy, etc.. In this formalism the classical dynamical variables appear generally as operators. If we look at the stationary equation (2.39) and the complete equation (2.41), and examine the Hamiltonian $\widehat{H} = -(\hbar^2/2m)\nabla^2 + V(\mathbf{r})$, it is clear that the operator $i\hbar\partial\psi/\partial t$ represents the energy, and $-i\hbar\nabla$ represents the linear momentum. We will later find other operators related to other dynamical variables. In Chap. 7, we will study properties of operators and their relation to physical observables from a more general perspective.

In the next chapters, we will solve the stationary Schrödinger equation for simple systems. Our goal is to introduce the quantum phenomenology and to present a comprehensive theory. We will calculate energy eigenvalues and eigenfunctions and we will define other very useful variables in the analysis of the quantum transport properties. We will address, when necessary, some interpretation issues. To avoid substitutions, simplify the mathematical handling of the differential-equation solutions, and gain some intuition, we will use the transfer matrix approach.

It is pertinent to mention that the rigorous derivation of the Schrödinger equation has been an important research problem. For the naive derivation presented here, we have used the wave equation and the empirical relation $p\lambda = h$. There have been many attempts to derive the Schrödinger equation from first principles. Standing out is the stochastic formulation of L. de la Peña and A. M. Cetto. The discussion of this topic exceeds the objectives of this text.

2.5 Solved Problems

Exercise 4 In Fig. 2.4 we show the diffraction pattern of a red-light beam scattered by a double slit system. Determine the light wavelength λ if the distance to the screen is $L = 17.5\,\mu\text{m}$, $B = 7.8\,\mu\text{m}$ and the slits separation is $d = 1.7\,\mu\text{m}$.

Solution The spherical waves coming out from slits 1 and 2 superpose constructively at point P. For a constructive interference, the difference ΔR between the path-length $\overline{1P}$ and the path-length $\overline{2P}$, must be equal or an integer multiple of the wave length λ. Close to the slits a small right triangle is formed with hypotenuse d and cathetus ΔR. If this triangle, is compared with the right triangle 1P01 (with hypotenuse $\overline{1P}$ and cathetus B), and B is the distance between first-neighbors maxima, the length ΔR must be equal to λ. Since

$$\alpha = \tan^{-1}\frac{B}{L} = \tan^{-1}\frac{7.8}{17.4} \simeq 24.15^\circ. \tag{2.42}$$

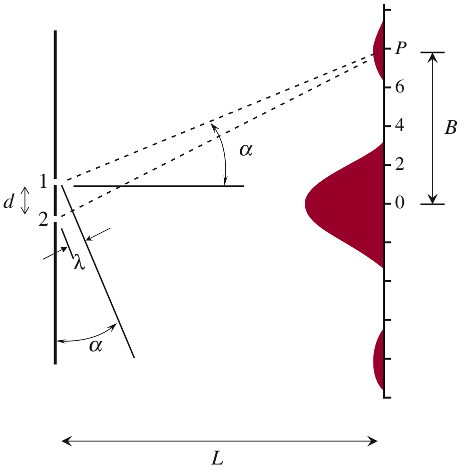


Fig. 2.4 Diffraction pattern of a red-light beam scattered by a double slit system

we have

$$\lambda = d \sin \alpha = 1.7 \sin 24.15^\circ = 0.695 \mu\text{m} = 695\text{nm}. \tag{2.43}$$

Exercise 5 Suppose that we accelerate electrons and they reach energies of the order of 100 GeV. Assuming that the relativistic relation $E = \sqrt{p^2c^2 + m^2c^4}$ holds, determine the de Broglie wavelength of those electrons.

Solution From $E = \sqrt{p^2c^2 + m^2c^4}$ it is easy to obtain the following relation

$$p = m_e v = \frac{\sqrt{E^2 - (m_e c^2)^2}}{c}. \tag{2.44}$$

Replacing the energy and the electron rest mass ($m_e c^2$, measured in eV) we have

$$\begin{aligned} m_e v &= \frac{\sqrt{(100\,10^9)^2 - (.510999\,10^6)^2}\text{eV s}}{2.99792458\,10^8\text{m}} \frac{1.60218\,10^{-19}\text{J}}{1\text{eV}}, \\ &= 5.344297\,10^{-18}\frac{\text{J s}}{\text{m}}, \end{aligned} \tag{2.45}$$

and the de Broglie wavelength will be

$$\begin{aligned} \lambda &= \frac{h}{p} = \frac{6.6261\,10^{-34}\text{J s m}}{5.344297\,10^{-18}\text{J s}}, \\ &= 1.23984610^{-6}\text{Å}. \end{aligned} \tag{2.46}$$

This negligible wavelength makes these electrons useful to explore systems with comparable dimensions, for instance the atomic nuclei whose dimensions are of the order of 10^{-15}m.

2.6 Problems

1. Prove the expression (2.13) in which
$$E_n = -\frac{2\pi^2 m e^4}{h^2 n^2}. \tag{2.47}$$
2. Consider the functions $\sin\bigl(2\pi\nu_1 t\bigr)$ and $\sin\bigl(2\pi\nu t(1 - v/v_p)\bigr)$, and show that the de Broglie stability condition
$$2\pi R = n\,\lambda \tag{2.48}$$
is equivalent to assuming that
$$\nu_1 = \frac{1}{n}\,\nu\bigl(1 - \beta^2\bigr). \tag{2.49}$$
Plot both functions and verify this conclusion.
3. Find the electron orbits radii of the Hydrogen atom.
4. Find the first, the second and the third energy levels in the Hydrogen atom.
5. From the equality of the phases, in the de Broglie analysis, derive the relation (2.22).
6. Show that the equation (2.26) can also be obtained from the Sommerfeld-Wilson-Ishiwara quantization rule
$$\oint p_\varphi d\varphi = nh, \tag{2.50}$$
where $p_\varphi = mR^2\omega$.
7. If the de Broglie wavelength of the first energy level electron is equal to the first circular orbit length, how fast does the electron move in that orbit?
8. How fast does one electron move in the second and the third energy levels, and how fast in the n-th energy level?
9. Show that (2.33) and (2.39) are essentially the same.

Chapter 3
Properties of the Stationary Schrödinger Equation

In this chapter we study some general properties of the solutions of the stationary equation. These properties will be observed in the specific examples that will be studied throughout this text. Among them, we will mention the relation between the number of classical-return points and the existence or not of energy quantization. We will study the free particle problem, where no quantization exists, and the infinite quantum well problem, where the particle energy quantizes. In this chapter we introduce also the Dirac notation and the particle current density.

3.1 Quantization as an Eigenvalue Problem

If the potential function is independent of time, the complete Schrödinger equation is separable. This means that we can propose a solution like

$$\psi(\mathbf{r}, t) = \varphi(\mathbf{r})\tau(t). \tag{3.1}$$

Replacing this function in the Schrödinger equation we obtain, on one hand the equation

$$i\hbar\frac{\partial\tau(t)}{\partial t} = E\tau(t), \tag{3.2}$$

with solution $\tau(t) = e^{-iEt/\hbar}$, and on the other the equation

$$-\frac{\hbar^2}{2m}\nabla^2\varphi + V(\mathbf{r})\varphi(\mathbf{r}) = E\varphi(\mathbf{r}), \tag{3.3}$$

whose solutions will be obtained once the potential function $V(\mathbf{r})$ and the boundary conditions will be given. In some cases, not always, solving this equation is essentially equivalent to solving an eigenvalue problem. In those cases one tries to obtain the

P. Pereyra, *Fundamentals of Quantum Physics*, Undergraduate Lecture Notes in Physics,
DOI: 10.1007/978-3-642-29378-8_3, © Springer-Verlag Berlin Heidelberg 2012

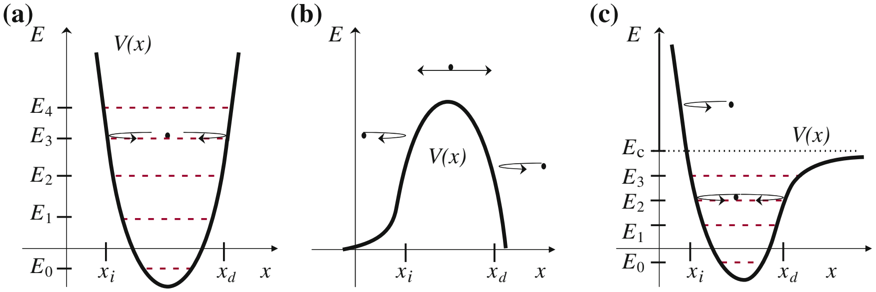


Fig. 3.1 Different types of one-dimensional potential. In (**a**) the potential is of a double classical return point and the energy discretizes. In (**b**) there is one return point for energies below the potential height, and no return points for energies above the potential height. In this case all values of the energy are possible. In (**c**) the energy discretizes only if $E < E_c$. Confining potentials discretize the energy E

eigenvalues $\{E_n\}$ and the eigenfunctions $\{\varphi_n\}$ that satisfy the eigenvalue equation

$$-\frac{\hbar^2}{2m}\nabla^2\varphi_n(\mathbf{r}) + V(\mathbf{r})\varphi_n(\mathbf{r}) = E_n\varphi_n(\mathbf{r}). \tag{3.4}$$

When Schrödinger's equation admits solutions for only a discrete set of energy values, we say that the energy is quantized. This is not a requirement for a system to be considered a quantum system. There is an uncountable number of systems which Schrödinger equation has solutions for any value of the energy. Other properties, like the tunneling effect and the spin, characterize also the quantum behavior. A few lines below we will comment on the physical conditions that must be present to quantize the energy. In any case, whether the energy quantizes or not, the solutions of the Schrödinger equation are required to satisfy the *continuity* and *finiteness* conditions and they must be *single-valued* and *square integrable* functions, so that integrals like

$$\int \varphi^*(\mathbf{r})\varphi(\mathbf{r})d^3r, \tag{3.5}$$

can be evaluated throughout the domain of definition of the physical system.

In the Schrödinger equation, the energy E plays the role of a parameter. It turns out that the energy quantization depends mainly on the form of the potential function $V(\mathbf{r})$. In one-dimensional (1D) potentials, like in Fig. 3.1, distinct results can be expected. For example:

1. the potential of Fig. 3.1a with two classical return points, x_i and x_d, confines the particle. In cases like this, the energy will quantize, and one can search for the eigenfunctions φ_n and the corresponding energy eigenvalues E_n.
2. the potential with a single classical return point, or no return point as in Fig. 3.1b, does not confine the classical particle. It can move away, either in one or both

directions. In cases like this, there will be no energy quantization, and the Schrödinger equation admits solutions for all values of E.

3. the potential in Fig. 3.1c, with two return points when $E < E_c$ and only one for energies $E > E_c$, leads to a mixed problem. For energies below E_c, the potential confines the classical particle, and the quantum particle lives in a discrete set of states corresponding to the eigenfunctions φ_n. When the energy is greater than E_c, there is no confining and no quantization.

3.2 Degenerate Eigenfunctions, Orthogonality and Parity

When the physical system has some symmetry it is possible to find more than one eigenfunction for the same energy value. In this case the eigenfunctions are called degenerate eigenfunctions. Otherwise, when only one eigenfunction φ_n exists, for each eigenvalue E_n, we have non-degenerate solutions. We will show now that when the eigenfunctions are non-degenerate and the potential V is real, the eigenfunctions are orthogonal.

Suppose that we have a set of *non-degenerate eigenfunctions* φ_n, normalized as follows

$$\int \varphi_n^*(\mathbf{r})\varphi_n(\mathbf{r})d^3r = 1. \tag{3.6}$$

Consider now the eigenvalue equation:

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\varphi_i = E_i\varphi_i, \tag{3.7}$$

and its complex conjugate

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\varphi_j^* = E_j\varphi_j^*, \tag{3.8}$$

with $j \neq i$. If we multiply the first of these equations by φ_j^* and the second by φ_i, subtracting from each other and integrating we obtain:

$$\left(E_i - E_j\right)\int\limits_{\mathrm{V}} \varphi_j^*\varphi_i d^3r = \frac{\hbar^2}{2m}\int\limits_{\mathrm{V}} \left(\varphi_j^*\nabla^2\varphi_i - \varphi_i\nabla^2\varphi_j^*\right) d^3r. \tag{3.9}$$

Since

$$\nabla\cdot\left(\varphi_j^*\nabla\varphi_i - \varphi_i\nabla\varphi_j^*\right) = \left(\varphi_j^*\nabla^2\varphi_i - \varphi_i\nabla^2\varphi_j^*\right), \tag{3.10}$$

we can write (3.9) in the form

$$\left(E_i - E_j\right) \int\limits_{\mathrm{V}} \varphi_j^*(\mathbf{r})\varphi_i(\mathbf{r}) d^3r = \frac{\hbar^2}{2m} \int\limits_{\mathrm{V}} \nabla \cdot \left(\varphi_j^* \nabla \varphi_i - \varphi_i \nabla \varphi_j^*\right) d^3r. \quad (3.11)$$

Using the divergence theorem, the integral on the right side becomes a closed surface integral. Thus

$$\left(E_i - E_j\right) \int\limits_{\mathrm{V}} \varphi_j^*(\mathbf{r})\varphi_i(\mathbf{r}) d^3r = \frac{\hbar^2}{2m} \oint\limits_{S} \left(\varphi_j^* \nabla \varphi_i - \varphi_i \nabla \varphi_j^*\right) \cdot dS. \quad (3.12)$$

We know from the divergence theorem that the only restriction on the surface integral is that the closed surface should contain the volume V. The closed surface can be zoomed, as much as one would like, keeping the flux $\varphi_j^* \nabla \varphi_i - \varphi_i \nabla \varphi_j^*$ constant. To evaluate the integral it is convenient to zoom the surface to infinity, where the functions and their gradients should tend to zero by the finiteness condition. Accordingly, we have

$$\left(E_i - E_j\right) \int\limits_{\mathrm{V}} \varphi_j^*(\mathbf{r})\varphi_i(\mathbf{r}) d^3r = 0. \quad (3.13)$$

Since $E_i \neq E_j$, we can conclude that the non-degenerate solutions are orthogonal, i.e. that

$$\left\langle \varphi_j | \varphi_i \right\rangle = \int\limits_{\mathrm{V}} \varphi_j^*(\mathbf{r})\varphi_i(\mathbf{r}) d^3r = 0. \quad (3.14)$$

We have introduced here the Dirac notation of “bra-kets” (angle parenthesis) to denote the integral of the product of φ_j^*, represented by the “bra” $\left\langle \varphi_j \right|$, with φ_i, represented by the “ket” $|\varphi_i\rangle$. This is a compact notation, especially useful when general derivations and obvious integrals are performed and one needs to carry out only the relevant variables. At the end of this chapter we will show schematically the most important relations between the standard functions and the bra-ket notation. If the orthogonal functions are normalized, we obtain

$$\left\langle \varphi_j | \varphi_i \right\rangle = \int\limits_{\mathrm{V}} \varphi_j^*(\mathbf{r})\varphi_i(\mathbf{r}) d^3r = \delta_{ij}, \quad (3.15)$$

with δ_{ij} the Kronecker symbol defined as

$$\delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j, \\ 0 & \text{if} \quad i \neq j. \end{cases} \quad (3.16)$$

The symmetries of the wave functions are compatible with those of the Hamiltonian function. When the potential function remains invariant under the parity transformation

$$\widehat{\mathbf{P}} : \mathbf{r}_o + \mathbf{r}' \mapsto \mathbf{r}_o - \mathbf{r}', \tag{3.17}$$

that changes the sign of the coordinate $\mathbf{r}$ with respect to the symmetry point $\mathbf{r}_o$, the wave function remains invariant or changes sign. If

$$\widehat{\mathbf{P}}\varphi_e(\mathbf{r}_o + \mathbf{r}') = \varphi_e(\mathbf{r}_o - \mathbf{r}') = \varphi_e(\mathbf{r}_o + \mathbf{r}'), \tag{3.18}$$

the wave function has even parity with respect to the symmetry point $\mathbf{r}_o$, if instead it satisfies the relation

$$\widehat{\mathbf{P}}\varphi_o(\mathbf{r}_o + \mathbf{r}') = \varphi_o(\mathbf{r}_o - \mathbf{r}') = -\varphi_o(\mathbf{r}_o + \mathbf{r}'), \tag{3.19}$$

the wave function has odd parity. To simplify the analysis of parity it is convenient to choose the origin at the symmetry point. In that case

$$\widehat{\mathbf{P}} : \mathbf{r} \longmapsto -\mathbf{r}, \tag{3.20}$$

and the wave functions with even parity satisfy the relation

$$\varphi_e(-\mathbf{r}') = \varphi_e(\mathbf{r}'), \tag{3.21}$$

while those with odd parity satisfy the relation

$$\varphi_o(-\mathbf{r}') = -\varphi_o(\mathbf{r}'). \tag{3.22}$$

When a function $\varphi(\mathbf{r})$ does not have a definite parity, it is always possible to write

$$\varphi(\mathbf{r}) = \varphi_e(\mathbf{r}) + \varphi_o(\mathbf{r}) \tag{3.23}$$

with

$$\varphi_e(\mathbf{r}) = \frac{1}{2}\Big(\varphi(\mathbf{r}) + \varphi(-\mathbf{r})\Big) \quad \text{and} \quad \varphi_o(\mathbf{r}) = \frac{1}{2}\Big(\varphi(\mathbf{r}) - \varphi(-\mathbf{r})\Big). \tag{3.24}$$

We will now study two simple but illustrative problems: the free particle and a particle inside an infinite quantum well. At the end of this chapter we will briefly comment on the physical interpretation of the Schrödinger equation solutions. As an example of the importance of having the eigenvalues and the eigenfunctions of a Schrödinger equation, we will introduce the particle current density. Other quantities, relevant in the quantum description, are generally evaluated in terms of eigenvalues and eigenfunctions. For this reason one of the first aims is to obtain eigenvalues and eigenfunctions.

3.3 The Free Particle

In the classical description a free particle moves without any interaction along a straight line. In the quantum description we have a free particle when the potential function is $V(\mathbf{r}) = 0$. This is in principle the simplest quantum system. Although the complete absence of interaction is impossible in the real world, the free particle model is often a good approximation when the local interaction is negligible. If the particle energy is E and the direction in which it moves is x, the stationary Schrödinger equation is simply

$$-\frac{\hbar^2}{2m}\frac{d^2\varphi(x)}{dx^2} = E\varphi(x), \quad -\infty < x < \infty. \tag{3.25}$$

In this system, without boundaries, the solutions should also fulfill the continuity and finiteness condition at any point, from $x = -\infty$ to $x = \infty$. As is usual for the second order differential equations with constant coefficients, we propose as a solution the function

$$\varphi(x) = e^{rx}. \tag{3.26}$$

If we substitute this function in (3.25) for $\varphi(x)$, we get the quadratic equation

$$r^2 = -\frac{2m}{\hbar^2}E, \tag{3.27}$$

that gives us the parameter r. It is evident that, for positive energy, the parameter r must be a purely imaginary number, i.e.

$$r = \pm i\sqrt{\frac{2m}{\hbar^2}E} = \pm ik. \tag{3.28}$$

The factor $k = \sqrt{2mE/\hbar^2}$ is the free particle's wave number.[1] Replacing these values of r in the proposed solution (3.26), we have the general solution

$$\varphi(x) = ae^{ikx} + be^{-ikx} = a\,\varphi^+(x) + b\varphi^-(x). \tag{3.31}$$

---

[1] If we take into account that $E = p^2/2m$ and use the relation $p = h/\lambda$, it is clear that this is precisely its nature. Indeed

$$\frac{1}{\lambda^2} = \frac{2m}{h^2}E. \tag{3.29}$$

Multiplying by $(2\pi)^2$ we have (remember that $\hbar = h/2\pi$)

$$\left(\frac{2\pi}{\lambda}\right)^2 = \frac{2m}{\hbar^2}E = k^2. \tag{3.30}$$

Given this solution, we also have the wave function $\psi(x,t)$

$$\psi(x,t) = ae^{i(kx-\omega t)} + be^{-i(kx+\omega t)} = a\psi^{+}(x,t) + b\psi^{-}(x,t). \tag{3.32}$$

The first term describes a wave propagating in the positive x direction and the second in the negative x direction.

As will be seen in the next chapters, it is some times convenient to represent the wave functions as vector functions. With the right- and left-moving components of the wave function, we can define, for instance, the state vector

$$\phi(x) = \begin{pmatrix} ae^{ikx} \\ be^{-ikx} \end{pmatrix} = \begin{pmatrix} a\varphi^{+}(x) \\ b\varphi^{-}(x) \end{pmatrix}. \tag{3.33}$$

Since

$$\begin{pmatrix} ae^{ikx_2} \\ be^{-ikx_2} \end{pmatrix} = \begin{pmatrix} e^{ik(x_2-x_1)} & 0 \\ 0 & e^{-ik(x_2-x_1)} \end{pmatrix} \begin{pmatrix} ae^{ikx_1} \\ be^{-ikx_1} \end{pmatrix} \tag{3.34}$$

$$\phi(x_2) = M(x_2 - x_1)\phi(x_1), \tag{3.35}$$

it is clear that knowing the state vector of a free particle at some point x_1, we can obtain the state vector at any other point x_2. The diagonal matrix $M(x_2 - x_1)$ provides the space evolution of the wave function phase.

The coefficients a and b define the relative amplitude of each component. Inasmuch as the coefficients are unknown, we have not yet solved the problem. We must try to determine the coefficients a and b. In some cases the initial conditions allow a first evaluation. For example, if we know that the free particles move only in one direction, one of the coefficients must vanish. We are then left with just one coefficient to determine. In that case, it will be determined by the wave function normalization. In most applications one needs normalized free-particle wave functions. As mentioned before, the wave functions φ^{+} and φ^{-} are normalized if

$$\int_{-\infty}^{\infty} \varphi^{\pm *}(x)\varphi^{\pm}(x)dx = 1. \tag{3.36}$$

It turns out that this is not a trivial problem for the free particle wave functions. It is easy to see that the normalization integral diverges. Indeed

$$\int_{-\infty}^{\infty} e^{\mp ikx}e^{\pm ikx}dx = \infty! \tag{3.37}$$

This makes the normalization of e^{ikx} and e^{-ikx} a major problem, in contrast with the ease in solving Schrödinger's free particle equation. To control this singularity, different criteria were proposed to normalize the free particle wave function. Among them we will discuss the Born and the Dirac normalization procedures. We will deal now with this problem. Before continuing, it is important to notice that in the case of free particles there is no energy quantization. This means that the problem admits solutions for all positive values of E. As mentioned before, normalizing a function $A\varphi(x)$, defined on the domain $(-a, a)$, means to find the factor A such that

$$A^*A\int_{-a}^{a}\varphi^*(x)\varphi(x)dx = 1. \tag{3.38}$$

In the Born normalization the domain of integration is restricted to a finite length L, which afterwards may tend to infinity. This is a highly controversial but widely used procedure. In this case

$$A^*A\int_{-L/2}^{L/2} e^{-ikx}e^{ikx}dx = A^*AL = 1. \tag{3.39}$$

If A is real and positive, the normalized free-particle wave function *a la Born* will be:

$$\varphi^{\pm}(x) = \frac{1}{\sqrt{L}}e^{\pm ikx}. \tag{3.40}$$

Another important normalization and also widely used is the Dirac normalization. It is based on the Dirac's delta function $\delta(x - x')$. This is a subtle and important mathematical object. To define the Dirac's delta function we mention here only the most relevant features. First of all, the delta function $\delta(x - x')$ is zero everywhere except at $x = x'$ and has, among others, the following properties:

$$\int_a^b \delta(x - x')dx = 1, \qquad a < x' < b; \tag{3.41}$$

$$\int_a^b f(x)\delta(x - x')dx = f(x'). \tag{3.42}$$

There are several specific representations for the delta function $\delta(x - x')$. Let us now obtain one of them using Fourier transforms. It is well-known that the Fourier transform of a function $f(x)$ is defined as

$$F(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ikx} dx, \tag{3.43}$$

and the inverse Fourier transform as

$$f(x') = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(k) e^{ikx'} dk. \tag{3.44}$$

Substituting (3.43) for $F(k)$ in (3.44), we have

$$f(x') = \int_{-\infty}^{\infty} f(x) \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x'-x)} dk \right] dx. \tag{3.45}$$

Comparing this equation with (3.41), it is clear that a possible representation of the Dirac's delta function is ($k = p/\hbar$)

$$\delta(x' - x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x'-x)} dk = \frac{1}{2\pi\hbar} \int_{-\infty}^{\infty} e^{ip(x'-x)/\hbar} dp. \tag{3.46}$$

In a similar way, with (3.43) and (3.44), we obtain

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ik'x} e^{ikx} dx = \delta(k' - k), \tag{3.47}$$

for the orthogonality condition of the free-particle wave functions with different wave vectors. This condition can also be written as

$$|A|^2 \int_{-\infty}^{\infty} e^{-ik'x} e^{ikx} dx = |A|^2 \int_{-\infty}^{\infty} e^{-ix(p'-p)/\hbar} dx = 2\pi\hbar |A|^2 \delta(p - p'). \tag{3.48}$$

If we choose the real coefficient A, such that $2\pi\hbar A^2 = 1$, the normalized free particle function *a la Dirac* will be

$$\varphi(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} = \frac{1}{\sqrt{2\pi\hbar}} e^{ikx}. \tag{3.49}$$

The one and three-dimensional free-particle wave functions can be written, respectively, as:

$$\varphi_k(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ikx} \quad \text{and} \quad \varphi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{2\pi\hbar}} e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{3.50}$$

or in the form

$$\varphi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \quad \text{and} \quad \varphi_{\mathbf{p}}(\mathbf{r}) = \frac{1}{\sqrt{2\pi\hbar}} e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}, \tag{3.51}$$

with the wave number (or the momentum) as a parameter. This notation is similar to that of the eigenfunctions, where the subindex is a quantum number, usually discrete. In conclusion, for the wave functions of free particles, we have the normalization condition

$$\langle \varphi_{p'} | \varphi_p \rangle = \int_{-\infty}^{\infty} \varphi^*_{p'}(x) \varphi_p(x) dx = \delta(p - p'), \tag{3.52}$$

that should not be confused with the closure condition

$$\langle \varphi_p(x') | \varphi_p(x) \rangle = \int_{-\infty}^{\infty} \varphi^*_p(x') \varphi_p(x) dp = \delta(x - x'), \tag{3.53}$$

of Hilbert's space of functions, to which the free particle functions[2] belong. The Dirac normalization is similar to (3.15). The divergency is not removed but put under control by means of a sui generis function, the Dirac delta function $\delta(p - p')$ that is zero everywhere and diverges at $p = p'$. The interesting property of the delta "functions" is that at the divergency point it increases in such a way that the integral at the singularity of the "function" is 1, as inferred from Eq. (3.41).

Another very useful normalization is based on the particle current density, that will be defined later. We will see that one has a unit particle current density when

$$\varphi^{\pm}(x) = \frac{1}{\sqrt{\hbar k/m}} e^{\pm ikx}. \tag{3.54}$$

This will be referred to as the unit-flux normalization.

3.3.1 The Physical Meaning of the Free Particle Solutions

One of the most discussed issues in the history of the quantum theory has been the physical interpretation of the Schrödinger equation solutions. Regardless of the conceptual problems and paradoxes (that will be considered briefly later), there is

---

[2] The Hilbert space of square integrable functions.

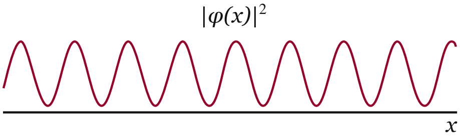


Fig. 3.2 The probability density function to find a free particle along the x-axis

an agreement that: (1) the most general solution $\varphi(\mathbf{r})$ must be expressed as a linear combination of linearly independent solutions $\varphi_i(\mathbf{r})$; (2) the function $\varphi(\mathbf{r})$ represents the probability density amplitude of finding the particle at point $\mathbf{r}$ and (3) $|\varphi(\mathbf{r})|^2$ gives the probability density function of finding the particle at point $\mathbf{r}$. In the absence of a confinement potential, the linearly independent solutions are usually complex. In the free particle case, the wave functions: $Ae^{i(kx-\omega t)}$ and $Ae^{-i(kx+\omega t)}$, describe particles propagating to the right and to the left, respectively. They represent the probability amplitude of finding the particle[3] at point x. The wave nature of these functions, is evident in the behavior of their real part $A\cos(kx \pm \omega t)$ and imaginary part $A\sin(kx \pm \omega t)$. If the physical conditions are such that the particles can move in both directions, the most general solution $\varphi(x)$ will be given by the superposition[4]

$$\varphi(x) = a\,Ae^{ikx} + bAe^{-ikx}, \tag{3.56}$$

and the probability density function of finding the particle at any point x will be:

$$|\varphi(x)|^2 = |A|^2\left(|a|^2 + |b|^2 + 2\Re\mathrm{e}\; ab^* e^{i2kx}\right). \tag{3.57}$$

When the right-moving and left-moving wave functions are normalized (with $A = 1/\sqrt{L}$ for the Born normalization, or $A = 1/\sqrt{2\pi h}$ for the Dirac normalization), the probability density function $|\varphi(x)|^2$, using the orthogonality condition in (3.47), will be normalized when $|a|^2 + |b|^2 = 1$. If a and b are equal, we have $a = b = 1/\sqrt{2}$. In this case, the probability density function $|\varphi(x)|^2$ will be (see Fig. 3.2):

$$|\varphi(x)|^2 = \begin{cases} \frac{1}{L}\cos 2kx + c_B & \text{for Born's normalization,} \\ \frac{1}{2\pi h}\cos 2kx + c_L & \text{for Dirac's normalization,} \end{cases} \tag{3.58}$$

with $c_B = 1/L$ and $c_L = 1/2\pi h$. When the particles move only to the right or only to the left, the probability density function to find them at point x is constant and equal to $|Ae^{i(kx\pm\omega t)}|^2 = A^2$.

---

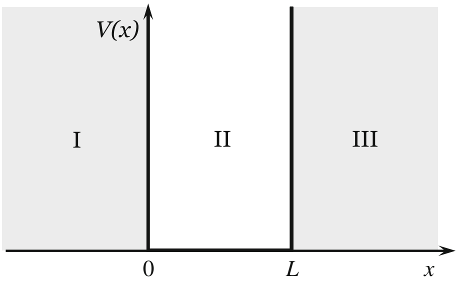
[3] Or finding the ensemble of particles at the point x, if we consider that the wave function to describe an ensemble of free particles.

[4] It is usual to write the superposition of the right- and left-moving solutions just as

$$\varphi(x) = a'\,e^{ikx} + b'\,e^{-ikx}. \tag{3.55}$$

It is easy to show that after normalizing the wave function $\varphi(x)$ one ends up with the same results (see Problem 3.1) as combining orthonormal functions like in (3.56).

Fig. 3.3 The infinite quantum well potential



3.4 The Infinite Quantum Well

In this example we have again a 1D system (whose generalization to 3D is immediate). In the 1D case the potential is

$$V(x) = \begin{cases} 0 & 0 < x < L; \\ \infty & x \leq 0 \ \text{ and } \ x \geq L. \end{cases} \tag{3.59}$$

In the regions where the potential is infinite (regions I and III in Fig. 3.3), the solutions and the probabilities of finding a particle, with finite energy, are zero. Therefore:

$$\varphi_{\mathrm{I}}(x) = \varphi_{\mathrm{III}}(x) = 0. \tag{3.60}$$

In region II, where the potential is zero, the Schrödinger equation is

$$-\frac{\hbar^2}{2m}\frac{d^2\varphi_{\mathrm{II}}(x)}{dx^2} = E\varphi_{\mathrm{II}}(x), \qquad 0 < x < L. \tag{3.61}$$

The solution must satisfy the following boundary conditions

$$\varphi_{\mathrm{II}}(0) = \varphi_{\mathrm{II}}(L) = 0. \tag{3.62}$$

The differential Eq. (3.61) is the same as for the free particle in (3.25). Thus our solutions can also be written as

$$\varphi_{\mathrm{II}}(x) = ae^{ikx} + be^{-ikx} \qquad \text{for} \qquad 0 < x < L. \tag{3.63}$$

The main difference between these solutions and those of the free particle resides in the boundary conditions (3.62). They make the problem completely different. The boundary condition $\varphi_{\mathrm{II}}(0) = 0$ implies that

$$a = -b. \tag{3.64}$$

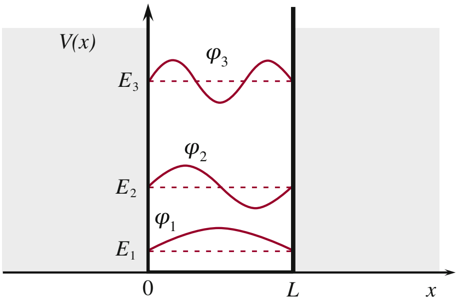


Fig. 3.4 The three lowest energy levels E_i and the corresponding eigenfunctions $\varphi_i(x)$ of the infinite quantum well

Replacing this relation in (3.63) we have

$$\varphi_{\mathrm{II}}(x) = A \sin kx. \tag{3.65}$$

At this point we have still to impose the condition $\varphi_{\mathrm{II}}(L) = 0$. The vanishing of $A \sin kx$ when $x = L$ is possible only if

$$kL = n\pi \quad n = 1, 2, \ldots \tag{3.66}$$

This is a quantization condition for the wave number k, and also for the energy. Introducing Eq. (3.28) we have now

$$E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{n^2 \hbar^2 \pi^2}{2mL^2}. \tag{3.67}$$

This is the first example where the quantization of the energy arises naturally by solving the Schrödinger equation. From our previous discussion on the classical turning points, we could certainly expect the quantization of the energy. As mentioned there the energy values for which the equation has solutions are the eigenvalues E_n. Replacing the wave number k_n it is easy to show that the solution that corresponds to the eigenvalue E_n is the eigenfunction

$$\varphi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi}{L} x. \tag{3.68}$$

In Fig. 3.4 we give the three lowest energy levels and the corresponding wave functions. The number of nodes of the eigenfunctions grows with the quantum number n. This means that the eigenfunctions of higher energies oscillate more than those of lower energies. This is a general characteristic that will be found always.

The infinite quantum well eigenfunctions possess well defined parity symmetries. The infinite quantum well potential of Fig. 3.4 has reflection symmetry around $x = L/2$, thus $x_o = L/2$ is a parity symmetry point. If we change x by $x_o + x'$, where x' is measured from the symmetry point, the eigenfunctions $\varphi_n(x)$ become

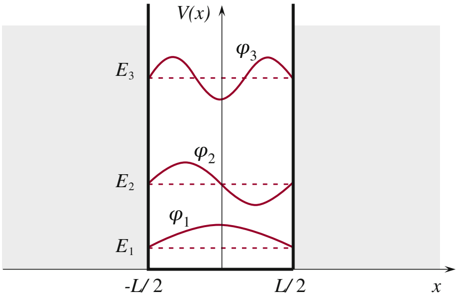


Fig. 3.5 The first energy levels and the corresponding eigenfunctions of an infinite quantum well. The origin is at the center of the well

$$\varphi_n(x_o + x') = \sqrt{\frac{2}{L}}\sin\Big(\frac{n\pi}{2} + \frac{n\pi}{L}x'\Big) \qquad \text{with} \qquad -L/2 < x' < L/2, \tag{3.69}$$

that can be written as

$$\varphi_n(x_o + x') = \begin{cases} (-1)^{\frac{n}{2}}\sqrt{\frac{2}{L}}\sin\frac{n\pi}{L}x' & \text{for} \quad n = 2, 4, \ldots \\ \\ (-1)^{\frac{n-1}{2}}\sqrt{\frac{2}{L}}\cos\frac{n\pi}{L}x' & \text{for} \quad n = 1, 3, \ldots \end{cases} \tag{3.70}$$

It is easy to verify that the eigenfunctions with n odd, like $\varphi_1(x)$ and $\varphi_3(x)$ satisfy the condition

$$\varphi_n(x_o - x') = \varphi_n(x_o + x'), \tag{3.71}$$

characteristic of even parity functions, according to (3.18), while the eigenfunctions with n even like $\varphi_2(x)$ and $\varphi_4(x)$, satisfy the condition

$$\varphi_n(x_o - x') = -\varphi_n(x_o + x'), \tag{3.72}$$

characteristic of odd parity functions with respect to the symmetry point, according to (3.19). One can simplify the analysis of the eigenfunctions parity by moving the origin of the coordinate system to the symmetry point as in Fig. 3.5 (see Problem 2). In that case, it is easy to verify that the eigenfunctions like $\varphi_1(x)$ and $\varphi_3(x)$ are even under the change of sign of x, and the eigenfunctions like $\varphi_2(x)$ and $\varphi_4(x)$ are odd under the change of sign of x.

What kind of information do the eigenfunctions and the eigenvalues provide? Does quantum mechanics predict the behavior of a particle or rather that of a collection of equivalent particles? The answers to some of these questions are well established in some senses, but in others it is a controversial subject. The eigenvalues of a physical quantity, for example the energy, are those values that the system allows for the quantum particle to have. In fact, if one measures the particle's energy, the experimental value will match one of the predicted eigenvalues. Therefore, the energy eigenvalues constitute the set of possible values for the particle energy. Before the experiment is performed, one does not know which energy state the particle is actually in. Hence, it should be described in terms of the most general solution of the Schrödinger equation, i.e. by the combination

$$\Phi(x) = c_1\varphi_1(x) + c_2\varphi_2(x) + c_3\varphi_3(x) + \ldots = \Sigma_i c_i \varphi_i(x). \tag{3.73}$$

The coefficients c_i of this superposition give the relative participation ratio, i.e. the weight with which each eigenstate participates in the total wave function Φ. In some problems the coefficients can be inferred by physical considerations, in others from the experimental observation. In a time-dependent potential, the coefficients can change with time.

The product $\Phi^*(x)\Phi(x)$ represents again the probability density function of finding the particle at point x. If the eigenfunctions $\varphi_i(x)$ are normalized, the integral

$$\langle \Phi|\Phi \rangle \equiv \int\limits_{-\infty}^{\infty} \Phi^*(x)\Phi(x)dx = \int\limits_{0}^{L} \Phi^*(x)\Phi(x)dx, \tag{3.74}$$

$$\langle \Phi|\Phi \rangle = \Sigma_{ij} c_i^* c_j \left\langle \varphi_i|\varphi_j \right\rangle = \Sigma_{ij} c_i^* c_j \delta_{ij} = \Sigma_i \left|c_i\right|^2, \tag{3.75}$$

represents the probability of finding the particle in the quantum well, which must certainly be equal to 1. Therefore we have

$$\langle \Phi|\Phi \rangle = \Sigma_i \left|c_i\right|^2 = 1. \tag{3.76}$$

This relation makes clear the physical meaning of the coefficients c_i in the superposition (3.73). While $|\varphi_i|^2$ gives the probability density function of finding the particle at point x, *when it is in the state* i, $|c_i|^2$ gives the probability of finding the particle in the state i, i.e. the relative weight of the state φ_i in the superposition (3.73).

When we solve the Schrödinger equation for a particle in the infinite quantum well, we obtain the energy eigenvalues E_n and the corresponding eigenfunctions φ_n. With these quantities the theory tells us which are the states where the particle could be found, but it cannot tell us in which state will actually the particle be found. When the experiment is done, the particle will certainly be found in one of the predicted states. This process, in which the knowledge of the state of the particle passes from certain level of ignorance to that where the system is observed in one of the predicted states, is known as the collapse of the wave function

$$\Phi(x) = \Sigma_i c_i \varphi_i(x) \longrightarrow \varphi_j(x). \tag{3.77}$$

What does this mean? If the measurement collapses the wave function, what was then the relationship between the superposition of states with the actual state of the system? The question is whether the particle, before the experiment, occupied at the same time the whole set of possible states and then jumped over into one of them when it was observed, or, on the contrary, the particle was already in the state where it was observed. Has really the measurement process influenced the resulting outcome? This kind of limitations in the theoretical prediction and the epistemological problems that emerge from these questions have been part of the debates on the completeness issue of the quantum theory. When the quantum theory is assumed complete, its predicting limitations are transferred into the nature of the quantum system, one can then suggest that the particle is such that it occupies simultaneously all of the accessible states, or, as in the one and two slits diffraction problem, the collapsed outcomes depend on how the system is observed. The correspondence between the actual system and the theoretical image gave rise to one of the most controversial articles of quantum theory: a paper published by Einstein, Podolsky and Rosen (the Einstein-Podolsky-Rosen theorem) that questions whether the quantum theory provides a complete description or not. Some paradoxes of quantum mechanics, like Schrödinger's cat and Wigner's friend, are related to this question.

On the other hand, one can also argue as follows. Suppose one has a system like the infinite quantum well with one electron inside. When we solve the Schrödinger equation, the general solution is written as

$$\Phi(x) = \Sigma_i c_i \varphi_i(x). \tag{3.78}$$

If this system is replicated N times, with $N \to \infty$, and we observe the electrons of this "ensemble" of systems, we will find that in a fraction f_1 of them (which will be approximately equal to $|c_1|^2/\langle\Phi|\Phi\rangle$), the electron is in the state φ_1, another fraction f_2 (which will also be approximately equal to $|c_2|^2/\langle\Phi|\Phi\rangle$), in the state φ_2 and so on. The relative frequency with which a state $\varphi_i(x)$ participates in the collection corresponds to $|c_i|^2$. In this case the meaning of the collapse of the wave-function changes. We do not need the particle occupying the whole set of states at the same time. But we have to accept that the wave function Φ describes not one particle but an ensemble of particles. In other words, if we say that the function Φ of (3.73) describes an ensemble of quantum wells containing one electron each, there will be a better agreement between the theory and the reality. This is called the statistical interpretation. While the Copenhagen school (with Niels Bohr in a prominent place) argues that the quantum theory describes completely a single system, the statistical interpretation school (in which A. Einstein stands out) argues that the quantum theory describes the behavior of an ensemble of quantum systems.

3.5 The Particle Current Density in Quantum Mechanics

As said before, the calculation of the energy eigenvalues and its corresponding eigenfunctions are essential for the quantum description. Without them, it is virtually impossible to assess other important physical variables. One of these physical variables of interest is the *particle current density* $\mathbf{j}$ that describes the number of particles crossing a unit area per unit time. We will see now that, to obtain the current density in the quantum mechanical formalism, it is necessary to know the eigenfunctions. In classical physics, an important equation where the current density appears is the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0. \tag{3.79}$$

Here ρ represents the density of particles. Suppose that the density ρ is given by the probability density function

$$\rho = \psi^*(x, t)\psi(x, t). \tag{3.80}$$

If we derive this function with respect to time

$$\frac{\partial \rho}{\partial t} = \psi \frac{\partial \psi^*}{\partial t} + \psi^* \frac{\partial \psi}{\partial t}, \tag{3.81}$$

and use the complete Schrödinger Eq. (2.42), we have

$$\frac{\partial \rho}{\partial t} = -\frac{i\hbar}{2m}\left(\psi \nabla^2 \psi^* - \psi^* \nabla^2 \psi\right). \tag{3.82}$$

Writing the last equation in the form

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \left[\frac{i\hbar}{2m}\left(\psi \nabla \psi^* - \psi^* \nabla \psi\right)\right] = 0, \tag{3.83}$$

we have the continuity equation, with the quantum current density given by

$$\mathbf{j} = \frac{i\hbar}{2m}\left(\psi \nabla \psi^* - \psi^* \nabla \psi\right). \tag{3.84}$$

When the wave function is factorized, i.e. when the quantum state is described by

$$\psi(\mathbf{r}, t) = \varphi(\mathbf{r}) e^{-i\omega t}, \tag{3.85}$$

the particle current density will simply be given by

$$\mathbf{j} = \frac{i\hbar}{2m}\left(\varphi \nabla \varphi^* - \varphi^* \nabla \varphi\right). \tag{3.86}$$

This is an important expression and is frequently used in the quantum transport theory. It is clear that when the function $\varphi(\mathbf{r})$ is real, the current density $\mathbf{j}$ is zero. A system with real wave functions is the infinite quantum well, therefore, the current density in each of the eigenstates is zero, and it is zero at all points of the well. This does not mean that the particles are static. They move but the flux of particles to the right is the same as to the left. In the case of free particles, the wave functions are complex. The combination is real when the amplitudes of the right and left moving waves are equal. Let us assume that we have a bunch of free particles moving towards the right, described by the function $\varphi^{+}(x) = Ae^{ikx}$ and let us leave, for the moment, the normalization constant undefined. The particle current density in this state will be

$$\begin{aligned}
\mathbf{j}^{+}(x) &= \frac{i\hbar}{2m}\left(\varphi^{+}(x)\frac{d}{dx}\,\varphi^{+*}(x) - \varphi^{+*}(x)\frac{d}{dx}\,\varphi^{+}(x)\right)\widehat{\mathbf{x}},\\
&= \frac{i\hbar}{2m}\left(\varphi^{+}(x)(-ik)\,A^{*}e^{-ikx} - \varphi^{+*}(x)(ik)\,Ae^{ikx}\right)\widehat{\mathbf{x}},\\
\mathbf{j}^{+}(x) &= \frac{\hbar\mathbf{k}}{m}\rho^{+}(x).
\end{aligned} \tag{3.87}$$

If we remember the relation $\hbar\mathbf{k} = \mathbf{p}$, it is evident that the density of free particles (moving to the right) is similar to the classical expression of particle current density $\mathbf{j} = \mathbf{v}\rho$ with $\mathbf{v} = \hbar\mathbf{k}/m$. Since $\rho^{+}(x) = |Ae^{ikx}|^2$, the current density of free particles (towards the right) is given by the expression

$$\mathbf{j}^{+}(x) = \frac{|A|^2\hbar\mathbf{k}}{m}. \tag{3.88}$$

The constant $|A|^2$ is a scale factor that depends on the normalization chosen for the wave functions. For the Born and Dirac normalization we have

$$\mathbf{j}^{+}(x) = \frac{|A|^2\hbar\mathbf{k}}{m} = \begin{cases} \hbar\mathbf{k}/Lm, & \text{with} \quad A = 1/\sqrt{L};\\ \mathbf{k}/2\pi m, & \text{with} \quad A = 1/\sqrt{2\pi\hbar}\,. \end{cases} \tag{3.89}$$

As mentioned before, the normalization constant can also be chosen such that

$$\mathbf{j}^{+}(x) = \frac{|A|^2\hbar\mathbf{k}}{m} = 1. \tag{3.90}$$

In that case, we have

$$A = \sqrt{m/\hbar k}. \tag{3.91}$$

Hence

$$\varphi^{\pm}(x) = \frac{1}{\sqrt{\hbar k/m}}e^{\pm ikx}. \tag{3.92}$$

The physical meaning that the normalization constant has in this case is rather clear. For some calculations knowing the normalization constant is fundamental.

The current density is an important quantity in quantum transport and will be very useful in the next chapter, where transmission and reflection coefficients will be obtained.

3.6 Dirac's Notation and Some Useful Relations

At various points throughout this text, we will use for compactness and simplicity reasons the bra-ket Dirac notation. This notation is flexible and consistent. We will show some examples. To visualize the flexible use of the Dirac notation, we will show on the right hand side, in bra-ket notation, some expressions written at the left hand side in the standard notation. In some cases we will indicate in a second row, the same expression, but using an alternative bra-ket notation form. The bras $\langle f|$ and the kets $|g\rangle$ are quantities with vector properties and are often referred to as state vectors.

For continuous variable functions we have:

standard form		bra-ket notation	simplified notation
$f(x)$	$\Longleftrightarrow$	$\langle x \vert f\rangle$	$\vert f\rangle$
$\varphi(\mathbf{r})$	$\Longleftrightarrow$	$\langle \mathbf{r} \vert \varphi\rangle$	$\vert \varphi\rangle$
$\varphi(\mathbf{p})$	$\Longleftrightarrow$	$\langle \mathbf{p} \vert \varphi\rangle$	$\vert \varphi\rangle$

In the last row the function is in the momentum representation. Generally, one works in the coordinates representation and the simplified notation is used without any confusion, however if any confusion possibility exists, the specific notation must be used. For the wave function of a free particle $\varphi_p(x)$ and for the eigenfunctions $\varphi_n(x)$, we will use the following notations

standard form		bra-ket notation	simplified notation
$\varphi_p(x)$	$\Longleftrightarrow$	$\langle x \vert \varphi_p\rangle$ $\langle x \vert p\rangle$	$\vert \varphi_p\rangle$
$\varphi_p^*(x)$	$\Longleftrightarrow$	$\langle \varphi_p \vert x\rangle$ $\langle p \vert x\rangle$	$\langle \varphi_p \vert$
$\varphi_n(x)$	$\Longleftrightarrow$	$\langle x \vert \varphi_n\rangle$ $\langle x \vert n\rangle$	$\vert \varphi_n\rangle$ $\vert n\rangle$
$\widehat{H}\varphi_n(x) = E_n\varphi_n(x)$	$\Longleftrightarrow$	$\widehat{H}\vert n\rangle = E_n\vert n\rangle$	

3.6.1 General Properties of Bras and Kets

1. The bras and kets not always represent vectors. When they represent scalar functions and an internal product $\langle\psi|\phi\rangle$ is evaluated, the first factor (the bra $\langle\psi|$) is the complex conjugate of the ket $|\psi\rangle$. When they represent vectors, the bra is the transpose conjugate of the ket . The operation that transposes and conjugates, is usually denoted with the symbol (†). Therefore, we have

$$\begin{array}{ll} |\psi\rangle^* = \langle\psi| & \langle\varphi|\psi\rangle^* = \langle\psi|\varphi\rangle \\ |\psi\rangle^\dagger = \langle\psi| & \langle\varphi|\psi\rangle^\dagger = \langle\psi|\varphi\rangle. \end{array} \tag{3.93}$$

2. The action of an operator $\widehat{Q}$ on a ket $|\psi\rangle$ is another ket, i.e.:

$$\widehat{Q}|\psi\rangle = |\phi\rangle. \tag{3.94}$$

3. The product $|\varphi\rangle\langle\varphi|$ is an operator, the projection operator $\widehat{P}$. When this operator acts on a ket, say on the ket $|\varphi_n\rangle$, it produces a new ket equal to $|\varphi\rangle$ multiplied by the internal product $\langle\varphi|\varphi_n\rangle$. In fact:

$$\widehat{P}|\varphi_n\rangle = |\varphi\rangle\langle\varphi|\varphi_n\rangle = \langle\varphi|\varphi_n\rangle|\varphi\rangle. \tag{3.95}$$

4. If $|\phi\rangle$ and $|\varphi\rangle$ are two arbitrary kets and a a complex number, we have

$$\langle\phi|a\varphi\rangle = a\langle\phi|\varphi\rangle; \qquad \langle a\phi|\varphi\rangle = a^*\langle\phi|\varphi\rangle. \tag{3.96}$$

5. The linear combination of kets and bras is another ket or bra and the following properties are fulfilled

$$\langle\phi|\big(|\varphi_1\rangle + a|\varphi_2\rangle\big) = \langle\phi|\varphi_1\rangle + a\langle\phi|\varphi_2\rangle; \tag{3.97}$$

$$\big(\langle a\phi_1| + b\langle\phi_2|\big)|\varphi\rangle = a^*\langle\phi_1|\varphi\rangle + b\langle\phi_2|\varphi\rangle. \tag{3.98}$$

6. Given an operator $\widehat{P}$, we have

$$\langle\alpha|\widehat{P}^\dagger|\beta\rangle = \big(\langle\beta|\widehat{P}|\alpha\rangle\big)^* = \big(\langle\beta|\widehat{P}|\alpha\rangle\big)^\dagger. \tag{3.99}$$

3.6.2 Some Useful Relations

To illustrate the use of the Dirac notation, let us present, in terms of bra-kets, some derivations that we know in the standard notation.

1. *The orthogonality condition*
   a. One alternative is writing

$$\begin{aligned}\int \varphi_{n'}^*(x)\varphi_n(x)dx &\Longleftrightarrow \int \langle\varphi_{n'}|x\rangle\langle x|\varphi_n\rangle dx\\ &= \langle\varphi_{n'}| \int |x\rangle\langle x|dx \ \ |\varphi_n\rangle\\ &\Longleftrightarrow \langle\varphi_{n'}|\varphi_n\rangle = \delta_{n'n}.\end{aligned} \tag{3.100}$$

In this expression we have used the unit operator

$$\int |x\rangle\langle x|dx = 1. \tag{3.101}$$

b. Another is

$$\begin{aligned}\int \varphi_{n'}^*(x)\varphi_n(x)dx &\Longleftrightarrow \int \langle n'|x\rangle\langle x|n\rangle dx\\ &= \langle n'| \int |x\rangle\langle x|dx \ \ |n\rangle\\ &\Longleftrightarrow \langle n'|n\rangle = \delta_{n'n}.\end{aligned} \tag{3.102}$$

2. *The closure property*
   If the unit operator of Eq. (3.101) is multiplied from the right by $|x'\rangle$, we have

$$\int |x\rangle\langle x|x'\rangle dx = |x'\rangle. \tag{3.103}$$

   This means that

$$\langle x|x'\rangle = \delta(x - x'). \tag{3.104}$$

3. *The development of the function $\psi(x)$ in a basis and the completeness of the basis*

   If we have a function $\psi(x)$ and the basis functions $\varphi_n(x)$, the development of $\psi(x)$ in either notation is

$$\psi(x) = \sum_n c_n\varphi_n(x) \Longleftrightarrow |\psi\rangle = \sum_n c_n\langle x|n\rangle. \tag{3.105}$$

   Let us now use Dirac's notation to obtain the coefficients c_n and to deduce the completeness of the basis.

   a. If we multiply the function $\psi(x)$ from the left with $\varphi_n^*(x)$ and integrate, we have

$$\begin{aligned}\int \varphi_n^*(x)\psi(x)dx \Longleftrightarrow \langle n|\psi\rangle &= \sum_{n'} c_{n'}\langle n| \int |x\rangle\langle x|dx|n'\rangle \\ &= \sum_{n'} c_{n'}\langle n|n'\rangle; \\ \langle n|\psi\rangle &= \sum_{n'} c_{n'}\delta_{n'n}; \end{aligned} \tag{3.106}$$

which gives us, after summing, the coefficient

$$c_n = \langle n|\psi\rangle. \tag{3.107}$$

b. When we substitute these coefficients in (3.105) we have

$$|\psi\rangle = \sum_n \langle n|\psi\rangle\ |n\rangle = \sum_n |n\rangle\langle n|\psi\rangle. \tag{3.108}$$

The factor that multiplies $|\psi\rangle$, in the last term, must be the unity, therefore

$$\sum_n |n\rangle\langle n| = 1. \tag{3.109}$$

This equation expresses the completeness of the basis $|n\rangle$.

3.6.3 Momentum Representation

With the Dirac notation it is easy to visualize a change of representation, for example from the coordinate to the momentum representation. This change of representations is done through the Fourier transformation

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} f(x)dx; \quad f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} \tilde{f}(k)dk. \tag{3.110}$$

Taking into account that $k = p/\hbar$, it is possible to rewrite these relations in the form

$$\begin{aligned} \tilde{f}(p) &= \tfrac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{ipx/\hbar} f(x)dx; \\ f(x) &= \tfrac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} e^{-ipx/\hbar} \tilde{f}(p)dp. \end{aligned} \tag{3.111}$$

Using Dirac's notation, with $\langle p|\varphi\rangle$, for the function $\varphi(p)$ defined in the p-space, and $\langle x|\varphi\rangle$, for the function $\varphi(x)$ defined in the x-space, the Fourier transformation can be written as:

$$\langle p|\varphi\rangle = \int \langle p|x\rangle\langle x|\varphi\rangle dx; \quad \langle x|\varphi\rangle = \int \langle x|p\rangle\langle p|\varphi\rangle dp; \tag{3.112}$$

with

$$\langle p|x\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar} \quad \text{and} \quad \langle x|p\rangle = \langle p|x\rangle^\dagger = \frac{1}{\sqrt{2\pi\hbar}}e^{-ipx/\hbar}, \tag{3.113}$$

the kernels of the corresponding transformations.

Let us see now a simple example. Suppose we have the free particle function

$$\varphi(x) = \langle x|\varphi\rangle = e^{-ikx}/\sqrt{2\pi\hbar} \tag{3.114}$$

and we want to determine its transformed function $\langle p|\varphi\rangle$. The function $\varphi(x)$ can be rewritten in the form

$$\varphi_{p'}(x) = e^{-ip'x/\hbar}/\sqrt{2\pi\hbar}. \tag{3.115}$$

Notice that the p' of this function, not necessarily is equal to p in the "transformation kernel" $\langle p|x\rangle$, in the first equation of (3.112). If we substitute, and take into account the second equality of (3.113), we have

$$\langle p|\varphi\rangle = \frac{1}{\sqrt{2\pi\hbar}}\int \langle p|x\rangle e^{-ip'x/\hbar}dx = \int \langle p|x\rangle\langle x|p'\rangle dx. \tag{3.116}$$

Rearranging the bras and kets in the last equality, to highlight the unit operator in the coordinate space, we have

$$\langle p|\int |x\rangle\langle x|dx|p'\rangle = \langle p|p'\rangle = \delta(p-p'). \tag{3.117}$$

Therefore

$$\varphi(p) = \langle p|\varphi\rangle = \delta(p-p'). \tag{3.118}$$

This means that a free particle, in the momentum space, is a definite momentum pulse, whereas in the coordinate space is an extended function from $-\infty$ to $+\infty$. This will be better understood with the Heisenberg uncertainty relation that will be seen later. We postpone to Chap. 7 the discussion of eigenvalues and operators in the momentum space.

3.7 Solved Problems

Exercise 6 Show that a bounded one-dimensional system, described by the 1D Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2\varphi(x)}{dx^2}+V(x)\varphi(x)=E\varphi(x), \tag{3.119}$$

is a non degenerate system.

Solution Suppose that $\varphi_1(x)$ and $\varphi_2(x)$ are two independent solutions of a degenerate system, i.e. solutions with the same eigenvalue. If $E_1 = E_2 = E$, the following equations should be satisfied

$$-\frac{\hbar^2}{2m}\frac{d^2\varphi_1(x)}{dx^2}+V(x)\varphi_1(x)=E\varphi_1(x), \tag{3.120}$$

and

$$-\frac{\hbar^2}{2m}\frac{d^2\varphi_2(x)}{dx^2}+V(x)\varphi_2(x)=E\varphi_2(x). \tag{3.121}$$

Multiplying the first of these equations by $\varphi_2(x)$, the second by $\varphi_1(x)$, and subtracting, we have

$$\varphi_2(x)\frac{d^2\varphi_1(x)}{dx^2}-\varphi_1(x)\frac{d^2\varphi_2(x)}{dx^2}=0, \tag{3.122}$$

that can be written in the form

$$\frac{d}{dx}\left(\varphi_2(x)\frac{d\varphi_1(x)}{dx}-\varphi_1(x)\frac{d\varphi_2(x)}{dx}\right)=0. \tag{3.123}$$

This means that at any point x the function in the parenthesis is constant, i.e.

$$\varphi_2(x)\frac{d\varphi_1(x)}{dx}-\varphi_1(x)\frac{d\varphi_2(x)}{dx}=\text{constant}. \tag{3.124}$$

We will show that the constant must be zero. If we find a point at which the first member is zero, we will conclude that the constant will be zero everywhere, since it is a constant.

We know that in a bounded system, the wave functions and their derivatives evaluated at $x = \infty$ must vanish, i.e.

$$\left.\varphi_2(x)\frac{d\varphi_1(x)}{dx}\right|_{x=\infty}=0, \tag{3.125}$$

hence

$$\varphi_2(x)\frac{d\varphi_1(x)}{dx} = \varphi_1(x)\frac{d\varphi_2(x)}{dx}, \tag{3.126}$$

which is satisfied if

$$\varphi_1(x) = c\ \varphi_2(x). \tag{3.127}$$

where c is an arbitrary constant. We conclude that $\varphi_1(x)$ and $\varphi_2(x)$ are essentially the same function. This proofs that there is no degenerate solutions in 1D bounded systems.

Exercise 7 Show that the eigenfunctions of a bounded one-dimensional system are real.

Solution This problem complements the previous one on non-degenerate systems. The demonstration will use some relations found before. If we substitute $\varphi_2(x)$ by $\varphi_1^*(x)$ in (3.123), and assume that the potential $V(x)$ is real, we have

$$\varphi_1^*(x)\frac{d\varphi_1(x)}{dx} = \varphi_1(x)\frac{d\varphi_1^*(x)}{dx}. \tag{3.128}$$

A relation that is fulfilled if $\varphi_1(x)$ is real.

Exercise 8 If a (free particle) wave function is written as

$$\varphi(x) = ae^{ikx} + be^{-ikx}, \tag{3.129}$$

determine the particle current density.

Solution In Sect. 3.5, we defined the current density

$$\mathbf{j}(x) = \frac{i\hbar}{2m}\left(\varphi\frac{d}{dx}\varphi^* - \varphi^*\frac{d}{dx}\varphi\right)\widehat{\mathbf{x}}. \tag{3.130}$$

Using the function $\varphi(x)$ of (3.129), we have

$$\begin{aligned}\mathbf{j}(x) = \frac{i\hbar}{2m}\Big(&\big(a\,e^{ikx} + b\,e^{-ikx}\big)\big(-ika^*\,e^{-ikx} + ikb^*\,e^{ikx}\big)\\ &-\big(a^*\,e^{-ikx} + b^*\,e^{ikx}\big)\big(ika\,e^{ikx} - ikb\,e^{-ikx}\big)\Big)\widehat{\mathbf{x}},\end{aligned} \tag{3.131}$$

which, after some reductions becomes

$$\begin{aligned}\mathbf{j}(x) = \frac{k\hbar}{2m}\Big(&\big(ae^{ikx} + b\,e^{-ikx}\big)\big(a^*e^{-ikx} - b^*e^{ikx}\big)\\ &-\big(a^*e^{-ikx} + b^*e^{ikx}\big)\big(-a\,e^{ikx} + be^{-ikx}\big)\Big)\widehat{\mathbf{x}}.\end{aligned} \tag{3.132}$$

It is easy to verify that this current can be written in the form

$$\mathbf{j}(x) = \frac{k\hbar}{2m}\Big(\big(|a|^2 - |b|^2 - 2\Re\mathrm{e}\, ab^* e^{2ikx}\big) \\ -\big(-|a|^2 + |b|^2 - 2\Re\mathrm{e}\, ab^* e^{2ikx}\big)\Big)\,\widehat{\mathbf{x}}, \tag{3.133}$$

which reduces to

$$\mathbf{j}(x) = \frac{k\hbar}{m}\big(|a|^2 - |b|^2\big)\,\widehat{\mathbf{x}}. \tag{3.134}$$

This is a very important result that basically says that when the wave function has two components, describing particles moving to the right and to the left, as in (3.129), the current density at any point x, is just the difference of the respective currents.

Exercise 9 In the last example, we saw that when the wave function is given by the combination $\varphi(x) = ae^{ikx} + be^{-ikx}$, the current density is proportional to $|a|^2 - |b|^2$. It is evident that for $a = \pm b$ the current density vanishes ($\mathbf{j}(x) = 0$). Based on this example, show that: a) the current is zero when the wave function is real or pure imaginary and b) the particle current moving to the right $\mathbf{j}^+(x)$ and the particle current moving to the left $\mathbf{j}^-(x)$ are equal in magnitude but with opposite signs.

Solution When $a = b$, the wave function $\varphi(x)$ becomes

$$\varphi(x) = a\left(e^{ikx} + e^{-ikx}\right) = 2a\cos kx, \tag{3.135}$$

whereas if $b = -a$, the wave function $\varphi(x)$ becomes

$$\varphi(x) = a\left(e^{ikx} - e^{-ikx}\right) = 2ia\sin kx. \tag{3.136}$$

In both cases, $\mathbf{j}(x) = 0$. We will calculate now the current associated to each component of the wave function when $b = \pm a$. To avoid confusion we will keep the coefficient b and replace it by $\pm a$ at the end. The currents associated to the components $\varphi^+(x) = a\,e^{ikx}$ and $\varphi^-(x) = be^{-ikx}$ are respectively

$$\mathbf{j}^+(x) = \frac{i\hbar}{2m}\left(\varphi^+(x)\frac{d}{dx}\varphi^{+*}(x) - \varphi^{+*}(x)\frac{d}{dx}\varphi^+(x)\right)\widehat{\mathbf{x}}; \tag{3.137}$$

and

$$\mathbf{j}^-(x) = \frac{i\hbar}{2m}\left(\varphi^-(x)\frac{d}{dx}\varphi^{-*}(x) - \varphi^{-*}(x)\frac{d}{dx}\varphi^-(x)\right)\widehat{\mathbf{x}}. \tag{3.138}$$

Replacing $\varphi^{+*}(x)$ and $\varphi^{-*}(x)$ we have

$$\begin{aligned}\mathbf{j}^{+}(x) &= \frac{i\hbar}{2m}\left(ae^{ikx}(-ik)\,a^{*}e^{-ikx} - a^{*}e^{-ikx}(x)(ik)ae^{ikx}\right)\widehat{\mathbf{x}},\\ &= \frac{\hbar\mathbf{k}}{m}|a|^{2},\end{aligned} \tag{3.139}$$

and

$$\begin{aligned}\mathbf{j}^{-}(x) &= \frac{i\hbar}{2m}\left(be^{-ikx}(ik)\,b^{*}e^{ikx} - b^{*}e^{ikx}(x)(-ik)\,be^{-ikx}\right)\widehat{\mathbf{x}},\\ &= -\frac{\hbar\mathbf{k}}{m}|b|^{2}.\end{aligned} \tag{3.140}$$

These are the two currents contained in the expression of current density that appears in (3.134). It is clear that when $b = \pm a$, the currents satisfy the relation $\mathbf{j}^{-}(x) = -\mathbf{j}^{+}(x)$, regardless of whether b is equal to $+a$ or $-a$.

3.8 Problems

1. Show that normalizing the wave function

$$\varphi(x) = a'e^{ikx} + b'e^{-ikx}, \tag{3.141}$$

one obtains

$$\varphi(x) = aAe^{ikx} + b\,Ae^{-ikx}, \tag{3.142}$$

with $|a|^2+|b|^2 = 1$ and the constant $A = 1/\sqrt{L}$, for the normalization procedure *a la Born*, or $A = 1/\sqrt{2\pi h}$, for the normalization procedure *a la Dirac*.

2. For the infinite quantum well in Fig. 3.5, with origin of the x-axis at the center of the well, obtain the first four eigenfunctions. Discuss the symmetries of the eigenfunctions and the Schrödinger equation under the parity transformation

$$\widehat{\mathbf{P}} : x \mapsto -x. \tag{3.143}$$

3. Show that the function

$$\varphi_E(x) = \frac{1}{\sqrt{\pi\hbar^2 k/m}}e^{ikx} \tag{3.144}$$

satisfies the condition

$$\langle\varphi_{E'}|\varphi_E\rangle = \int_{-\infty}^{\infty} \varphi^*_{E'}(x)\varphi_E(x)dx = \delta(E - E'),$$

with $E = \hbar^2k^2/2m$, the free particle energy.

4. Show that the wave functions in Eq. (3.68) are certainly the normalized free particle eigenfunctions of the infinite quantum well, whose width is L.
5. If the particle is an electron in an infinite quantum well, determine the eigenvalues of the first four levels when $L = 10$ nm and also when $L = 20$ nm and $L = 5$ nm. What is the effect of increasing the width of the well and what of decreasing?
6. Draw the first four eigenfunctions for the well widths of the previous problem.
7. Show that the particle current density at any level of the infinite quantum well is zero.
8. Determine the free particle current density if they are described by

$$\varphi(x) = a\frac{e^{ikx}}{\sqrt{\hbar k/m}} + b\frac{e^{-ikx}}{\sqrt{\hbar k/m}}.$$

9. If we have initially an electron in the fourth energy level of an infinite quantum well and it makes a transition to the first level, determine the width L of the well so that the emitted photon has wavelength $\lambda \approx 500$ nm, and what will be the widths if the photons have wavelengths $\lambda \approx 600$ nm and $\lambda \approx 700$ nm? To which colors do these wavelengths correspond?

Chapter 4
The Tunneling Effect and Transport Properties

In this chapter we will solve the Schrödinger equation for simple one-dimensional examples. We will show that these systems exhibit important quantum properties like the energy quantization and the tunneling effect. The one-dimensional examples that will be considered here are: the step potential; the finite rectangular quantum well and the rectangular potential barrier. In these systems the potential functions are piecewise constant, with abrupt discontinuities at two or three points. For many years this kind of potentials were systems of academic interest and used to model real systems, where the mathematical procedures required more involved calculations. In the current nano-structure physics, the rectangular quantum wells and the rectangular potential barriers are not any more systems of purely academic interest. The actual semiconductor structures, contain barriers and wells (grown with atomic layer precision), and the potential profiles "felt" by the conduction electrons and valence band holes in LEDs,[1] quantum well lasers[2] and other optoelectronic devices, of mesoscopic[3] or nanoscospic[4] sizes, are rather similar to those considered here and the following chapter.

As mentioned before, the quantum wells confine. Hence, the energy of a quantum particle in a quantum well will quantize. In this case, we will be interested in the energy eigenvalues and the corresponding eigenfunctions. The step potential and the potential barrier imply open systems, with only one classical turning point. In these systems the quantum particles, approaching from the left or from the right, will get reflected and eventually, depending on the energy, they will get transmitted. In open systems, we will be more interested in the transport properties and in the wave functions describing the quantum particles from $-\infty$ to $+\infty$. We will see that for particles in a quantum well, with energies larger than the confining potential

---

[1] Acronym of light emitting diode.

[2] Acronym of light amplification by stimulated emission of radiation.

[3] Mesoscopic refers to intermediate size systems, between the macroscopic and the atomic dimensions. Therefore the systems are described by quantum theory.

[4] Systems with dimensions of the order of 10^{-9} m, i.e. of a few atomic diameters.

P. Pereyra, *Fundamentals of Quantum Physics*, Undergraduate Lecture Notes in Physics,
DOI: 10.1007/978-3-642-29378-8_4,

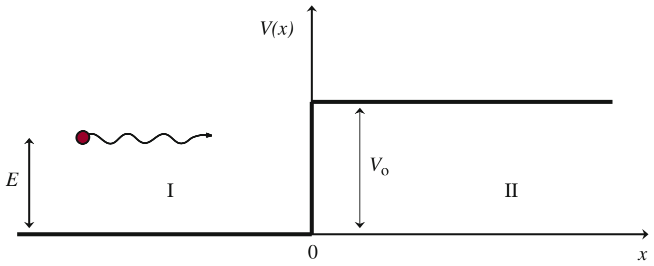


Fig. 4.1 Parameters of the step potential

depth, it will be also possible to talk of transport properties. In this chapter we will use, systematically, the transfer matrix method and we will see that, despite the physical differences among the quantum wells and the potential barriers, the transition matrices that enclose the continuity conditions, can equally well be used when we are solving their Schrödinger equations. This is one of the attractive features of the transfer matrix method.

4.1 The 1D Step Potential

Except for the constant potential, the simplest potential profile that one can imagine is the step potential shown in Fig. 4.1. This is essentially a constant potential with one discontinuity point, thus, with one classical turning point. If the discontinuity is at the origin, the potential is described by the function

$$V(x) = \begin{cases} 0, & x \leq 0\,; \\ V_o, & x > 0\,. \end{cases} \tag{4.1}$$

Electrons and holes experience a change of the potential energy like that of Eq. (4.1) when they move from one (highly doped) semiconductor to another, for example, when the conduction band electrons reach the interface of a GaAs layer with an AlAs layer, in the GaAs/AlAs structure. If we have a beam of particles[5] with energy E, approaching the potential step from the left, we can expect different results depending on whether the energy E is smaller or larger than the step potential height V_o. We will discuss these cases separately.

1. *When the particle's energy is less than* V_o, the Schrödinger equations in regions I and II are, respectively,

$$-\frac{\hbar^2}{2m}\frac{d^2\varphi_{\rm I}(x)}{dx^2} = E\,\varphi_{\rm I}(x), \qquad x \leq 0, \tag{4.2}$$

---

[5] We will systematically assume the independent particle approximation. In this approximation the interactions among the quantum particles, are neglected. We shall also assume that quantum solutions, according with the statistical interpretation, describe well the behavior of the beam.

and

$$\frac{\hbar^2}{2m}\frac{d^2\,\varphi_{\mathrm{II}}(x)}{dx^2} = (V_o - E)\,\varphi_{\mathrm{II}}(x), \qquad x > 0. \tag{4.3}$$

We will begin with Eq. (4.2). As we saw in Chap. 3 the solutions in this case are the oscillatory functions

$$\varphi_{\mathrm{I}}(x) = a_1 e^{ikx} + b_1 e^{-ikx}, \tag{4.4}$$

with $k = \sqrt{2mE/\hbar^2}$, while, the solutions of (4.3) are the exponential functions

$$\varphi_{\mathrm{II}}(x) = a_2 e^{qx} + b_2 e^{-qx}, \tag{4.5}$$

with $q = \sqrt{2m\,(V_o - E)\,/\hbar^2}$. These solutions, $\varphi_{\mathrm{I}}(x)$ and $\varphi_{\mathrm{II}}(x)$, are continuous functions in their respective domains and they are part of the solution of the Schrödinger equation for the step potential, which is defined from $-\infty$ to ∞. As mentioned before, the Schrödinger equation solutions must fulfill the finiteness and continuity requirements everywhere. Including points where the potential function is discontinuous, like the point $x = 0$ of our example. At points like this, we have to impose the continuity requirement on the wave function and on its first order derivative.

These continuity requirements are

$$\varphi_{\mathrm{I}}(0) = \varphi_{\mathrm{II}}(0) \qquad \text{and} \qquad \frac{d\varphi_{\mathrm{I}}(0)}{dx} = \frac{d\varphi_{\mathrm{II}}(0)}{dx}. \tag{4.6}$$

Using the explicit forms of $\varphi_{\mathrm{I}}(x)$ and $\varphi_{\mathrm{II}}(x)$, we have

$$a_1 + b_1 = a_2 + b_2, \tag{4.7}$$

$$ik(a_1 - b_1) = q(a_2 - b_2). \tag{4.8}$$

We will solve these equations for two coefficients out of four. The linear combinations (4.4) and (4.5) will be forced to match smoothly at $x = 0$. The remaining coefficients behave as independent variables. In this problem, the first term of $\varphi_{\mathrm{I}}(x)$ (i.e. the function $a_1 e^{ikx}$), describes the incident particles. Since the incident flux can be set at will, it is natural to choose a_1 as an independent variable. As will be seen in the oncoming examples, the number of coefficients is reduced by imposing additional conditions, as finiteness of the wave functions or specific physical conditions like setting the incoming particles only from the left or only from the right hand side. We will see, in the quantum well problem, that the energy quantization emerges from the fulfillment of the confining requirements. In the step potential problem, the finiteness problem is an issue to take care. Indeed, if we look at (4.5), we have two exponential functions and one of them diverges when $x \to \infty$. To prevent this singularity, we choose $a_2 = 0$.

But we will not make this coefficient zero right now. We will keep it, temporarily. We need first to express all coefficients on the right side of the discontinuity point in terms of those on the left. In this case, the coefficients a_2 and b_2 in terms of $a_1$and b_1. Multiplying (4.7) by q and summing or subtracting with (4.8), we obtain

$$a_2 = \frac{1}{2q}(q+ik)\,a_1 + \frac{1}{2q}(q-ik)\,b_1, \tag{4.9}$$

$$b_2 = \frac{1}{2q}(q-ik)\,a_1 + \frac{1}{2q}(q+ik)\,b_1. \tag{4.10}$$

At this point we can introduce the matrix representation and define the transfer matrix. In this problem, one can not yet perceive any advantage of using this representation. On the contrary, the student may feel that it distracts the calculation procedure. The advantages will be recognized later, dealing with slightly more complex systems. We introduce the matrix representation when we write the continuity conditions (4.9) and (4.10) in the form

$$\begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = \frac{1}{2q}\begin{pmatrix} q+ik & q-ik \\ q-ik & q+ik \end{pmatrix}\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}. \tag{4.11}$$

In this representation, the coefficients $\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}$ and $\begin{pmatrix} a_2 \\ b_2 \end{pmatrix}$ are nothing else than the state vectors

$$\phi_{\rm I}(x) = \begin{pmatrix} a_1 e^{ikx} \\ b_1 e^{-ikx} \end{pmatrix} \quad \text{and} \quad \phi_{\rm II}(x) = \begin{pmatrix} a_2 e^{qx} \\ b_2 e^{-qx} \end{pmatrix}, \tag{4.12}$$

evaluated at $x = 0 - \epsilon \equiv 0^-$ and at $x = 0 + \epsilon \equiv 0^+$, respectively. The matrix

$$\frac{1}{2q}\begin{pmatrix} q+ik & q-ik \\ q-ik & q+ik \end{pmatrix} \equiv M(0^+, 0^-), \tag{4.13}$$

connects the physics on the left side of the discontinuity point with the physics on the right hand side of this point. This matrix captures the continuity requirements at $x = 0$. In a compact notation, we write (4.11) also as

$$\phi_{\rm I}(0^+) = M(0^+, 0^-)\phi_{\rm I}(0^-). \tag{4.14}$$

The matrix $M(0^+, 0^-)$ is a specific example of a transfer matrix $M(x_2, x_1)$, that fulfills in general a relation like

$$\phi(x_2) = M(x_2, x_1)\phi(x_1) = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \phi(x_1), \tag{4.15}$$

that connects the state vectors $\phi(x_1)$ and $\phi(x_2)$ at any two points x_1 and x_2. A matrix like $M(0^+, 0^-)$, defined for two points, infinitesimally close to the discontinuity point, will be called *transition matrix*. We will see later that the same transition matrix will appear each time that the potential energy changes from V_1 to $V_2 = V_1 + V_o$. We have this kind of changes in piecewise constant potential systems like the rectangular quantum well and the potential barrier.

Let us now turn back into the step potential problem and recall that to prevent divergency at $x = \infty$ we have to make the coefficient a_2 zero. This leaves us with

$$\begin{pmatrix} 0 \\ b_2 \end{pmatrix} = \frac{1}{2q} \begin{pmatrix} q + ik & q - ik \\ q - ik & q + ik \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, \tag{4.16}$$

where a_1 is the natural candidate for playing the role of an independent variable. If we multiply the matrix times the state vector $\phi_{\mathrm{I}}(0^-) = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}$, on the right hand side, we find in the first place that

$$(q + ik)a_1 + (q - ik)b_1 = 0, \tag{4.17}$$

which gives

$$b_1 = -\frac{q + ik}{q - ik} a_1. \tag{4.18}$$

This coefficient multiplies the wave function e^{-ikx}, which describes particles moving to the left of the step, and defines the amplitude of the reflected particles. Since the factor that multiplies the coefficient a_1 in (4.18) is a complex number with magnitude 1, we can state that, for energies E less than V_o, all particles moving towards the step potential get reflected. If that is the case, what do we have on the right side? To understand fully the physics of the step potential, for energies E less than V_o, we still need to determine b_2. Multiplying the second row of the transition matrix with the state vector $\phi_{\mathrm{I}}(0^-)$, we have

$$b_2 = \frac{1}{2q} \Big((q - ik)a_1 + (q + ik)b_1 \Big). \tag{4.19}$$

Using here the coefficient b_1 from (4.18), we obtain

$$b_2 = \frac{1}{2q} \left((q - ik)a_1 - (q + ik)\frac{q + ik}{q - ik} a_1 \right), \tag{4.20}$$

that can also be written as

$$b_2 = -\frac{2ik}{q - ik} a_1 = -\frac{2ik}{q^2 + k^2} (q + ik) a_1. \tag{4.21}$$

What are the implications of these results? If we use b_1 in the wave function of Eq. (4.4), we have

$$\varphi_{\mathrm{I}}(x) = a_1 e^{ikx} - \frac{q + ik}{q - ik} a_1 e^{-ikx} \tag{4.22}$$

Since the amplitude of the complex number $(q + ik)/(q - ik)$ is 1, it can be written as

$$\frac{q + ik}{q - ik} = e^{i2\theta}, \qquad \text{with} \qquad \theta = \tan^{-1} \frac{k}{q}, \tag{4.23}$$

the wave function $\varphi_{\mathrm{I}}(x)$, on the left side of the step potential, becomes

$$\varphi_{\mathrm{I}}(x) = a_1 e^{ikx} - a_1 e^{i2\theta} e^{-ikx}. \tag{4.24}$$

This function can also be written as follows

$$\varphi_{\mathrm{I}}(x) = a_1 e^{i\theta} \left(e^{-i\theta} e^{ikx} - e^{i\theta} e^{-ikx} \right). \tag{4.25}$$

If we define a new constant, say $a = -2ia_1 e^{i\theta}$, the solution in region I takes the form

$$\varphi_{\mathrm{I}}(x) = -a \sin(kx - \theta). \tag{4.26}$$

This is a stationary function whose current density is zero (see the illustrative problem 3.2). This shows that for $E < V_o$ the current density towards the right is equal to the current density towards the left. In this system the whole flux that reaches the step potential gets reflected. What will we have then for the wave function $\varphi_{\mathrm{II}}(x)$, defined for $x > 0$? In this region $a_2 = 0$ and b_2 can be written as

$$b_2 = -\frac{2ik}{\sqrt{q^2 + k^2}} e^{i\theta} a_1. \tag{4.27}$$

If we use the constants defined before, we have

$$\varphi_{\mathrm{II}}(x) = a \frac{k}{\sqrt{q^2 + k^2}} e^{-qx}. \tag{4.28}$$

This is a function whose amplitude decreases exponentially with x, as shown in Fig. 4.2 (note that for $x >> 0$ the flat line has zero amplitude). The oscillating function matches perfectly with the decaying exponential function at $x = 0$.

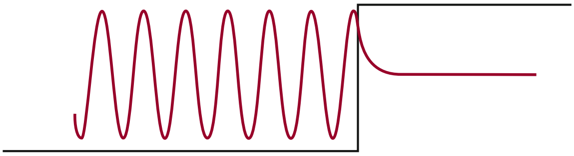


Fig. 4.2 The wave function at the *left* and *right* sides of the discontinuity point, for energies $E < V_o$

According to (4.28) the probability $|\varphi_{\rm II}(x)|^2$ of finding the particle at a point x within the step potential is proportional to e^{-2qx} and *is different from zero*. A result like this, for particles with mass, is impossible to meet in the classical physics for particles with mass. The penetration depth in the step potential is proportional to $1/q \approx \hbar/\sqrt{2m\big(V_o - E\big)}$ and becomes infinite when $E = V_o$. To conclude this part, it is important to emphasize that in the step potential problem there is no energy quantization (the potential has a single return point) and the solutions are propagating wave functions in region I, and evanescent (exponentially decreasing) functions in region II. Let us now consider the other case.

2. *When* $E > V_o$, the differential equations on the left and the right hand sides of the discontinuity point are similar and both have propagating wave solutions. For region I we have exactly the same differential equation as before, but since the differential equation on the right hand side changes, we will have different coefficients, thus, different physics. The differential equation for $x > 0$ is now

$$-\frac{\hbar^2}{2m}\frac{d^2\varphi_{\rm II}(x)}{dx^2} = \big(E - V_o\big)\varphi_{\rm II}(x), \qquad \text{for} \quad x > 0, \tag{4.29}$$

with $(E - V_o) > 0$. Our solutions are now

$$\varphi_{\rm I}(x) = a_1 e^{ik_1 x} + b_1 e^{-ik_1 x}, \tag{4.30}$$

and

$$\varphi_{\rm II}(x) = a_2 e^{ik_2 x} + b_2 e^{-ik_2 x}, \tag{4.31}$$

with $k_1 = \sqrt{2mE/\hbar^2}$ and $k_2 = \sqrt{2m\big(E - V_o\big)/\hbar^2}$. With these solutions there is no divergency risk and it makes no sense to impose a finiteness condition. It is clear that, in principle, we can have particles approaching the discontinuity point from the left or from the right, and their flux amplitudes depend on the coefficients a_1 and b_2. Without loss of generality, we can assume that particles come towards the discontinuity point only from the left. If this is the case, it is not physically possible to have, on the right hand side, particles moving towards the left. Therefore, the second term in (4.31) must vanish. This means that b_2 should be zero. Before making this coefficient zero, we will impose the continuity conditions

$$\varphi_{\rm I}(0) = \varphi_{\rm II}(0) \qquad \text{and} \qquad \varphi'_{\rm I}(0) = \varphi'_{\rm II}(0), \tag{4.32}$$

that lead us to

$$a_1 + b_1 = a_2 + b_2, \tag{4.33}$$

$$ik_1(a_1 - b_1) = ik_2(a_2 - b_2). \tag{4.34}$$

We write these equations, in the matrix representation, as

$$\begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = \frac{1}{2k_2} \begin{pmatrix} k_2 + k_1 & k_2 - k_1 \\ k_2 - k_1 & k_2 + k_1 \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}. \tag{4.35}$$

Here the transition matrix differs slightly from the transition matrix in (4.11). The wave number q, there, is here replaced by ik_2. If we now make $b_2 = 0$, we have

$$\begin{pmatrix} a_2 \\ 0 \end{pmatrix} = \frac{1}{2k_2} \begin{pmatrix} k_2 + k_1 & k_2 - k_1 \\ k_2 - k_1 & k_2 + k_1 \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}. \tag{4.36}$$

This leads us, after multiplying the matrix rows with the state vector, to

$$\frac{1}{2k_2}\left(k_2 - k_1\right) a_1 + \frac{1}{2k_2}\left(k_2 + k_1\right) b_1 = 0, \tag{4.37}$$

that allows us to obtain the coefficient

$$b_1 = \frac{k_1 - k_2}{k_1 + k_2} a_1 < a_1, \tag{4.38}$$

and, on the other hand, to

$$a_2 = \frac{1}{2}\left(1 + \frac{k_1}{k_2}\right) a_1 + \frac{1}{2}\left(1 - \frac{k_1}{k_2}\right) b_1, \tag{4.39}$$

that, using b_1 from (4.38), becomes

$$a_2 = \frac{1}{2}\left(1 + \frac{k_1}{k_2}\right) a_1 - \frac{1}{2}\left(1 - \frac{k_1}{k_2}\right) \frac{k_2 - k_1}{k_2 + k_1} a_1. \tag{4.40}$$

This equation can be simplified, and takes the form

$$a_2 = \frac{2k_1}{k_1 + k_2} a_1. \tag{4.41}$$

In this way, given b_1 and a_2, we can finally write the solutions in regions I and II as

$$\varphi_{\mathrm{I}}(x) = a_1 e^{ik_1 x} + a_1 \frac{k_1 - k_2}{k_1 + k_2} e^{-ik_1 x}, \tag{4.42}$$

and as

$$\varphi_{\mathrm{II}}(x) = a_1 \frac{2k_1}{k_1 + k_2} e^{ik_2 x}, \tag{4.43}$$

respectively. These functions describe all particles, the incident, reflected and transmitted ones. The first term of $\varphi_{\mathrm{I}}(x)$ describes the incoming particles, the second term describes the fraction of the particles that gets reflected by the step potential, and the function $\varphi_{\mathrm{II}}(x)$ describes the fraction of particles that gets transmitted. Using these functions, one can easily evaluate the incident, reflected and transmitted currents

$$\begin{aligned}
\mathbf{j}_{\mathrm{inc}} &= \frac{\hbar k_1 |a_1|^2}{m} \widehat{\mathbf{x}}, \\
\mathbf{j}_{\mathrm{refl}} &= -\frac{\hbar k_1 |a_1|^2}{m} \left(\frac{k_1 - k_2}{k_1 + k_2}\right)^2 \widehat{\mathbf{x}}, \\
\mathbf{j}_{\mathrm{trans}} &= \frac{\hbar k_2 |a_1|^2}{m} \left(\frac{2k_1}{k_1 + k_2}\right)^2 \widehat{\mathbf{x}}.
\end{aligned} \tag{4.44}$$

Notice that the reflected and transmitted currents are proportional to the incident current $\hbar k_1 |a_1|^2/m$. We can also observe from these results that when the height of the potential step $V_o \to 0$ (which means $k_1 \simeq k_2$) the reflected current tends to 0, and the transmitted current becomes equal to the incident current. We found here that, for energies larger than the step potential height, the wave functions at the left and right hand sides of the discontinuity point are propagating functions. To understand the meaning of these solutions, it is convenient to introduce the *reflection and the transmission coefficients*. One way to define these coefficients is as follows:

$$R = \frac{|\mathbf{j}_{\mathrm{refl}}|}{|\mathbf{j}_{\mathrm{inc}}|} = \frac{(k_1 - k_2)^2}{(k_1 + k_2)^2} \quad \text{and} \quad T = \frac{|\mathbf{j}_{\mathrm{trans}}|}{|\mathbf{j}_{\mathrm{inc}}|} = \frac{4k_1 k_2}{(k_1 + k_2)^2}. \tag{4.45}$$

These coefficients fulfill an important property: the *flux conservation principle*. This principle states that, when the flux of particles is conserved, we must have

$$R + T = 1. \tag{4.46}$$

It is easy to verify that this requirement is fulfilled by the reflection and transmission coefficients, R and T, defined in (4.45).

The transmission and reflection coefficients are important quantities in the analysis of transport properties through potential regions. These coefficients are closely related to the reflection and the transmission amplitudes (r and t) of scattering theory. To introduce the basic concepts of the scattering theory and to understand the

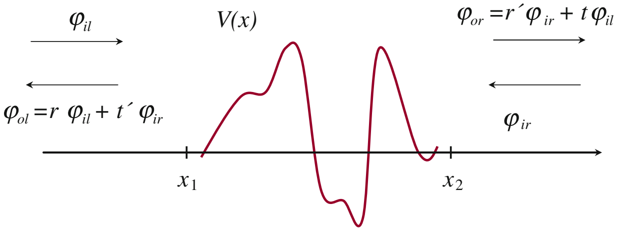


Fig. 4.3 The incident and the scattered (reflected and transmitted) waves by a potential function $V(x)$

relation between the "scattering" amplitudes and the transfer matrix elements, we will make, in the next section, a brief digression.

4.2 Scattering Amplitudes and the Transfer Matrix

An alternative description of the transport phenomenon is possible using the physical quantities and concepts of the scattering theory. In this approach, the physics of particles moving across a potential region $V(x)$ is visualized as a scattering process with incoming and outgoing particles (towards and from) the scattering region. The wave functions describing outgoing particles are related, through a scattering matrix S, with those describing incoming particles. If we have the scattering process shown in Fig. 4.3, the incoming particles, from the left and right sides, are described by $\varphi_{il}(x_1)$ and by $\varphi_{ir}(x_2)$, and the outgoing particles by $\varphi_{ol}(x_1)$ and $\varphi_{or}(x_2)$. Because of the interaction potential $V(x)$, part of the incoming particles are transmitted and part reflected. Using r and t to represent the reflected and transmitted amplitudes, when particles come from the left, and r' and t', when particles come from the right, the outgoing functions on the left and right sides can be written as

$$\varphi_{ol} = r\,\varphi_{il} + t'\varphi_{ir}, \tag{4.47}$$

$$\varphi_{or} = t\,\varphi_{il} + r'\varphi_{ir}. \tag{4.48}$$

In a matrix representation we have

$$\begin{pmatrix} \varphi_{ol}(x_1) \\ \varphi_{or}(x_2) \end{pmatrix} = \begin{pmatrix} r & t' \\ t & r' \end{pmatrix} \begin{pmatrix} \varphi_{il}(x_1) \\ \varphi_{ir}(x_2) \end{pmatrix} = S \begin{pmatrix} \varphi_{il}(x_1) \\ \varphi_{ir}(x_2) \end{pmatrix}, \tag{4.49}$$

Here S is the scattering matrix, the vectors $\begin{pmatrix} \varphi_{il}(x_1) \\ \varphi_{ir}(x_2) \end{pmatrix}$ and $\begin{pmatrix} \varphi_{ol}(x_1) \\ \varphi_{or}(x_2) \end{pmatrix}$ represent the incoming- and outgoing-particle amplitude, ϕ_i and ϕ_o, respectively. When the norm of the incoming-particles amplitude is preserved, we have

$$\phi_i^\dagger \phi_i = \phi_o^\dagger \phi_o = \phi_i^\dagger S^\dagger S \phi_i, \tag{4.50}$$

which means that the scattering matrix S is a unitary matrix that fulfils the condition

$$S^\dagger S = I, \tag{4.51}$$

where I is the unit matrix. Thus

$$|r|^2 + |t|^2 = 1, \qquad |r'|^2 + |t'|^2 = 1 \qquad \text{and} \qquad \left(\frac{t}{r}\right)^* = -\frac{t'}{r'}. \tag{4.52}$$

With the wave functions $\varphi_{il}(x_1)$ and $\varphi_{ol}(x_1)$ on the left side of the scattering region, and the wave functions $\varphi_{ir}(x_2)$ and $\varphi_{or}(x_2)$ on the right side, we can define the state vectors

$$\phi(x_1) = \begin{pmatrix} \varphi_{il}(x_1) \\ \varphi_{ol}(x_1) \end{pmatrix} \qquad \text{and} \qquad \phi(x_2) = \begin{pmatrix} \varphi_{or}(x_2) \\ \varphi_{ir}(x_2) \end{pmatrix}, \tag{4.53}$$

at x_1 and x_2, respectively, and put them in a relation like this

$$\begin{pmatrix} \varphi_{or}(x_2) \\ \varphi_{ir}(x_2) \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} \varphi_{il}(x_1) \\ \varphi_{ol}(x_1) \end{pmatrix} = M(x_2, x_1) \begin{pmatrix} \varphi_{il}(x_1) \\ \varphi_{ol}(x_1) \end{pmatrix}, \tag{4.54}$$

where $M(x_2, x_1)$ is the transfer matrix that connects the state vector at x_1 with the state vector at x_2. The transfer matrix $M(x_2, x_1)$ behaves as a propagator of the physical information from one point to another.

The scattering matrix S and the transfer matrix, have among them a simple and well defined relation. Indeed, if we combine (4.49) and (4.54), one can easily obtain the following relations

$$(\alpha + \beta r - t)\varphi_{il}(x_1) = (r' - \beta t')\varphi_{ir}(x_2), \tag{4.55}$$

and

$$(\gamma + \delta r)\varphi_{il}(x_1) = (1 - \delta t')\varphi_{ir}(x_2). \tag{4.56}$$

Since $\varphi_{il}(x_1)$ and $\varphi_{ir}(x_2)$ are linearly independent functions, we have

$$\begin{aligned} \alpha + \beta r - t = 0, &\qquad r' - \beta t' = 0, \\ \gamma + \delta r = 0, &\qquad 1 - \delta t' = 0. \end{aligned} \tag{4.57}$$

We can then solve for the scattering amplitudes

$$t = \alpha - \beta\frac{1}{\delta}\gamma, \quad r = -\frac{1}{\delta}\gamma, \quad t' = \frac{1}{\delta} \quad \text{and} \quad r' = \beta\frac{1}{\delta}. \tag{4.58}$$

or for the transfer matrix elements

$$\alpha = t - r' \frac{1}{t'} r, \quad \beta = r' \frac{1}{t'}, \quad \gamma = \frac{1}{t'} r \quad \text{and} \quad \delta = \frac{1}{t'}. \tag{4.59}$$

In the particular case of incidence from the left side only, the reflected and transmitted functions can be written as

$$\varphi_{\text{refl}} = r\varphi_{inc} \qquad \text{and} \qquad \varphi_{\text{trans}} = t\varphi_{inc}, \tag{4.60}$$

and the reflected and transmitted current densities take the form

$$\mathbf{j}_{\text{refl}} = |r|^2 \, \mathbf{j}_{\text{inc}} \qquad \text{and} \qquad \mathbf{j}_{\text{trans}} = |t|^2 \, \mathbf{j}_{\text{inc}}. \tag{4.61}$$

For most of the systems of interest, the interactions are invariant under time reversal. In this case (see appendix A) $\gamma = \beta^*$ and $\delta = \alpha^*$ and the scattering matrix is symmetric. Hence, when flux is conserved and the system is time reversal invariant, the transfer matrix M becomes unimodular ($\det M = 1$), and

$$t = \frac{1}{\alpha^*} \quad \text{and} \quad r = -\frac{\beta^*}{\alpha^*}. \tag{4.62}$$

These are important and simple relations. It is then clear that once we have the transfer matrix, we can immediately evaluate not only the transmission and reflection amplitudes, but also the *transmission and reflection coefficients*. In fact, if we consider the transmission and reflection coefficient, defined in (4.45) in terms of the incident, transmitted and reflected current densities, we have

$$T = \frac{|\mathbf{j}_{\text{trans}}|}{|\mathbf{j}_{\text{inc}}|} = tt^* = \frac{1}{|\alpha|^2} \qquad \text{and} \qquad R = \frac{|\mathbf{j}_{\text{refl}}|}{|\mathbf{j}_{\text{inc}}|} = rr^* = \frac{|\beta|^2}{|\alpha|^2}. \tag{4.63}$$

It is important to notice and to keep clear the relation and difference between (transmission and reflection) coefficients and (transmission and reflection) amplitudes.[6] This possibility of easy and straightforward calculation of transport properties is one of the advantages of the transfer matrix representation. We will find other properties. Let us mention just a couple of them.

A property that makes the transfer matrix extremely useful is the composition rule of two matrices

$$M(x_3, x_1) = M_2(x_3, x_2) M_1(x_2, x_1), \tag{4.64}$$

which is called the multiplicative property of transfer matrices. Using this property one can build up transfer matrices for almost any potential profile. We will use

---

[6] It is worth noticing also that while the transmission and reflection coefficients are real quantities, the transmission and reflection amplitudes are complex, in general.

this property in the piecewise constant potential systems, and in evaluating wave functions.

Some times we need to connect state vectors at two points of a constant potential region. In this case, the transfer matrix takes a simple form. Suppose that x_1 and x_2 are points of this type, with state vectors

$$\begin{pmatrix} ae^{ikx_2} \\ be^{-ikx_2} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} ae^{ikx_1} \\ be^{-ikx_1} \end{pmatrix}. \tag{4.65}$$

Notice that the coefficients a and b are the same because the differential equation is the same at x_1 and at x_2. Thus, we can relate these state vectors as follows:

$$\begin{pmatrix} ae^{ikx_2} \\ be^{-ikx_2} \end{pmatrix} = \begin{pmatrix} e^{ik(x_2-x_1)} & 0 \\ 0 & e^{-ik(x_2-x_1)} \end{pmatrix} \begin{pmatrix} ae^{ikx_1} \\ be^{-ikx_1} \end{pmatrix} \tag{4.66}$$

The matrix that connects state vectors at two points of a constant potential region, is diagonal and carries on the phase evolution of the exponential functions. When the solutions in the constant potential region are propagating functions, like in the free particle case mentioned in Chap. 3, the transfer matrix is

$$M_p(x_2, x_1) = \begin{pmatrix} e^{ik(x_2-x_1)} & 0 \\ 0 & e^{-ik(x_2-x_1)} \end{pmatrix}, \tag{4.67}$$

and, when the solutions are exponential functions, we have

$$\begin{pmatrix} ae^{qx_2} \\ be^{-qx_2} \end{pmatrix} = \begin{pmatrix} e^{q(x_2-x_1)} & 0 \\ 0 & e^{-q(x_2-x_1)} \end{pmatrix} \begin{pmatrix} ae^{qx_1} \\ be^{-qx_1} \end{pmatrix}. \tag{4.68}$$

Hence

$$M_p(x_2, x_1) = \begin{pmatrix} e^{q(x_2-x_1)} & 0 \\ 0 & e^{-q(x_2-x_1)} \end{pmatrix}. \tag{4.69}$$

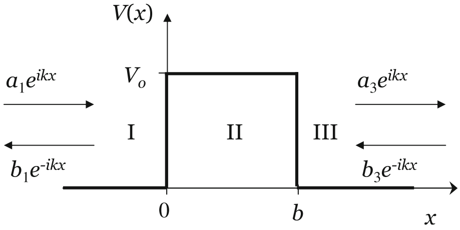
In the following sections, we will need and use these type of phase-propagating transfer matrices.

4.3 The Rectangular Potential Barrier

Let us now consider the rectangular potential barrier shown in Fig. 4.4. This potential energy corresponds to

$$V(x) = \begin{cases} 0, & x \le 0, \quad x \ge b; \\ V_o, & 0 < x < b. \end{cases} \tag{4.70}$$

Fig. 4.4 The rectangular potential barrier, its potential parameters and the propagating solutions at the *left* and *right*



As mentioned before, a potential profile like this was used to model repulsive potentials. The rectangular potential barrier is easy to solve and gives rise to important qualitative results, characteristic of quantum behavior. It will, in particular, make clear one of the most striking and significant quantum phenomena, *the tunneling effect*. The actual potential barriers and wells in metals, semiconductors and insulators appear in an inexhaustible variety of shapes. Even if one assumes the effective mass approximation within a layer of metal or semiconductor,[7] we can not avoid for the conduction band electrons, in layered structures, the strong effect of repulsive and attractive potentials as they pass from one medium to another. If, for example, on top of a $GaAs$ substrate one grows a layer of $Ga_{1-x}Al_xAs$ with thickness b and on top of this a layer of $GaAs$, the conduction band edge of the $GaAs/Al_xGa_{1-x}As/GaAs$ heterostructure will have (in the effective mass approximation) a potential profile that looks like a rectangular potential barrier. In this type of systems the height V_o depends essentially on the energy gaps[8] of the alternating semiconductor layers. In our example, it depends on the Aluminum concentration x and on the $GaAs$ energy gap E_g ($\simeq$1.5 eV). For a concentration of $\simeq$30 %, the barrier height V_o is $\simeq$0.23 eV.

We will solve the Schrödinger equation for the rectangular barrier, first, when the incident particle's energy E is less than V_o and, then, when it is greater than V_o. We will study these cases separately. To simplify the discussion of this problem we will assume, without loss of generality, that the incoming particles approach the potential barrier only from the left side.

In this problem, and in the oncoming piecewise constant potential problems, the differential equations are similar to those of the step potential. To avoid repetitions, we will use results already known, whenever it will be possible.

1. If $E < V_o$, we can immediately write down the Schrödinger equations and the corresponding solutions in each of the three regions that naturally we have in the potential domain. In regions I and III, the Schrödinger equation has the form

$$-\frac{\hbar^2}{2m}\frac{d^2\varphi(x)}{dx^2} = E\,\varphi(x). \tag{4.71}$$

---

[7] Because of the periodic potential and the tunneling phenomenon, electrons and holes in metals and semiconductors behave as if they will be moving in a constant potential but with a different mass. The effective mass.

[8] The energy gap E_g is the distance between the upper edge of the valence band and the lower edge of the conduction band. A gap of forbidden levels opens, generally, in a periodic system.

This is the same as in (4.2). Thus, the solutions are the wave functions[9]

$$\varphi_{\mathrm{I}}(x) = a_1 e^{ikx} + b_1 e^{-ikx}, \quad \text{for} \quad x \leq 0, \tag{4.72}$$

and

$$\varphi_{\mathrm{III}}(x) = a_3 e^{ikx} + b_3 e^{-ikx}, \quad \text{for} \quad x \geq b, \tag{4.73}$$

with $k = \sqrt{2mE/\hbar^2}$. In region II the Schrödinger equation is

$$\frac{\hbar^2}{2m}\frac{d^2\varphi_{\mathrm{II}}(x)}{dx^2} = \big(V_o - E\big)\,\varphi_{\mathrm{II}}(x). \tag{4.74}$$

This equation coincides with (4.3), therefore, the solution is

$$\varphi_{\mathrm{II}}(x) = a_2 e^{qx} + b_2 e^{-qx}, \quad \text{for} \quad 0 < x < b, \tag{4.75}$$

with $q = \sqrt{2m(V_o - E)/\hbar^2}$. Functions (4.72), (4.73) and (4.75) are formal solutions of the differential equations. We have still the problem of determining the coefficients a_i and b_i. They are generally obtained with the help of the fulfillment of continuity and boundary conditions. Once the coefficients are fixed and the continuity and physical conditions are imposed, the wave functions $\varphi_i(x)$, take a specific form and will become the actual solutions of the potential-barrier Schrödinger equation. Only then, we will be able to describe the physics of the quantum particles in the presence of the rectangular potential barrier.

Let us now consider the continuity conditions at $x = 0$ and $x = b$. From the continuity requirements at $x = 0$, we have

$$a_1 + b_1 = a_2 + b_2, \tag{4.76}$$

$$ik(a_1 - b_1) = q(a_2 - b_2). \tag{4.77}$$

and from the continuity requirements at $x = b$, we have

$$a_2 e^{qb} + b_2 e^{-qb} = a_3 e^{ikb} + b_3 e^{-ikb}, \tag{4.78}$$

$$q\big(a_2 e^{qb} - b_2 e^{-qb}\big) = ik\big(a_3 e^{ikb} - b_3 e^{-ikb}\big). \tag{4.79}$$

These equations allow us to obtain four coefficients out of six. In an open system like the potential barrier, the wave functions with coefficients a_1 and b_3 describe particles approaching the barrier from the left and from the right hand side, respectively. These amplitudes depend on the experimental set up, and can be fixed at will. For example, if we fix the experimental set up with incident particles coming only from the right hand side, we have to choose $a_1 = 0$. If we fix instead the

---

[9] In this text, we refer to these solutions as wave functions because we have in mind that they come always multiplied by the time dependent function $e^{-i\omega t}$.

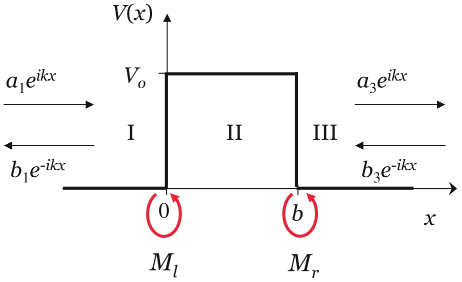


Fig. 4.5 The transfer matrix M_l connects state vectors at the *left* and *right* of $x = 0$, and the transfer matrix M_r connects state vectors at the *left* and *right* of $x = b$

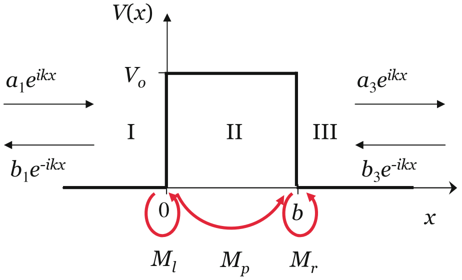


Fig. 4.6 Here the transfer matrix M_p connects the state vectors at the *right* of $x = 0$ with the state vectors at the *left* of $x = b$

experimental set up with incoming particles only from the left side, we have to choose $b_3 = 0$. Therefore, a_1 and b_3 play the role of independent parameters and one searches to express the other coefficients as functions of a_1 and b_3. The algebraic substitutions that we need to solve this kind of problems grow geometrically with the number of equations and become unmanageable. We will see that using the transfer matrix, the algebraic manipulation will diminish drastically.

4.3.1 Transfer Matrix of the Rectangular Potential Barrier

As in the step potential problem, the continuity relations (4.77)–(4.79) at $x = 0$ and $x = b$, can be written in the compact forms

$$\begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = \frac{1}{2q}\begin{pmatrix} q+ik & q-ik \\ q-ik & q+ik \end{pmatrix}\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = M_l(0^+,0^-)\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}; \tag{4.80}$$

and

$$\begin{pmatrix} a_3e^{ikb} \\ b_3e^{-ikb} \end{pmatrix} = \frac{1}{2k}\begin{pmatrix} k-iq & k+iq \\ k+iq & k-iq \end{pmatrix}\begin{pmatrix} a_2e^{qb} \\ b_2e^{-qb} \end{pmatrix} = M_r(b^+,b^-)\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}. \tag{4.81}$$

In Fig. 4.5, we sketch the connections established with the transfer matrices M_l and M_r. It is clear that a third matrix M_p is lacking. We need to connect, as shown in Fig. 4.6, the state vector at $x = 0 + \epsilon = 0^+$ with the state vector at $x = b - \epsilon = b^-$, where ϵ is an infinitesimally small increment. This connection is done by the matrix M_p. To obtain this matrix, we observe that

$$\begin{pmatrix} a_2 e^{qb} \\ b_2 e^{-qb} \end{pmatrix} = \begin{pmatrix} e^{qb} & 0 \\ 0 & e^{-qb} \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}, \tag{4.82}$$

thus

$$M_p(b^-, 0^+) = \begin{pmatrix} e^{qb} & 0 \\ 0 & e^{-qb} \end{pmatrix}. \tag{4.83}$$

Having this matrix, we are able to establish the relation

$$\phi_{\text{III}}(b^+) = M_b(b^+, 0^-)\phi_{\text{I}}(0^-). \tag{4.84}$$

with

$$M_b(b^+, 0^-) = M_r(b^+, b^-)M_p(b^-, 0^+)M_l(0^+, 0^-). \tag{4.85}$$

Notice that $M_p(b^-, 0^+)$ is a particular case of the transfer matrix (4.69), mentioned at the end of the last section. Replacing (4.80), (4.81) and (4.83) in (4.85), we have

$$M_b(b^+, 0^-) = \frac{1}{4qk} \begin{pmatrix} k - iq & k + iq \\ k + iq & k - iq \end{pmatrix} \begin{pmatrix} e^{qb} & 0 \\ 0 & e^{-qb} \end{pmatrix} \begin{pmatrix} q + ik & q - ik \\ q - ik & q + ik \end{pmatrix}. \tag{4.86}$$

Each matrix in this sequence plays a role in carrying the physical information from the left side to the right, across the barrier. Multiplying matrices and simplifying we have, finally, the transfer matrix of the rectangular potential barrier

$$M_b(b^+, 0^-) = \begin{pmatrix} \cosh qb + i\dfrac{k^2 - q^2}{2qk}\sinh qb & -i\dfrac{k^2 + q^2}{2qk}\sinh qb \\ i\dfrac{k^2 + q^2}{2qk}\sinh qb & \cosh qb - i\dfrac{k^2 - q^2}{2qk}\sinh qb \end{pmatrix}. \tag{4.87}$$

This is a *bona fide* transfer matrix that behaves as a propagator of the physical information across the barrier, from 0^- to b^+. Notice that these points are outside the barrier.

In the transfer matrix formalism, this is an important result of frequent reference. For convenience and compactness we denote this matrix as

$$M_b(b,0)=\begin{pmatrix}\alpha_b & \beta_b\\ \beta_b^* & \alpha_b^*\end{pmatrix}, \tag{4.88}$$

with

$$\begin{aligned}\alpha_b &= \cosh qb + i\frac{k^2-q^2}{2qk}\sinh qb,\\ \beta_b &= -i\frac{k^2+q^2}{2qk}\sinh qb.\end{aligned} \tag{4.89}$$

The structure of the matrix reflects, as mentioned before, the time reversal invariance symmetry of the potential barrier. Using this compact notation, we can write (4.83) as

$$\begin{pmatrix}a_3e^{ikb}\\ b_3e^{-ikb}\end{pmatrix}=\begin{pmatrix}\alpha_b & \beta_b\\ \beta_b^* & \alpha_b^*\end{pmatrix}\begin{pmatrix}a_1\\ b_1\end{pmatrix}. \tag{4.90}$$

In the transfer matrix we have all the physical information of the quantum particles and the potential function. We will use now the transfer matrices obtained in this section not only to define the wave functions, but also to obtain other results. Using M_b and the relation of this matrix with the scattering amplitudes r_b and t_b, we will obtain some transport properties like the transmission and reflection coefficients.

One can easily verify that the transition matrix in (4.81) is the inverse of the transition matrix in (4.80).

In the following we will avoid heavy notation and we will try to use a notation as light as possible. In fact, if we have to evaluate, say the function $f(x)$, at $a\pm\epsilon$ with $\epsilon\to 0$, we will denote it as $f(a^{\pm})$ or just as $f(a)$.

4.3.2 The Wave Functions in the Rectangular Potential Barrier

An important part of our objectives, solving the Schrödinger equation and deriving the transfer matrices, is to obtain the wave function, defined from $-\infty$ to $+\infty$. Our purpose in this section is to determine this wave function using the results obtained so far, especially the transfer matrix M_b. If we assume that the incident particles approach the potential barrier only from the left, we must have $b_3=0$, and (4.90) transforms into

$$\begin{pmatrix}a_3e^{ikb}\\ 0\end{pmatrix}=\begin{pmatrix}\alpha_b & \beta_b\\ \beta_b^* & \alpha_b^*\end{pmatrix}\begin{pmatrix}a_1\\ b_1\end{pmatrix}. \tag{4.91}$$

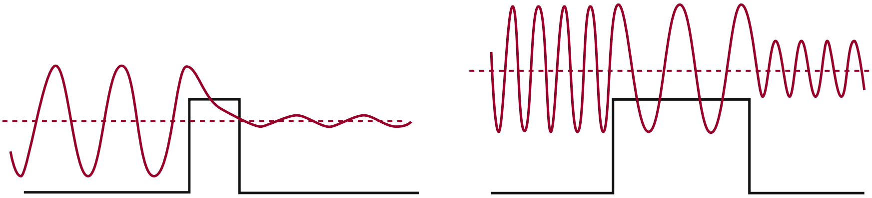


Fig. 4.7 Wave functions along the potential barrier and the tunneling effect. In the figure on the *left*, the wave function is plotted for $E = 0.18\,\mathrm{eV}$ ($< V_o = 0.23\,\mathrm{eV}$) and, on the *right* the wave function amplitude is plotted for $E = 0.28\,\mathrm{eV}$ ($> V_o$). The potential widths in the *left* and *right* figures are $b = 6$ nm and $b = 20$ nm, respectively. Notice that in the *left* figure, part of the incoming flux transmits through the barrier, even though the energy is below the barrier height. This is the tunneling effect

Multiplying, on the right hand side, we have

$$b_1 = -\frac{\beta_b^*}{\alpha_b^*} a_1, \tag{4.92}$$

and

$$a_3 = \left(\alpha_b - \beta_b \frac{\beta_b^*}{\alpha_b^*}\right) a_1 e^{-ikb} = \frac{1}{\alpha_b^*} a_1 e^{-ikb}, \tag{4.93}$$

where the flux conservation requirement $\alpha_b \alpha_b^* - \beta_b \beta_b^* = 1$ was used (see problem 4.3). With these coefficients we practically have the wave function. With b_1 we have the reflected fraction of the incident wave function, and with c_3 the transmitted part of the incident wave function. These coefficients are compatible with the reflection and transmission amplitudes $r = -\beta^*/\alpha^*$ and $t = 1/\alpha^*$, derived in Sect. 4.2. If we substitute b_1 in (4.72), we obtain the wave function on the left side of the barrier as

$$\varphi_{\mathrm{I}}(x) = a_1 e^{ikx} - \frac{\beta_b^*}{\alpha_b^*} a_1 e^{-ikx}, \qquad \text{for} \qquad x \le 0. \tag{4.94}$$

If we replace a_3 of (4.93) in the wave function $\varphi_{\mathrm{III}}(x)$, we have the transmitted wave function

$$\varphi_{\mathrm{III}}(x) = \frac{1}{\alpha_b^*} a_1 e^{ik(x-b)} \qquad \text{with} \qquad x \ge b. \tag{4.95}$$

To obtain the wave function in the barrier we use

$$\begin{pmatrix} a_2 e^{qx} \\ b_2 e^{-qx} \end{pmatrix} = \begin{pmatrix} e^{qx} & 0 \\ 0 & e^{-qx} \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}. \tag{4.96}$$

together with (4.80). After replacing the coefficients b_1, we have

$$\varphi_{\mathrm{II}}(x) = a_1 \left[\left(1 - \frac{\beta_b^*}{\alpha_b^*}\right) \cosh qx + i\frac{k}{q}\left(1 + \frac{\beta_b^*}{\alpha_b^*}\right) \sinh qx \right]. \tag{4.97}$$

In the left part of Fig. 4.7, we plot the absolute value of the wave functions given in (4.94), (4.95) and (4.97), when the energy is $E = 0.18$ eV. In the right part of this figure, we plot also the absolute value of the wave function when the energy is larger than V_o, using results that will be obtained below. It is important to notice that even though the energy is less than the potential height, we have a nonzero transmitted wave function. This property is related with the tunneling effect and with the penetration depth in the step potential.

4.3.3 *Reflection and Transmission Coefficients for Rectangular Potential Barriers*

Before obtaining the transmission and reflection coefficients, we recall that b_1 and a_3 can also be written as

$$b_1 = ra_1 \quad \text{and} \quad a_3 = ta_1 e^{-ikb}. \tag{4.98}$$

Hence, the wave functions $\varphi_{\mathrm{I}}(x)$ and $\varphi_{\mathrm{III}}(x)$, in (4.94) and (4.95), take the form (for $a_1 = 1$)[10]

$$\begin{aligned} \varphi_{\mathrm{I}}(x) &= e^{ikx} + r\, e^{-ikx} && \text{for} \quad x \le 0; \\ \varphi_{\mathrm{III}}(x) &= t\, e^{ik(x-b)} && \text{for} \quad x \ge b. \end{aligned} \tag{4.99}$$

It is clear from these expressions the role played by the reflection and the transmission amplitudes. Since

$$T = tt^* \quad \text{and} \quad R = rr^*, \tag{4.100}$$

we can immediately write down the transmission and reflection coefficients of the rectangular potential barrier. We have for the transmission coefficient the following expression

$$T = \frac{1}{|\alpha_b|^2} = \frac{1}{\left| \cosh qb + i\dfrac{k^2 - q^2}{2qk} \sinh qb \right|^2}, \tag{4.101}$$

while for the reflection coefficient we have

---

[10] This selection of the incident amplitude does not affect the subsequent results.

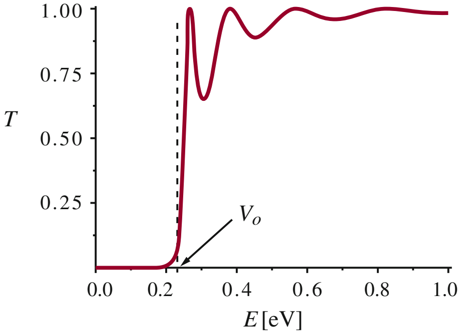


Fig. 4.8 The transmission coefficient as function of the energy for a potential barrier with $V_o = 0.23$ eV and $b = 6$ nm. It can be seen that transmission does not vanish for energies below the potential height V_o. The transmission coefficient oscillates for energies greater than V_o and becomes 1 only for some values of the energy

$$R = \frac{\left|\beta_b^*\right|^2}{\left|\alpha_b^*\right|^2} = \left(-\frac{k^2+q^2}{2kq}\sinh qb\right)^2 T. \tag{4.102}$$

These results, and, in particular the wave function $\varphi_{\text{III}}(x)$ in (4.99), exhibit an important difference with the classical physics description. While in classical physics the probability of finding a particle on the other side of the barrier for energies smaller than the barrier height is null, in the quantum description this kind of particles can tunnel through the barrier and then appear on the other side of the potential barrier with a probability given by

$$|\varphi_{\text{III}}(x)|^2_{x=b} = \frac{|a_1|^2}{\left|\alpha_b^*\right|^2} = |a_1|^2\, T, \tag{4.103}$$

different from zero, proportional to the transmission coefficient T. This phenomenon is *the tunneling effect*, mentioned before. It is a characteristic property of the quantum behavior. Starting from the transmission coefficient (4.101), it is possible to show that, in the limit $qb \gg 1$, the leading order term is

$$T \cong e^{-\frac{2b}{\hbar}\sqrt{2m(V_o-E)}}. \tag{4.104}$$

The transmission probability tends to 1 when the energy approaches V_o and decreases exponentially to zero when the product $V_o b^2$ grows.

In Fig. 4.8 we plot the transmission coefficient T of Eq. (4.101) as a function of the energy for $V_o = 0.23$ eV and $b = 6$ nm. Notice that for energies below V_o the transmission probability is small but different from zero.

The transmission coefficient in (4.101) was obtained for energies below the potential height. In principle this formula is valid only for $E < V_o$, however, if the plotting program used can handle the analytical continuation[11] one can extend the domain of application of this equation for energies larger than V_o.

---

[11] In this case, it means using the identities $\cosh ix = \cos x$ and $\sinh ix = i \sin x$.

We will study now the potential barrier for energies larger than V_o and will derive the transmission coefficient for $E > V_o$ and explain the resonant behavior in the high energies region.

2. If we have $E > V_o$, the solutions of the Schrödinger equation are:

$$\varphi_{\mathrm{I}}(x) = a_1 e^{ik_1x} + b_1 e^{-ik_1x}, \qquad \text{for} \qquad x \le 0; \tag{4.105}$$

$$\varphi_{\mathrm{II}}(x) = a_2 e^{ik_2x} + b_2 e^{-ik_2x}, \qquad \text{for} \qquad 0 < x < b; \tag{4.106}$$

$$\varphi_{\mathrm{III}}(x) = a_3 e^{ik_1x} + b_3 e^{-ik_1x}, \qquad \text{for} \qquad x \ge b. \tag{4.107}$$

Here $k_1 = \sqrt{\frac{2m}{\hbar^2} E}$ and $k_2 = \sqrt{\frac{2m}{\hbar^2} (E - V_o)}$. The only change that we have, compared with the solutions for $E < V_o$ is in $\varphi_{\mathrm{II}}(x)$. As in the step potential problem, this solution becomes an oscillatory function. All the expressions derived for the case $E < V_o$, will then have the wave number q replaced by ik_2. Taking into account this change, the transfer matrix of Eq. (4.87) becomes

$$M_b(b^+, 0^-) = \begin{pmatrix} \cos k_2 b + i\dfrac{k_1^2 + k_2^2}{2k_1k_2} \sin k_2 b & -i\dfrac{k_1^2 - k_2^2}{2k_1k_2} \sin k_2 b \\ i\dfrac{k_1^2 - k_2^2}{2k_1k_2} \sin k_2 b & \cos k_2 b - i\dfrac{k_1^2 + k_2^2}{2k_1k_2} \sin k_2 b \end{pmatrix}. \tag{4.108}$$

The hyperbolic functions become trigonometric functions. Therefore, the transmission coefficient, for energies above the barrier, is

$$T = \frac{1}{\left| \cos k_2 b + i\dfrac{k_1^2 + k_2^2}{2k_1k_2} \sin k_2 b \right|^2}. \tag{4.109}$$

This function can also be written in the form

$$T = \frac{\mathrm{I}}{1 + \dfrac{(k_1^2 - k_2^2)^2}{4k_1^2k_2^2} \sin^2 k_2 b}. \tag{4.110}$$

This result explains the oscillating behavior of the transmission coefficient in Fig. 4.8 for energies above the barrier height. It is clear that whenever the oscillating function $\sin k_2 b$ vanishes, the transmission coefficient T reaches its maximum value 1. The energies for which this phenomenon occurs are the resonant energies, and they are given by

$$E_l = \hbar^2 l^2 \pi^2 / 2mb + V_o, \qquad \text{with } l \text{ integer}. \tag{4.111}$$

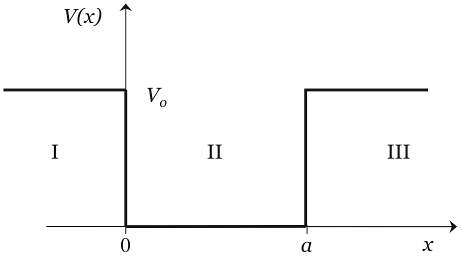


Fig. 4.9 Potential well with finite depth V_o and width a

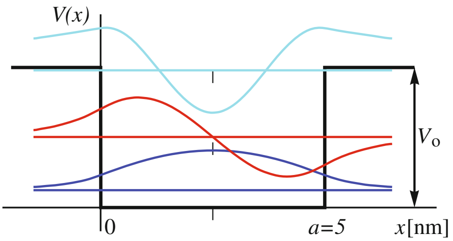


Fig. 4.10 The eigenfunctions φ_i in a finite potential well penetrate the walls and decay exponentially

These resonant energies should not be confused with energy eigenvalues. Recall that in this problem there is no energy quantization.

4.4 The Rectangular Potential Well

Another solvable and widely used system is the finite rectangular potential well. In Fig. 4.9 we have one example of this kind of systems. If we have a quantum particle in a potential well, we expect a rather different behavior from what we had in the potential barrier. When the energy is less than V_o, we have a confined or trapped particle whose energy quantizes. The potential function that corresponds to this system is

$$V(x) = \begin{cases} V_o, & x \le 0, \quad x \ge a; \\ 0, & 0 < x < a. \end{cases} \tag{4.112}$$

Potential energies like this, are found for electrons in the conduction (or holes in the valence) band of a semiconductor heterostructure like[12] $Ga_{1-x}Al_xA_s$/Ga A_s/ $Ga_{1-x}Al_xA_s$. The quantum wells have a wide range of applications, in particular in optoelectronic devices.[13] To study this system, it is also convenient to analyze the case of energy less than V_o and, separately, the case of energy greater than V_o.

---

[12] This is true in the effective mass approximation.

[13] In a quantum well laser the confined electrons of the conduction band recombine with confined holes of the valence band, emitting a photon whose energy $h\nu$ depends on the difference of the electron and hole energies.

1. When $0 < E < V_o$, the solutions of the Schrödinger equation in regions I, II and III are, respectively,

$$\begin{aligned} \varphi_{\rm I}(x) &= a_1 e^{qx} + b_1 e^{-qx}, && \text{for} \quad x \leq 0, \\ \varphi_{\rm II}(x) &= a_2 e^{ikx} + b_2 e^{-ikx}, && \text{for} \quad 0 < x < a, \\ \varphi_{\rm III}(x) &= a_3 e^{qx} + b_3 e^{-qx}, && \text{for} \quad a \leq x. \end{aligned} \tag{4.113}$$

with $q = \sqrt{2m(V_o - E)/\hbar^2}$ and $k = \sqrt{2mE/\hbar^2}$. Now, the solutions in regions I and III are exponential functions, and oscillating functions in region II. Since the wave functions $\varphi_{\rm I}(x)$ and $\varphi_{\rm III}(x)$ diverge when $x \to -\infty$ and $x \to \infty$, respectively, we have to choose $b_1 = a_3 = 0$. But we will, temporarily, retain all the coefficients and once the transfer matrix of the quantum well is obtained we will make $b_1 = a_3 = 0$.

4.4.1 Continuity and the Transfer Matrix of the Rectangular Potential Well

Unlike the infinite quantum well, where the wave functions vanish at the infinite walls, in the finite quantum well the exponential functions penetrate the finite walls, as shown in Fig. 4.10.

Imposing the continuity requirement on the wave functions and their first order derivatives, we have, at $x = 0$, the following equations

$$\begin{aligned} a_1 + b_1 &= a_2 + b_2, \\ q(a_1 - b_1) &= ik(a_2 - b_2), \end{aligned} \tag{4.114}$$

which, in the matrix representation, can be written as

$$\begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = \frac{1}{2k} \begin{pmatrix} k - iq & k + iq \\ k + iq & k - iq \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = M_l(0^+, 0^-) \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}. \tag{4.115}$$

Notice that the transition matrix that we have here, for the barrier-well interface, is the same as the transition matrix $M_r(b^+, b^-)$ that we had in the potential barrier. We will see now that the transition matrix that relates the state vectors at $x = a$, is equal to the transition matrix that we had on the left side of the potential barrier. Indeed, from the continuity conditions at $x = a$ we have

$$\begin{aligned} a_3 e^{qa} + b_3 e^{-qa} &= a_2 e^{ika} + b_2 e^{-ika}, \\ q\big(a_3 e^{qa} - b_3 e^{-qa}\big) &= ik\big(a_2 e^{ika} - b_2 e^{-ika}\big). \end{aligned} \tag{4.116}$$

These equations, in matrix and state vector representation, take the form

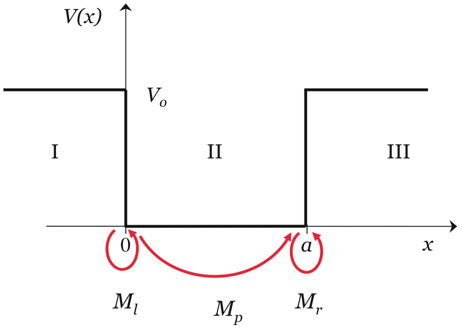


Fig. 4.11 The transfer matrices that connect the physics on the *left* with the physics on the *right* of the square quantum well

$$\begin{aligned}\begin{pmatrix} a_3e^{qa} \\ b_3e^{-qa} \end{pmatrix} &= \frac{1}{2q}\begin{pmatrix} q+ik & q-ik \\ q-ik & q+ik \end{pmatrix}\begin{pmatrix} a_2e^{ika} \\ b_2e^{-ika} \end{pmatrix}, \\ &= M_r(a^+,a^-)\begin{pmatrix} a_2e^{ika} \\ b_2e^{-ika} \end{pmatrix}, \end{aligned} \tag{4.117}$$

with a transition matrix, for the well-barrier interface, equal to the transition matrix $M_l(0^+,0^-)$ in the potential barrier. To connect the state vector at $x = 0^-$ with the state vector at $x = a^+$, we still need to connect the state vector $\phi_{\text{II}}(0^+)$ on the left end of the quantum well with the state vector $\phi_{\text{II}}(a^-)$ on the right. In other words, we need a transfer matrix M_p to carry the phase at the bottom of the quantum well. It is easy to verify that

$$\begin{pmatrix} a_2e^{ika} \\ b_2e^{-ika} \end{pmatrix} = \begin{pmatrix} e^{ika} & 0 \\ 0 & e^{-ika} \end{pmatrix}\begin{pmatrix} a_2 \\ b_2 \end{pmatrix}. \tag{4.118}$$

Therefore

$$M_p(a^-,0^+) = \begin{pmatrix} e^{ika} & 0 \\ 0 & e^{-ika} \end{pmatrix}, \tag{4.119}$$

With this matrix, we have all the necessary relations to connect the state vector in region III with the state vector in region I. Indeed, combining (4.115), (4.117) and (4.118), we obtain

$$\begin{pmatrix} a_3e^{qa} \\ b_3e^{-qa} \end{pmatrix} = \frac{1}{4qk}\begin{pmatrix} q+ik & q-ik \\ q-ik & q+ik \end{pmatrix}\begin{pmatrix} e^{ika} & 0 \\ 0 & e^{-ika} \end{pmatrix}\begin{pmatrix} k-iq & k+iq \\ k+iq & k-iq \end{pmatrix}\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}. \tag{4.120}$$

In this case, as for the potential barrier, we have the sequence of transition and transfer matrices

$$M_a(a^+,0^-) = M_r(a^+,a^-)M_p(a^-,0^+)M_l(0^+,0^-). \tag{4.121}$$

These matrices propagate along the quantum well, one after the other (see Fig. 4.11), carrying the physical information contained in the state vectors. After multiplying these matrices, and simplifying, we have

$$\begin{pmatrix} a_3 e^{qa} \\ b_3 e^{-qa} \end{pmatrix} = \begin{pmatrix} \cos ka + \dfrac{q^2 - k^2}{2qk} \sin ka & -\dfrac{k^2 + q^2}{2qk} \sin ka \\ \dfrac{k^2 + q^2}{2qk} \sin ka & \cos ka - \dfrac{q^2 - k^2}{2qk} \sin ka \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \quad (4.122)$$

Therefore the *transfer matrix of the rectangular quantum well* is

$$M_a(a^+, 0^-) = \begin{pmatrix} \cos ka + \dfrac{q^2 - k^2}{2qk} \sin ka & -\dfrac{k^2 + q^2}{2qk} \sin ka \\ \dfrac{k^2 + q^2}{2qk} \sin ka & \cos ka - \dfrac{q^2 - k^2}{2qk} \sin ka \end{pmatrix}, \quad (4.123)$$

This matrix will be written, in compact notation, as

$$M_a(a, 0) = \begin{pmatrix} \alpha_a & \beta_a \\ -\beta_a & \delta_a \end{pmatrix}. \quad (4.124)$$

Therefore

$$\begin{pmatrix} a_3 e^{qa} \\ b_3 e^{-qa} \end{pmatrix} = \begin{pmatrix} \alpha_a & \beta_a \\ -\beta_a & \delta_a \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}. \quad (4.125)$$

The transfer matrix structure is apparently different from that expected for transfer matrices satisfying the time reversal symmetry. We will see in the second part of this section, that for energies $E > V_o$ one has to replace the wave vector q by $i\kappa$, simultaneously the time reversal invariance structure, that we have seen for the transfer matrices in (4.90) and (4.108), reappears. What happens with the time reversal symmetry for energies $E < V_o$? The transfer matrix $M_a(a^+, 0^-)$ in (4.123) fulfills also the time reversal invariance requirement. It is known that this symmetry imposes, on transfer matrices, a more general condition. We will not deduce that condition here. It is possible however to show that to fulfill the time reversal symmetry, a transfer matrix M should satisfy the condition (see the Appendix A)

$$M \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} M^* = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (4.126)$$

One can easily verify that the matrix $M_a(a, 0)$, obtained above, satisfies this condition.

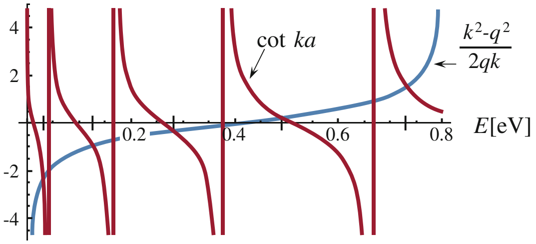


Fig. 4.12 A graphical method to determine the eigenvalues fulfilling the eigenvalues equation (4.130)

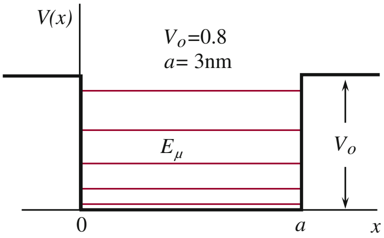


Fig. 4.13 Eigenvalues for a potential well whose depth is $V_o = 0.8$ eV

4.4.2 *Eigenvalues and Wave Functions in Rectangular Potential Wells*

Let us now come back into the main objective solving this problem, the calculation of eigenvalues and eigenfunctions. As mentioned earlier, the finiteness of our solutions require that $b_1 = a_3 = 0$. If we use these coefficients in (4.125), we have

$$\begin{pmatrix} 0 \\ b_3 e^{-qa} \end{pmatrix} = \begin{pmatrix} \alpha_a & \beta_a \\ -\beta_a & \delta_a \end{pmatrix} \begin{pmatrix} a_1 \\ 0 \end{pmatrix}, \tag{4.127}$$

which leads, on one side, to

$$a_1 \alpha_a = 0, \tag{4.128}$$

and, on the other, to

$$b_3 e^{-qa} = -\beta_a a_1. \tag{4.129}$$

From the first of these equations, and taking into account that a_1 can not be zero, we end up with the very important equation

$$\left(\cos ka + \frac{q^2 - k^2}{2qk} \sin ka\right) = 0, \tag{4.130}$$

that can be written also as

$$\cot ka = \frac{k^2 - q^2}{2qk}. \tag{4.131}$$

This equation is nothing else than *the energy eigenvalues equation*, that is fulfilled only by a discrete set of values of E, the energy eigenvalues of the quantum well. One way to solve this transcendental equation is plotting, separately, $\cot ka$ and $(k^2 - q^2)/2qk$, as functions of the energy. In Fig. 4.12 we have precisely this plot for $a = 3\,\text{nm}$ and $V_o = 0.8\,\text{eV}$. Both functions coincide at the eigenvalues E_μ. It is also possible to solve the eigenvalue equation numerically. In Fig. 4.13 the energy levels E_μ are shown for the quantum well with $a = 3\,\text{nm}$ and $V_o = 0.8\,\text{eV}$. It is worth noticing that the equation (4.131) provides all the eigenvalues that are generally obtained in other textbooks from two implicit equations, one associated to odd functions and the other to even functions. From (4.129), we get

$$b_3 = -e^{qa}\beta_a a_1. \tag{4.132}$$

This means that the wave functions in the left and right side walls can be written as

$$\varphi_{\text{I}}(x) = a_1 e^{qx}, \quad \text{for} \quad x < 0, \tag{4.133}$$

and

$$\varphi_{\text{III}}(x) = a_1 \frac{k^2 + q^2}{2qk} \sin ka \; e^{-q(x-a)}, \quad \text{for} \quad x > a, \tag{4.134}$$

respectively. As expected, both functions fall down exponentially as one goes deeper in the lateral walls. Before we come to the eigenfunction issue, let us see the solution φ_{II} in the well. If we return to (4.115) with $b_1 = 0$, we obtain the coefficients

$$a_2 = \frac{k - iq}{2k} a_1, \tag{4.135}$$

$$b_2 = \frac{k + iq}{2k} a_1, \tag{4.136}$$

which, replaced in $\varphi_{\text{II}}(x)$, leave us with the function

$$\varphi_{\text{II}}(x) = a_1 \left[\cos kx + \frac{q}{k} \sin kx\right]. \tag{4.137}$$

In this way, the solutions in each of the three regions have been written in terms of only one coefficient, the coefficient a_1, that can be fixed later, after normalizing the eigenfunctions (see Exercise 10 in the illustrative Problems section). Having these functions and the energy eigenvalues $\{E_\mu\}$, we are ready to obtain the eigenfunctions of the quantum well. The eigenfunctions $\varphi_\mu(x)$ are obtained when the wave functions $\varphi_{\text{I}}(x, E)$, $\varphi_{\text{II}}(x, E)$ and $\varphi_{\text{III}}(x, E)$, given in (4.133), (4.134) and (4.137), are evaluated for $E = E_\mu$. In other words, the eigenfunction that corresponds

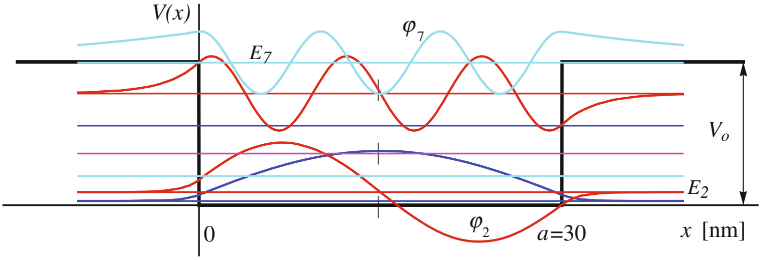


Fig. 4.14 The eigenvalues E_i and the corresponding eigenfunctions φ_i of a quantum well. In this well the depth is $V_o = 0.23$ eV and the width is $a = 30$ nm. These parameters are different from those in Fig. 4.10, where the well depth is greater but the width smaller. It is clear that increasing the width of the well, the energy levels diminish and increase their density. The parity of the eigenfunctions is defined with respect to the center of the well: The even functions alternate with the odd ones

to the energy eigenvalue E_μ is

$$\varphi_\mu(x) = \begin{cases} \varphi_{\mathrm{I}}(x, E_\mu), & \text{for} \quad x < 0, \\ \varphi_{\mathrm{II}}(x, E_\mu), & \text{for} \quad 0 \le x < a, \\ \varphi_{\mathrm{III}}(x, E_\mu); & \text{for} \quad x \ge a. \end{cases} \tag{4.138}$$

In Figs. 4.10 and 4.14 we plotted some eigenfunctions, together with their corresponding eigenvalues. To visualize the effect of the well width on the quantum levels, we consider a larger well width a in Fig. 4.14 than that in Fig. 4.13. It is clear from these figures that as we increase the well width, the energy levels come down and the eigenvalues' density (or level density) increases. There are many other properties that come out from our results. Among them, we notice that the eigenfunctions are real, thus the current density is zero everywhere and the parity of the eigenfunctions $\varphi_\mu(x)$ agrees with the parity of the quantum number μ. Even though the particles are mainly confined in the quantum well, the probability to find them beyond the lateral walls is different from zero. We will discuss now the quantum well problem when the energy is larger than the height of the lateral walls.

2. When $E > V_o$, the particle is no longer confined in the well and the solutions are all propagating functions, i.e.

$$\begin{aligned} \varphi_{\mathrm{I}}(x) &= a_1 e^{ik_2 x} + b_1 e^{-ik_2 x}, \quad \text{for} \quad x < 0, \\ \varphi_{\mathrm{II}}(x) &= a_2 e^{ik_1 x} + b_2 e^{-ik_1 x}, \quad \text{for} \quad 0 < x < a, \\ \varphi_{\mathrm{III}}(x) &= a_3 e^{ik_2 x} + b_3 e^{-ik_2 x}, \quad \text{for} \quad x > a. \end{aligned} \tag{4.139}$$

   Here $k_1 = \sqrt{\frac{2m}{\hbar^2} E}$ and $k_2 = \sqrt{\frac{2m}{\hbar^2}(E - V_o)}$. In these functions, the wave number q that we had in (4.114) is replaced by ik_2 and k by k_1. If we make these changes in (4.115) and in (4.117) we have

$$\begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = \frac{1}{2k_1} \begin{pmatrix} k_1 + k_2 & k_1 - k_2 \\ k_1 - k_2 & k_1 + k_2 \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, \tag{4.140}$$

and

$$\begin{pmatrix} a_3 e^{ik_2 a} \\ b_3 e^{-ik_2 a} \end{pmatrix} = \frac{1}{2k_2} \begin{pmatrix} k_2 + k_1 & k_2 - k_1 \\ k_2 - k_1 & k_2 + k_1 \end{pmatrix} \begin{pmatrix} a_2 e^{ik_1 a} \\ b_2 e^{-ik_1 a} \end{pmatrix}. \tag{4.141}$$

Therefore, the relation (4.120), that we had to connect the state vector $\phi_{\text{III}}(x)$ (at $x = a^+$) and the state vector $\phi_{\text{I}}(x)$ (at $x = 0^-$), becomes now

$$\begin{aligned} \begin{pmatrix} a_3 e^{ik_2 a} \\ b_3 e^{-ik_2 a} \end{pmatrix} &= \frac{1}{4k_1 k_2} \begin{pmatrix} k_2 + k_1 & k_2 - k_1 \\ k_2 - k_1 & k_2 + k_1 \end{pmatrix} \begin{pmatrix} e^{ik_1 a} & 0 \\ 0 & e^{-ik_1 a} \end{pmatrix} \begin{pmatrix} k_1 + k_2 & k_1 - k_2 \\ k_1 - k_2 & k_1 + k_2 \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, \\ &= M_a(a^+, 0^-) \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}. \end{aligned} \tag{4.142}$$

After multiplying and simplifying, the transfer matrix $M_a(a^+, 0^-)$ takes the form

$$M_a(a^+, 0^-) = \begin{pmatrix} \cos k_1 a + i \dfrac{k_1^2 + k_2^2}{2k_1 k_2} \sin k_1 a & i \dfrac{k_1^2 - k_2^2}{2k_1 k_2} \sin k_1 a \\ -i \dfrac{k_1^2 - k_2^2}{2k_1 k_2} \sin k_1 a & \cos k_1 a - i \dfrac{k_1^2 + k_2^2}{2k_1 k_2} \sin k_1 a \end{pmatrix}, \tag{4.143}$$

where the time reversal symmetry structure reappeared. In a compact notation we have

$$\begin{pmatrix} a_3 e^{ik_2 a} \\ b_3 e^{-ik_2 a} \end{pmatrix} = \begin{pmatrix} \alpha_a & \beta_a \\ \beta_a^* & \alpha_a^* \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}. \tag{4.144}$$

with

$$\begin{aligned} \alpha_a &= \cos k_1 a + i \frac{k_1^2 + k_2^2}{2k_1 k_2} \sin k_1 a, \\ \beta_a &= i \frac{k_1^2 - k_2^2}{2k_1 k_2} \sin k_1 a. \end{aligned} \tag{4.145}$$

We still have several coefficients to determine. We can get some of them if we impose additional physical conditions. For example, if the particles approach only from the left, the wave function $b_3 e^{-ik_2 x}$, describing particles moving towards the left in region III, must vanish. In this case we have to choose $b_3 = 0$, and (4.144) gives us the relation

$$\beta_a^* a_1 + \alpha_a^* b_1 = 0, \tag{4.146}$$

which implies that

$$b_1 = -\frac{\beta_a^*}{\alpha_a^*} a_1. \tag{4.147}$$

This coefficient has the functional form of the reflection amplitude multiplied by a_1. Indeed, the wave function $b_1 e^{-ikx}$ describes particles moving to the left in region I. These are particles reflected by the potential well. We will come back to this issue some lines below. Having b_1, we can express the wave function, at the left hand side of the well, as

$$\varphi_{\mathrm{I}}(x) = a_1 \left(e^{ik_2 x} - \frac{\beta_a^*}{\alpha_a^*} e^{-ik_2 x} \right). \tag{4.148}$$

From (4.144) and (4.147) we also have

$$a_3 e^{ika} = \alpha_a a_1 - \beta_a \frac{\beta_a^*}{\alpha_a^*} a_1. \tag{4.149}$$

Since the transfer matrix M_a is unimodular, with determinant $\alpha_a \alpha_a^* - \beta_a \beta_a^* = 1$, the transmitted wave function, in the right hand side of the well, can be written as

$$\varphi_{\mathrm{III}}(x) = \frac{1}{\alpha_a^*} a_1 e^{ik_2(x-a)}, \tag{4.150}$$

with $1/\alpha_a^*$ the transmission amplitude. Again, we have a coefficient that is formally similar to that found for the potential barrier. It is important to notice that even though the explicit function α_a is different than the explicit function α_b, of the potential barrier, *the functional dependence of physical quantities, like the transmission and reflection amplitudes, on the transfer matrix elements* α, β, *etc., is the same*. For example, the transmission amplitude in the potential barrier is $t_b = 1/\alpha_b^*$, the transmission amplitude in the quantum well is $t_a = 1/\alpha_a^*$. Moreover, they share the same functional dependence on α as the transmission amplitude in (4.62).This is another advantage of the transfer matrix formalism. The physical quantities are formally given by the same functions of α's and β's, as was found in Sect. 4.2.

Using (4.140) together with (4.147), one can finally express $\varphi_{\mathrm{II}}(x)$ as

$$\varphi_{\mathrm{II}}(x) = \frac{a_1}{2k_1} \left[(k_1 + k_2) e^{ik_1 x} - (k_1 - k_2) \frac{\beta_a^*}{\alpha_a^*} e^{-ik_1 x} \right]. \tag{4.151}$$

The coefficient a_1, is a common (scale) factor for $\varphi_{\mathrm{I}}(x)$, $\varphi_{\mathrm{II}}(x)$ and $\varphi_{\mathrm{III}}(x)$. It can be determined from the normalization condition. This constant has no effect on physical quantities like transmission and reflection coefficients.

It is worth noticing that for $E > V_o$ there is no energy quantization and except for some resonant energies, where the transmission is complete, the incoming particles always have a probability different from zero to get reflected. This behavior is better described by the reflection and transmission coefficients, that we will obtain now. As we know already, the transmission coefficient is given by the quotient $1/|\alpha_a|^2$, which means

$$T = \frac{1}{\cos^2 k_1 a + \left(\frac{k_1^2 + k_2^2}{2k_1 k_2}\right)^2 \sin^2 k_1 a}. \tag{4.152}$$

and the reflection coefficient by $|-\beta_a^*/\alpha_a^*|^2$. Therefore

$$R = \frac{\left(\frac{k_1^2 - k_2^2}{2k_1 k_2}\right)^2 \sin^2 k_1 a}{\cos^2 k_1 a + \left(\frac{k_1^2 + k_2^2}{2k_1 k_2}\right)^2 \sin^2 k_1 a}. \tag{4.153}$$

The student can easily check that $R + T = 1$. It is evident from these results that whenever

$$\sin^2 k_1 a = 0, \tag{4.154}$$

or

$$k_1 a = n\pi, \tag{4.155}$$

we have a resonant transmission with $T = 1$. Hence the resonant transmission through the potential well occurs when the de Broglie wavelengths $\lambda_B = 2\pi/k_1$ are such that

$$a = n\frac{\lambda_B}{2}. \tag{4.156}$$

Since $k_1^2 = 2mE/\hbar^2$, the resonant energies are

$$E_n = \frac{\hbar^2}{2m}\frac{n^2\pi^2}{a^2}. \tag{4.157}$$

These energies should not be confused with energy eigenvalues. If we recall that $E \geq V_o$ the integers n should satisfy the relation

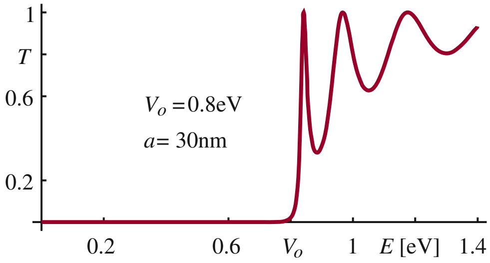


Fig. 4.15 The transmission coefficient of a potential well, with depth $V_o = 0.8$ eV and width $a = 30$ nm. These transport quantities have physical meaning only when the particle energy E is larger than V_o

$$n \geq \sqrt{\frac{2mV_o}{\hbar^2}}\frac{a}{\pi}. \tag{4.158}$$

In the following chapter we will see other potential profiles with different configurations of barriers and wells. Some of the results obtained in this chapter will be used, and we will be able to attest the advantages of the transfer matrix method that will greatly facilitate some calculations, which otherwise would become very cumbersome.

4.5 Solved Problems

Exercise 10 Determine the normalization constant for the eigenfunctions of the rectangular potential well, when the energy is less than the lateral wall height.

Solution All we need is to fulfill the normalization requirement

$$\int_{-\infty}^{\infty} |\varphi_\mu(x)|^2 dx = 1. \tag{4.159}$$

For the potential well, the eigenfunction was defined as

$$\varphi_\mu(x) = \begin{cases} \varphi_{\rm I}(x, E_\mu); & \text{for} \quad x < 0, \\ \varphi_{\rm II}(x, E_\mu); & \text{for} \quad 0 \leq x < a, \\ \varphi_{\rm III}(x, E_\mu); & \text{for} \quad x \geq a. \end{cases} \tag{4.160}$$

Thus, the normalization condition becomes

$$\int_{-\infty}^{0} |\varphi_{\rm I}(x, E_\mu)|^2 dx + \int_{0}^{a} |\varphi_{\rm II}(x, E_\mu)|^2 dx + \int_{a}^{\infty} |\varphi_{\rm III}(x, E_\mu)|^2 dx = 1. \tag{4.161}$$

If we use the explicit functions, we have

$$a_1^2\left[\int_{-\infty}^{0} e^{2q_\mu x}dx+\int_0^a\left(\cos k_\mu x+\frac{q_\mu}{k_\mu}\sin k_\mu x\right)^2 dx\right.$$
$$\left.+\int_a^\infty\left(\frac{k_\mu^2+q_\mu^2}{2q_\mu k_\mu}\right)^2\sin^2 k_\mu a\; e^{-2q_\mu(x-a)}dx\right]=1, \qquad (4.162)$$

with $k_\mu=\sqrt{2mE_\mu/\hbar^2}$ and $q_\mu=\sqrt{2m(V_o-E_\mu)/\hbar^2}$. All the integrals here can be evaluated analytically, hence the student can easily obtain the normalization constants. Notice that we have one constant for each eigenfunction.

Exercise 11 Show that at any point x of the potential well, and independently of the energy level E_μ, the current density is zero.

Solution For each energy level E_μ we have an eigenfunction $\varphi_\mu(x)$ defined from $-\infty$ to ∞. In the quantum well region the eigenfunction is given by

$$\varphi_{\mathrm{II}\mu}(x)=a_1\left[\cos k_\mu x+\frac{q_\mu}{k_\mu}\sin k_\mu x\right], \qquad (4.163)$$

with k_μ and q_μ the wave numbers evaluated for $E=E_\mu$. The current density is zero because the function ($\varphi_{\mathrm{II}\mu}(x)=\varphi^*_{\mathrm{II}\mu}(x)$) is real at each point x of the well, regardless of the index μ.

Exercise 12 Let us consider the asymmetric barrier of Fig. 4.16, with potential height V_i on the left and $V_d\neq V_i$ on the right of the barrier. Show that for energies greater than V_i and V_d, the transfer matrix

$$M(x_d,x_i)=\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}, \qquad (4.164)$$

also has the structure

$$M(x_d,x_i)=\begin{pmatrix}\alpha & \beta\\ \beta^* & \alpha^*\end{pmatrix}. \qquad (4.165)$$

Solution For systems like this, and more complex ones, we must use unit-flux wave functions (see Sect. 3.3). Therefore, the solutions in regions I, II and III are:

$$\begin{aligned}\varphi_{\mathrm{I}}(x)&=a_1\frac{e^{ik_l x}}{\sqrt{k_l}}+b_1\frac{e^{-ik_l x}}{\sqrt{k_l}}, &&\text{for}\quad x<0,\\ \varphi_{\mathrm{II}}(x)&=a_2\frac{e^{qx}}{\sqrt{q}}+b_2\frac{e^{-qx}}{\sqrt{q}}, &&\text{for}\quad 0<x<b,\\ \varphi_{\mathrm{III}}(x)&=a_3\frac{e^{ik_r x}}{\sqrt{k_r}}+b_3\frac{e^{-ik_r x}}{\sqrt{k_r}}, &&\text{for}\quad x>b.\end{aligned} \qquad (4.166)$$

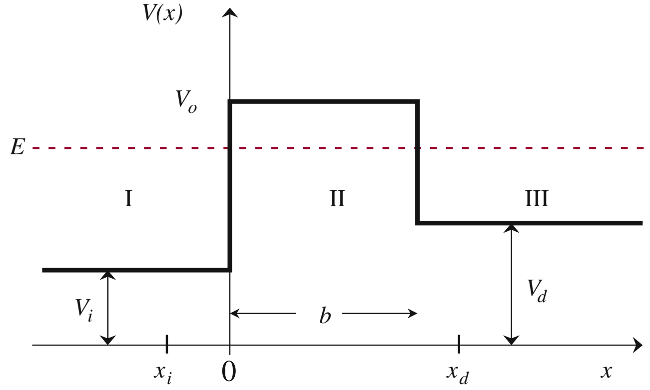


Fig. 4.16 An asymmetric potential barrier with different potential parameters at the *left* and *right* sides

with $k_l = \sqrt{2m(E-V_i)/\hbar^2}$, $k_r = \sqrt{2m(E-V_d)/\hbar^2}$ and $q = \sqrt{2m(V_o-E)/\hbar^2}$. In the same way as for the symmetric potential barrier, studied before, we can establish the relation

$$\begin{pmatrix} a_3 e^{ik_r b^+} \\ b_3 e^{-ik_r b^+} \end{pmatrix} = M_{ab}(b^+, 0^-) \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}. \tag{4.167}$$

In this case (the subindex *ab* stands for asymmetric barrier)

$$M_{ab}(b^+, 0^-) = \frac{1}{4qk_r} \begin{pmatrix} k_r - iq & k_r + iq \\ k_r + iq & k_r - iq \end{pmatrix} \begin{pmatrix} e^{qb} & 0 \\ 0 & e^{-qb} \end{pmatrix} \begin{pmatrix} q + ik_l & q - ik_l \\ q - ik_l & q + ik_l \end{pmatrix} \sqrt{\frac{k_r}{k_l}} \tag{4.168}$$

is the matrix that transfers the state vector from the left to the right of the barrier (from $x_l = 0^-$ to $x_r = b^+$). Multiplying the matrices in (4.168) and simplifying, yields the transfer matrix

$$M_{ab}(b^+, 0^-) = \begin{pmatrix} \alpha_{ab} & \beta_{ab} \\ \beta^*_{ab} & \alpha^*_{ab} \end{pmatrix}, \tag{4.169}$$

with

$$\alpha_{ab} = \left(\frac{k_l + k_r}{2k_r} \cosh qb + i \frac{k_l k_r - q^2}{2qk_r} \sinh qb \right) \sqrt{\frac{k_r}{k_l}}, \tag{4.170}$$

$$\beta_{ab} = \left(\frac{k_r - k_l}{2k_r} \cosh qb - i \frac{k_l k_r + q^2}{2qk_r} \sinh qb \right) \sqrt{\frac{k_r}{k_l}}. \tag{4.171}$$

Notice also that, making $k_l = k_r$, we recover the transfer matrix of the symmetric rectangular potential barrier.

Exercise 13 Assume that the potential is asymmetric, as in the previous problem, and that the particles approach only from one side. If they come from the left, the transmission amplitude (from left to right) will be denoted as t_{rl}, and if they approach from the right, the transmission amplitude (from right to left) will be denoted as t_{lr}. Show that the amplitudes and the corresponding transmission coefficients satisfy the following relations

$$t_{rl} = t_{lr}^* \qquad \text{and} \qquad T_{rl} = T_{lr} \tag{4.172}$$

Solution In the previous exercise we had the relation (4.167) where $\phi_{\mathrm{III}}(b^+)$ appears as a function of $\phi_{\mathrm{I}}(0^-)$. We can reverse this relationship to have

$$\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} = M_{ba}(0^-, b^+)\begin{pmatrix} a_3 e^{ik_r b^+} \\ b_3 e^{-ik_r b^+} \end{pmatrix}, \tag{4.173}$$

where $M_{ba}(0^-, b^+)$ is the inverse of $M_{ab}(b^+, 0^-)$, thus, the product of the inverse matrices in reverse order, i.e.

$$M_{ba}(0^-, b^+) = \frac{1}{4qk_l}\begin{pmatrix} k_l - iq & k_l + iq \\ k_l + iq & k_l - iq \end{pmatrix}\begin{pmatrix} e^{-qb} & 0 \\ 0 & e^{qb} \end{pmatrix}\begin{pmatrix} q + ik_r & q - ik_r \\ q - ik_r & q + ik_r \end{pmatrix}\sqrt{\frac{k_l}{k_r}}. \tag{4.174}$$

This matrix, after multiplying, takes the form

$$M_{ba}(0^-, b^+) = \begin{pmatrix} \alpha_{ba} & \beta_{ba}, \\ \beta_{ba}^* & \alpha_{ba}^* \end{pmatrix}, \tag{4.175}$$

with

$$\alpha_{ba} = \left(\frac{k_l + k_r}{2k_l}\cosh qb - i\frac{k_l k_r - q^2}{2qk_l}\sinh qb\right)\sqrt{\frac{k_l}{k_r}}, \tag{4.176}$$

$$\beta_{ba} = \left(\frac{k_r - k_l}{2k_i}\cosh qb + i\frac{k_l k_r + q^2}{2qk_l}\sinh qb\right)\sqrt{\frac{k_l}{k_r}}. \tag{4.177}$$

These matrix elements are similar to those in Eqs. (4.170) and (4.171) with k_l changed by k_r, and vice versa. It is easy to see from these expressions that

$$\alpha_{ba} = \alpha_{ab}^*. \tag{4.178}$$

The transmission amplitudes t_{rl} (left to right) and t_{lr} (right to left) have a very simple relation with the matrix elements α of M_{ab} and M_{ba}, respectively. Indeed

$$t_{rl} = \frac{1}{\alpha^*_{ab}} \quad \text{and} \quad t_{lr} = \frac{1}{\alpha^*_{ba}}. \tag{4.179}$$

If we combine Eqs. (4.178) and (4.179), we obtain the expected results

$$t_{rl} = t^*_{lr} \quad \text{and} \quad T_{rl} = T_{lr}. \tag{4.180}$$

These results are consistent with (4.58), where transmission amplitudes t and t' incidence from left and right were given in general.

4.6 Problems

1. Derive the currents $\boldsymbol{j}_{\text{inc}}$, $\boldsymbol{j}_{\text{ref}}$, and $\boldsymbol{j}_{\text{trans}}$ in Eq. (4.37).
2. Show that the transfer matrix of the potential barrier is given by Eq. (4.74).
3. Show that the current density on the left side of the symmetric potential barrier is the same as in the right hand side. Show also that

$$\text{Det} M_b = 1, \tag{4.181}$$

   where Det stands for the determinant.
4. Obtain the reflected and transmitted wave functions in terms of the reflection and transmitted amplitudes and show that

$$T = tt^* = \frac{1}{|\alpha|^2} \qquad \text{and} \qquad R = rr^* = \frac{|\beta|^2}{|\alpha|^2}. \tag{4.182}$$

5. Show that for time reversal invariant potential $V(x)$ defined from x_1 to x_2, the transfer matrix has the structure

$$M(x_2, x_1) = \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix}. \tag{4.183}$$

   Assume that in regions I (at the left of x_1) and III (at the right of x_2) the potential is zero. Show that, when the current densities $\mathbf{j}_{\text{I}}(x)$ and $\mathbf{j}_{\text{III}}(x)$ are equal, one has

$$M^\dagger \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad \text{Det}\; M(x_2, x_1) = \begin{vmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{vmatrix} = 1. \tag{4.184}$$

6. Verify that the transfer matrix of the potential barrier, for $E > V_o$, takes the form given in (4.87).

7. Show that when the particle energy is $E = V_o/2$, the transmission coefficient of the rectangular potential barrier becomes

$$T = \frac{1}{1 + \sinh^2 qb}. \tag{4.185}$$

8. Show that time reversal invariance implies for the quantum-well transfer matrix $\begin{pmatrix} \alpha & \beta \\ -\beta & \delta \end{pmatrix}$ the fulfillment of

$$\alpha\delta + \beta^2 = 1, \tag{4.186}$$

and verify that this condition is certainly satisfied by the transfer matrix M_a in (4.123).

9. Assume that the potential well is defined as

$$V(x) = \begin{cases} 0, & x \leq 0, \quad x \geq a; \\ -V_o, & 0 < x < a. \end{cases} \tag{4.187}$$

Prove that using this potential, one obtains the same results of Sect. 4.4, with $q = \sqrt{\frac{2m|E|}{\hbar^2}}$ and $k = \sqrt{\frac{2m(|V_o|-|E|)}{\hbar^2}}$.

10. Check that the wave function in the potential barrier is as shown in Eq. (4.97).
11. Using the solutions in (4.148), (4.151) and (4.150), check that:
    (i) at $x = 0 \quad \varphi_{\text{II}}(0) = \varphi_{\text{I}}(0) = a_1$;
    (ii) at $x = a \quad \varphi_{\text{II}}(a) = \varphi_{\text{III}}(a) = a_1 \frac{k^2+q^2}{2qk} \sin ka$.
12. Show that when the width and height of a potential barrier are such that $qb \gg 1$, the transmission coefficient tends to the exponential function

$$T \cong e^{-\frac{2b}{\hbar}\sqrt{2m(V_o-E)}}. \tag{4.188}$$

13. Show that the minima of the square-barrier transmission coefficient oscillations, for energies greater than V_o, is described by the function

$$T_{\min} = \frac{1}{\left[\left(\frac{2}{bk_1}\right)^2 + 1\right]\left(\frac{k_1^2 + k_2^2}{2k_1k_2}\right)^2} \tag{4.189}$$

14. Show that the transfer matrix of the asymmetric potential barrier, of exercise 12, has the form given in (4.169). Obtain also the transfer matrix of an asymmetric potential well.

Chapter 5
Quantum Coherence and Energy Levels Splitting

In the previous chapter we solved the Schrödinger equations for the potential barrier and the quantum well. These important and soluble systems, revealed fundamental properties present in the actual quantum systems. Some of the results obtained for these systems will reappear, in this and the coming chapters, and will enhance the insight into the quantum phenomena and their physical meaning. In this chapter we will study systems with slightly more complex structure, but still piecewise constant potentials. We will begin with the double quantum well with infinite walls, and we will continue with the double potential barrier and the double quantum well with finite lateral walls. In these systems a new phenomenon will appear: *the splitting of the energy levels*. We will conclude this chapter with a brief introduction to the finite periodic systems theory, applied to the Kronig–Penney model, taken as a finite sequence of rectangular wells and barriers. We will see that the energy levels splitting is responsible for the formation of energy bands, an essential property closely related to the quantum phase coherence, that makes it possible for us to see fundamental differences in the physical behavior of metals, semiconductors and insulators.

In quantum systems, as in any other physical system described by the wave equation, the wave functions $\psi(\mathbf{r}, t)$ evolve in space and time. The evolution is coherent when the amplitudes and phases, at any point $(\mathbf{r}_2, t_2)$, can be obtained once they are given at some other point $(\mathbf{r}_1, t_1)$. Phase coherence implies processes where the wave function interference is well defined.[1] This is not the case in disordered systems where the random fluctuations tend, generally, to destroy the phase coherence. In these systems, the phase coherence length ℓ_ϕ measures the average distance beyond which the phase coherence becomes negligible.

---

[1] Usually the superposition of waves gives rise to constructive and destructive interferences. These interferences of well defined phases, are responsible of interesting effects like the energy levels splitting and the band structure in periodic systems.

P. Pereyra, *Fundamentals of Quantum Physics*, Undergraduate Lecture Notes in Physics, DOI: 10.1007/978-3-642-29378-8_5, © Springer-Verlag Berlin Heidelberg 2012

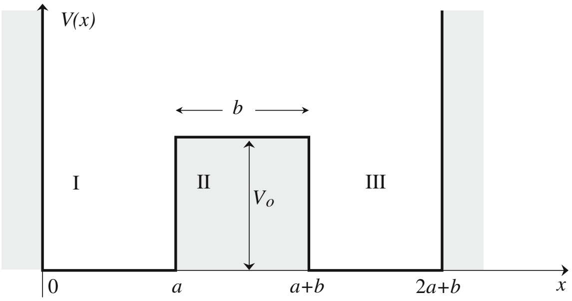


Fig. 5.1 A double well bounded by infinite walls. If the energy is less than V_o there are two confinement regions. The energy levels split and are pushed up by the repulsive infinite walls

5.1 A Rectangular Double Well Bounded by Hard Walls

Suppose we have the potential shown in Fig. 5.1. This double well potential is defined by the potential energy function

$$V(x) = \begin{cases} 0, & \text{for} \quad 0 < x < a \quad \& \quad a + b < x < 2a + b, \\ V_o, & \text{for} \quad a \le x \le a + b, \\ \infty, & \text{for} \quad x < 0 \quad \& \quad x > 2a + b. \end{cases} \tag{5.1}$$

This can be seen as a potential barrier inside an infinite quantum well. The potential barrier together with the infinite walls, located at a distance a to the left and to the right from the barrier, produce the double quantum well system. When the energy is such that $E < V_o$, we have two confining regions, but when $E > V_o$ only the infinite walls confine. Although we can expect energy quantization for energies below and above the barrier height V_o, we can also expect slight differences between the two cases (see problem 1). We will solve the double quantum well problem only for $E < V_o$.

5.1.1 Continuity and the Double-Well Transfer Matrix

In this system as in the potential barrier case, the solutions of the Schrödinger equation, in regions I, II and III, are:

$$\begin{aligned} \varphi_{\mathrm{I}}(x) &= a_1 e^{ikx} + b_1 e^{-ikx}; \\ \varphi_{\mathrm{II}}(x) &= a_2 e^{qx} \ + b_2 e^{-qx}; \\ \varphi_{\mathrm{III}}(x) &= a_3 e^{ikx} + b_3 e^{-ikx}. \end{aligned} \tag{5.2}$$

With $k = \sqrt{\frac{2m}{\hbar^2} E}$ and $q = \sqrt{\frac{2m}{\hbar^2} (V_o - E)}$. The infinite walls force the functions $\varphi_{\mathrm{I}}(x)$ and $\varphi_{\mathrm{III}}(x)$ to vanish at $x = 0$ and at $x = 2a + b$, respectively. This physical condition will be used below. These functions should also satisfy the continuity conditions at $x = a$ and $x = a + b$. In the potential barrier problem the discontinuity points were at $x = 0$ and $x = b$. If we take into account this difference, it is clear that the same transfer matrix M_b of Eq. (4.88), connects the state vectors $\phi_{\mathrm{III}}(a + b^+)$ and $\phi_{\mathrm{I}}(a^-)$, i.e:

$$\phi_{\mathrm{III}}(a + b^+) = \begin{pmatrix} \alpha_b & \beta_b \\ \beta_b^* & \alpha_b^* \end{pmatrix} \phi_{\mathrm{I}}(a^-), \tag{5.3}$$

or, more explicitly, we must have

$$\begin{pmatrix} a_3 e^{ik(a+b)} \\ b_3 e^{-ik(a+b)} \end{pmatrix} = \begin{pmatrix} \cosh qb + i\dfrac{k^2 - q^2}{2qk}\sinh qb & -i\dfrac{k^2 + q^2}{2qk}\sinh qb \\ i\dfrac{k^2 + q^2}{2qk}\sinh qb & \cosh qb - i\dfrac{k^2 - q^2}{2qk}\sinh qb \end{pmatrix} \begin{pmatrix} a_1 e^{ika} \\ b_1 e^{-ika} \end{pmatrix}. \tag{5.4}$$

We already know that the state vectors at different points of the same constant potential region, differ only in their phases. Indeed, for the state vectors at the extreme points, $x = 0^+$ and $x = a^-$, of the left side well, we have

$$\begin{pmatrix} a_1 e^{ika} \\ b_1 e^{-ika} \end{pmatrix} = \begin{pmatrix} e^{ika} & 0 \\ 0 & e^{-ika} \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}. \tag{5.5}$$

Similarly, for the extreme points of the right side well, $x = a + b^+$ and $x = 2a + b^-$, we have

$$\begin{pmatrix} a_3 e^{ik(2a+b)} \\ b_3 e^{-ik(2a+b)} \end{pmatrix} = \begin{pmatrix} e^{ika} & 0 \\ 0 & e^{-ika} \end{pmatrix} \begin{pmatrix} a_3 e^{ik(a+b)} \\ b_3 e^{-ik(a+b)} \end{pmatrix}. \tag{5.6}$$

Using (5.4) and (5.5) in this equation, we end up with the following relation

$$\begin{pmatrix} a_3 e^{ik(2a+b)} \\ b_3 e^{-ik(2a+b)} \end{pmatrix} = \begin{pmatrix} e^{ika} & 0 \\ 0 & e^{-ika} \end{pmatrix} \begin{pmatrix} \alpha_b & \beta_b \\ \beta_b^* & \alpha_b^* \end{pmatrix} \begin{pmatrix} e^{ika} & 0 \\ 0 & e^{-ika} \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}. \tag{5.7}$$

This relation between the vectors $\phi_{\mathrm{III}}(2a + b^-)$ and $\phi_{\mathrm{I}}(0^+)$, at the ends of the double well, can be written after multiplying matrices as

$$\begin{pmatrix} a_3 e^{ik(2a+b)} \\ b_3 e^{-ik(2a+b)} \end{pmatrix} = \begin{pmatrix} \alpha_b e^{i2ka} & \beta_b \\ \beta_b^* & \alpha_b^* e^{-i2ka} \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, \tag{5.8}$$

The transfer matrix of this equation contains already the continuity conditions at $x = a$ and $x = a + b$. We have still to introduce the conditions at $x = 0$ and $x = 2a + b$. At these points the functions $\phi_{\mathrm{I}}(0)$ and $\phi_{\mathrm{III}}(2a + b)$ must vanish, i.e.

$$\varphi_{\mathrm{I}}(0) = a_1 + b_1 = 0, \tag{5.9}$$

$$\varphi_{\mathrm{III}}(2a+b) = a_3 e^{ik(2a+b)} + b_3 e^{-ik(2a+b)} = 0. \tag{5.10}$$

Therefore

$$b_1 = -a_1, \tag{5.11}$$

$$b_3 = -a_3 e^{i2k(2a+b)}. \tag{5.12}$$

Here we have two alternatives: we can substitute these coefficients in (5.2), or replace them in (5.8). In the first case, $\varphi_{\mathrm{I}}(x)$ and $\varphi_{\mathrm{III}}(x)$ become

$$\varphi_{\mathrm{I}}(x, E) = a_1\big(e^{ikx} - e^{-ikx}\big) = A_1 \sin kx, \qquad \text{for} \qquad 0 < x < a; \tag{5.13}$$

and

$$\begin{aligned}\varphi_{\mathrm{III}}(x, E) &= a_3\Big(e^{ikx} - e^{-ikx} e^{i2k(2a+b)}\Big)\\ &= A_3 \sin k\big(x - (2a+b)\big), \qquad \text{for} \qquad a+b < x < 2a+b;\end{aligned} \tag{5.14}$$

where the constants $A_1 = 2ia_1$ and $A_3 = 2ia_3 e^{ik(2a+b)}$ were defined. These functions do not have yet their final form and are written in terms of two different constants. Written as sinusoidal functions of x they vanish at the infinite walls. In the illustrative problems section, we will discuss, with more detail, these functions and the double quantum well eigenfunctions issue.

5.1.2 *Energy Eigenvalues in the Double Quantum Well*

Let us now consider the second alternative. If we replace the coefficients of (5.11) and (5.12) in (5.8), we have

$$\begin{pmatrix} a_3 e^{ik(2a+b)} \\ -a_3 e^{ik(2a+b)} \end{pmatrix} = \begin{pmatrix} \alpha_b e^{i2ka} & \beta_b \\ \beta_b^* & \alpha_b^* e^{-i2ka} \end{pmatrix} \begin{pmatrix} a_1 \\ -a_1 \end{pmatrix}, \tag{5.15}$$

which means that

$$\alpha_b e^{2ika} - \beta_b = -\Big(\beta_b^* - \alpha_b^* e^{-2ika}\Big), \tag{5.16}$$

or equivalently, that

$$\mathrm{Im}\big(\alpha_b e^{2ika} - \beta_b\big) = 0. \tag{5.17}$$

We got here an important equation, the double-well energy eigenvalues equation. This equation came out in a very simple way, because of the transfer matrix representation. It would not be so easy to derive using other methods. Taking into account the explicit

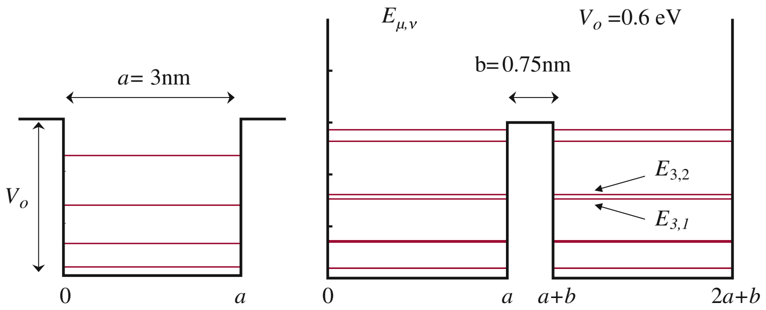


Fig. 5.2 Splitting and repulsion of the energy levels in the double well potential, bounded by infinite walls. In the *left side* the eigenvalues of a single well, with the same well depth V_o and the same width a, as in the double well, are shown as reference. In the double quantum well, the well's interference and the phase coherence are responsible for the energy levels splitting. We use the notation $E_{\mu,\nu}$ for the energy eigenvalues. The first index labels the levels before they split, and the index ν the split levels. $E_{1,1}$ and $E_{1,2}$ are the closest energy levels. The infinite walls push the energy levels a little up

functions of α_b and β_b, the eigenvalues equation takes the form

$$\cosh qb\ \sin 2ka + \frac{k^2 - q^2}{2qk} \sinh qb\ \cos 2ka = -\frac{k^2 + q^2}{2qk} \sinh qb. \tag{5.18}$$

This analytical expression is also an implicit equation that can be solved numerically. In Fig. 5.2 we plot the first energy levels. To visualize the double well and the infinite walls effect on the energy eigenvalues, we show also the eigenvalues of a single quantum well, with the same potential parameters, depth $V_o = 0.6\,\text{eV}$ and well width $a = 30\,\text{Å}$. The most evident feature, of the double well spectrum, is the energy levels splitting. The splitting is better perceived at the higher energy levels. It can be seen that the split energy levels lie almost at the same positions as the energy levels in the single well. The double well energy levels are displaced slightly upwards, because of the infinite walls repulsion. This is a clear example of energy levels splitting. With the examples that come next, we will better understand this phenomenon.

When the number of wells and barriers increases, the levels splitting will lead also to an increasing number of energy levels. Each energy level of the single quantum well (QW) will become a band in a multiple quantum well (MQW) system. Since each band contains as many levels as wells has the multiple quantum well system, it is convenient to label the energy eigenvalues as $E_{\mu,\nu}$, with two indices. The first index denotes the band and the second the intra-band levels. In the double well system we have $\mu = 1, 2, 3, \ldots$ and $\nu = 1, 2$. The lowest energy pair in Fig. 5.2 corresponds to $E_{1,1}$ and $E_{1,2}$, we then have the pair of energy levels, $E_{2,1}$ and $E_{2,2}$, and so on. Therefore, solving the Schrödinger equation (5.18) for a MQW system, one obtains a set of energy eigenvalues $\{E_{\mu,\nu}\}$. The energy eigenvalues represent an important part of the solution, another important part is the calculation of the eigenfunctions

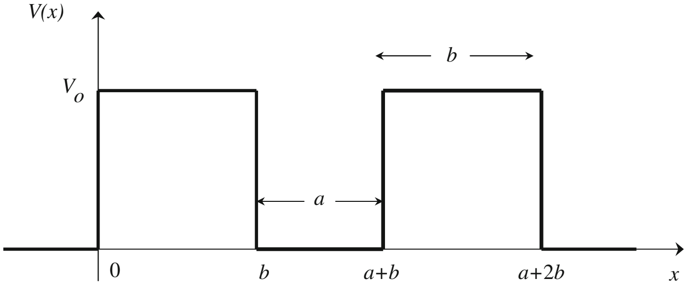


Fig. 5.3 A double rectangular barrier with its parameters

$\{\varphi_{\mu,\nu}(x)\}$. To determine the eigenfunctions one can also use the transfer matrix method. At the end of this chapter we will discuss this issue with an illustrative example.

5.2 The Double Rectangular Potential Barrier

The double potential barrier, as the one shown in Fig. 5.3, is another of the most studied systems, with many applications in optoelectronic and electronic transport. This is an interesting system. It contains not only a confining potential that tends to trap particles, but also finite-width barriers that particles tunnel to enter and to escape from the confining region. Solving this system and calculating the basic transport properties we will find an interesting physical behavior.[2] As in the single barrier case, some of the incoming particles, approaching the double barrier from the left or the right side, will get transmitted. The transmitted and reflected particles will in general suffer multiple reflections. Important information on this process can be obtained from the transmission and reflection coefficients.

5.2.1 *Continuity and the Double-Barrier Transfer Matrix*

We know that the state vectors at points just outside a rectangular potential barrier, as those in Fig. 5.3, are related by the transfer matrix M_b given in (4.87) for energies below the barrier height V_o, or in (4.108) for energies above the barrier height V_o. Those transfer matrices were written in a compact form as

$$M_b = \begin{pmatrix} \alpha_b & \beta_b \\ \beta_b^* & \alpha_b^* \end{pmatrix}. \tag{5.19}$$

---

[2] In the double barrier systems of the electronic devices, the potential function contains also the bias potential energy Fx.

This means that

$$\phi(b^+) = \begin{pmatrix} a_3 e^{ikb} \\ b_3 e^{-ikb} \end{pmatrix} = \begin{pmatrix} \alpha_b & \beta_b \\ \beta_b^* & \alpha_b^* \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix},$$

$$= \begin{pmatrix} \alpha_b & \beta_b \\ \beta_b^* & \alpha_b^* \end{pmatrix} \phi(0^-), \tag{5.20}$$

and also that

$$\phi(a+2b^+) = \begin{pmatrix} a_5 e^{ik(a+2b)} \\ b_5 e^{-ik(a+2b)} \end{pmatrix} = \begin{pmatrix} \alpha_b & \beta_b \\ \beta_b^* & \alpha_b^* \end{pmatrix} \begin{pmatrix} a_3 e^{ik(a+b)} \\ b_3 e^{-ik(a+b)} \end{pmatrix},$$

$$= \begin{pmatrix} \alpha_b & \beta_b \\ \beta_b^* & \alpha_b^* \end{pmatrix} \phi(a+b^-). \tag{5.21}$$

The specific functions α_b and β_b were given in (4.87) and (4.108). To obtain the transfer matrix of the whole double barrier system, we need still to connect $\phi(a+b^-)$ with $\phi(b^+)$. Since both vectors are defined at points of the same constant potential region, they are related by a diagonal transfer matrix, i.e. related as follows

$$\phi(a+b^-) = \begin{pmatrix} a_3 e^{ik(a+b)} \\ b_3 e^{-ik(a+b)} \end{pmatrix} = \begin{pmatrix} e^{ika} & 0 \\ 0 & e^{-ika} \end{pmatrix} \begin{pmatrix} a_3 e^{ikb} \\ b_3 e^{-ikb} \end{pmatrix},$$

$$= \begin{pmatrix} e^{ika} & 0 \\ 0 & e^{-ika} \end{pmatrix} \phi(b^+). \tag{5.22}$$

Having this matrix, it is clear that

$$\begin{pmatrix} a_5 e^{ik(2b+a)} \\ b_5 e^{-ik(2b+a)} \end{pmatrix} = \begin{pmatrix} \alpha_b & \beta_b \\ \beta_b^* & \alpha_b^* \end{pmatrix} \begin{pmatrix} e^{ika} & 0 \\ 0 & e^{-ika} \end{pmatrix} \begin{pmatrix} \alpha_b & \beta_b \\ \beta_b^* & \alpha_b^* \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}. \tag{5.23}$$

After multiplying the matrices on the right side of this equation, we obtain the transfer matrix of the double-barrier potential. One can easily verify that it will have the same structure as the transfer matrix of a single barrier, i.e., it will be of the form

$$M_D = \begin{pmatrix} \alpha_D & \beta_D \\ \beta_D^* & \alpha_D^* \end{pmatrix}, \tag{5.24}$$

with

$$\alpha_D = \alpha_b^2 e^{ika} + |\beta_b|^2 e^{-ika} \qquad \text{and} \qquad \beta_D = \alpha_b \beta_b e^{ika} + \alpha_b^* \beta_b e^{-ika}. \tag{5.25}$$

Having the transfer matrix, we can obtain information on the transport properties.

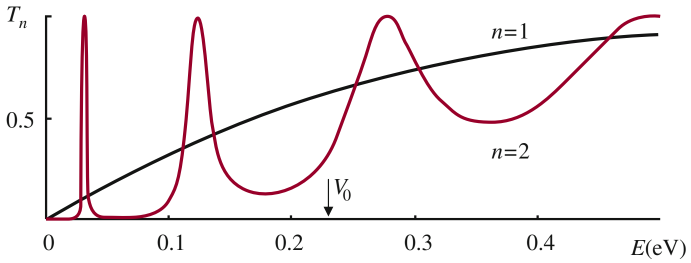


Fig. 5.4 The transmission coefficient of the double barrier potential (*red line*) and a single barrier (*black line*) when the potential parameters are $a = 2\,\text{nm}$, $b = 10\,\text{nm}$ and $V_o = 0.23\,\text{eV}$

5.2.2 *Transport Properties in the Double Rectangular Potential Barrier*

We saw before in our brief reference to the scattering theory, and in the examples studied so far, that the transmission and the reflection amplitudes are related to the transfer matrix elements of the scattering system. Using these relations, the transmission and reflection amplitudes of the double potential barrier, t_D and r_D, are

$$t_D = \frac{1}{\alpha_D^*} \qquad \text{and} \qquad r_D = -\frac{\beta_D^*}{\alpha_D^*}. \tag{5.26}$$

Since $T_D = |t_D|^2$, the double barrier transmission coefficient is given by

$$T_D = \frac{1}{|\alpha_b|^4 + |\beta_b|^4 + 2\,|\beta_b|^2\,\Re e(\alpha_b^2 e^{i2ka})} \equiv T_2, \tag{5.27}$$

where $\Re e$ stands for the real part. In Fig. 5.4 we plot this coefficient as a function of the incoming energy together with the single barrier transmission coefficient T_1. The double barrier transmission coefficient, unlike the single barrier coefficient T_1 (characterized by a monotonic behavior), presents a resonant behavior and reaches at the resonant energies the maximum value 1, even though the energy is below V_o. To explain the origin of this resonant behavior, it is helpful to write the previous equation in the form

$$T_D = \frac{1}{1 + 2\,|\beta_b|^2\left(|\alpha_b|^2 + \Re e\ \alpha_b^2 e^{i2ka}\right)}. \tag{5.28}$$

It is now evident that the transmission coefficient oscillates and becomes equal to 1 when

$$|\alpha_b|^2 + \Re e\ \alpha_b^2 e^{i2ka} = 0. \tag{5.29}$$

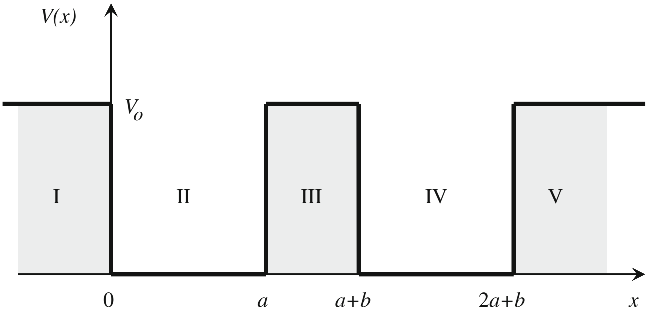


Fig. 5.5 The double well with finite lateral walls

The roots of this equation, which we denote as E^*_μ with $\mu = 1, 2 \ldots$, are the resonant energies. These energies correspond to the confining energies in the quantum well between the two potential barriers. Unlike the true energy eigenvalues, which are stable states, the resonant states are *meta-stable* states. The possibility of tunneling through the barrier allows the particles to enter and to escape. The resonant transmission needs not only the tunneling effect, requires also the confining potential with quasi-bounded or trapping states, that we will call resonant states.

When the incoming particle energy coincides with a resonant energy (i.e. with the resonant energies E^*_μ), the particle is allowed not only to stay in the resonant state, it can also use the state to pass through the system as if the potential barriers did not exist, i.e. with transmission probability equal to 1! Otherwise, the particle will hardly cross the system. In fact, if we look at Fig. 5.4, we can see that for energies other than the resonant ones the transmission probability is low, much smaller than the transmission probability through a single barrier: it tends to zero. Notice also that the resonant behavior occurs also for energies greater than the barriers height, which in our example is $V_o = 0.23\,\mathrm{eV}$ (see the arrow in the figure), these resonant states are called resonant states in the continuum. To conclude this section, it is worth mentioning that in the double barrier problem there is no energy quantization. The incident particles energy has no restriction and can take any positive value.

5.3 The Finite Double Quantum Well

Let us now study the double quantum well shown in Fig. 5.5. This is the simplest example of the so-called multiple quantum well (MQW) systems. Although the results, at the end are independent of the choice of origin, it will be convenient in this case to choose the zero of the energy E as in the simple quantum well problem studied in Chap. 4, i.e. at the bottom of the potential wells. Once we choose the origin, we must be careful and consistent with this choice.

5.3.1 Continuity and the Double-Well Transfer Matrix

In the simple quantum well case, for energies smaller than the barrier height ($E < V_o$), we had the transfer matrix

$$M_a(a^+,0^-) = \begin{pmatrix} \cos ka + \dfrac{q^2-k^2}{2qk}\sin ka & -\dfrac{k^2+q^2}{2qk}\sin ka \\ \dfrac{k^2+q^2}{2qk}\sin ka & \cos ka - \dfrac{q^2-k^2}{2qk}\sin ka \end{pmatrix}, \\ \equiv \begin{pmatrix} \alpha_a & \beta_a \\ -\beta_a & \delta_a \end{pmatrix}, \tag{5.30}$$

to connect the state vectors at points just outside of the well, at the left and right sides. This matrix, with the corresponding change $q \to ik_2$ is also valid when the energy is larger than V_o. In any case, whether the energy is less than or greater than V_o, the transfer matrix $M_a(a^+, 0^-)$ is the same as the transfer matrix $M_a(2a+b^+, a+b^-)$, that connects state vectors at $a+b^-$ and $2a+b^+$. Thus, we have the same matrix in

$$\begin{pmatrix} a_3 e^{qa^+} \\ b_3 e^{-qa^+} \end{pmatrix} = \begin{pmatrix} \alpha_a & \beta_a \\ -\beta_a & \delta_a \end{pmatrix}\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, \tag{5.31}$$

as in

$$\begin{pmatrix} a_5 e^{q(2a+b^+)} \\ b_5 e^{-q(2a+b^+)} \end{pmatrix} = \begin{pmatrix} \alpha_a & \beta_a \\ -\beta_a & \delta_a \end{pmatrix}\begin{pmatrix} a_3 e^{q(a+b^-)} \\ b_3 e^{-q(a+b^-)} \end{pmatrix}. \tag{5.32}$$

Again, if we want to connect $\phi_{\text{V}}(2a+b^+)$ with $\phi_{\text{I}}(0^-)$, we still need to connect the state vectors $\phi_{\text{III}}(a+b^-)$ and $\phi_{\text{III}}(a^+)$. Since these vectors are in the same constant-potential region, we have

$$\begin{pmatrix} a_3 e^{q(a+b)} \\ b_3 e^{-q(a+b)} \end{pmatrix} = \begin{pmatrix} e^{qb} & 0 \\ 0 & e^{-qb} \end{pmatrix}\begin{pmatrix} a_3 e^{qa} \\ b_3 e^{-qa} \end{pmatrix}. \tag{5.33}$$

With this relation, we are ready to connect $\phi_{\text{V}}(2a+b^+)$ with $\phi_{\text{I}}(0^-)$. Indeed, if we use (5.33) and (5.31) in (5.32), we have

$$\begin{pmatrix} a_5 e^{q(2a+b)} \\ b_5 e^{-q(2a+b)} \end{pmatrix} = \begin{pmatrix} \alpha_a & \beta_a \\ -\beta_a & \delta_a \end{pmatrix}\begin{pmatrix} e^{qa} & 0 \\ 0 & e^{-qa} \end{pmatrix}\begin{pmatrix} \alpha_a & \beta_a \\ -\beta_a & \delta_a \end{pmatrix}\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, \tag{5.34}$$

that, after multiplying in the right hand side, yields

$$\begin{pmatrix} a_5 e^{q(2a+b)} \\ b_5 e^{-q(2a+b)} \end{pmatrix} = \begin{pmatrix} \alpha_D & \beta_D \\ -\beta_D & \delta_D \end{pmatrix}\begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, \tag{5.35}$$

with

$$\begin{aligned}\alpha_D &= \alpha_a^2 e^{qb} - \beta_a^2 e^{-qb},\\ \beta_D &= \alpha_a\beta_a e^{qb} + \beta_a\delta_a e^{-qb}, \qquad (5.36)\\ \delta_D &= -\beta_a^2 e^{qb} + \delta_a^2 e^{-qb}. \qquad (5.37)\end{aligned}$$

Notice that the matrix structure is the same as that of M_a.

When the incident particles energy is larger than the height of the potential barriers, i.e. when $E > V_o$, we can obtain the double quantum well transfer matrix, exactly in the same way as for energies below V_o. It is clear that in this case we have to use the transfer matrix of the finite rectangular quantum well, that was given as

$$M_a = \begin{pmatrix}\alpha_a & \beta_a\\ \beta_a^* & \alpha_a^*\end{pmatrix} = \begin{pmatrix}\cos ka + i\dfrac{k^2+k_2^2}{2kk_2}\sin ka & i\dfrac{k^2-k_2^2}{2kk_2}\sin ka\\ -i\dfrac{k^2-k_2^2}{2kk_2}\sin ka & \cos ka - i\dfrac{k^2+k_2^2}{2kk_2}\sin ka\end{pmatrix}. \qquad (5.38)$$

Notice that this matrix connects also the state vector at $x = a + b^-$ with the state vector at $x = 2a + b^+$. We still need a matrix to connect the state vectors at $x = a^+$ and $x = a + b^-$. This matrix comes from

$$\begin{pmatrix}a_3 e^{ik_2(a+b)}\\ b_3 e^{-ik_2(a+b)}\end{pmatrix} = \begin{pmatrix}e^{ik_2 b} & 0\\ 0 & e^{-ik_2 b}\end{pmatrix}\begin{pmatrix}a_3 e^{ik_2 a}\\ b_3 e^{-ik_2 a}\end{pmatrix}. \qquad (5.39)$$

We can now establish the connection between the state vectors just outside (at the outer ends of) the double quantum well for energies above the barrier height. This connection is given as follows

$$\begin{pmatrix}a_5 e^{ik_2(2a+b)}\\ b_5 e^{-ik_2(2a+b)}\end{pmatrix} = \begin{pmatrix}\alpha_a & \beta_a\\ \beta_a^* & \alpha_a^*\end{pmatrix}\begin{pmatrix}e^{ik_2 b} & 0\\ 0 & e^{-ik_2 b}\end{pmatrix}\begin{pmatrix}\alpha_a & \beta_a\\ \beta_a^* & \alpha_a^*\end{pmatrix}\begin{pmatrix}a_1\\ b_1\end{pmatrix}. \qquad (5.40)$$

Notice that the variable q, in α_a, β_a and in the diagonal matrix of Eq. (5.34), is replaced here by ik_2. After multiplying, the transfer matrix of the double quantum well becomes

$$M_D = \begin{pmatrix}\alpha_D & \beta_D\\ \beta_D^* & \alpha_D,\end{pmatrix}, \qquad (5.41)$$

with

$$\alpha_D = \alpha_a^2 e^{ik_2 b} + |\beta_a|^2 e^{-ik_2 b},$$

$$\beta_D = \alpha_a\beta_a e^{ik_2 b} + \alpha_a^*\beta_a e^{-ik_2 b}. \qquad (5.42)$$

Given the transfer matrices $M_D(2a+b^+, 0^-)$ and the possibility of defining transfer matrices like $M_x(x, -\infty)$, from $-\infty$ to any point x, one can determine the coef-

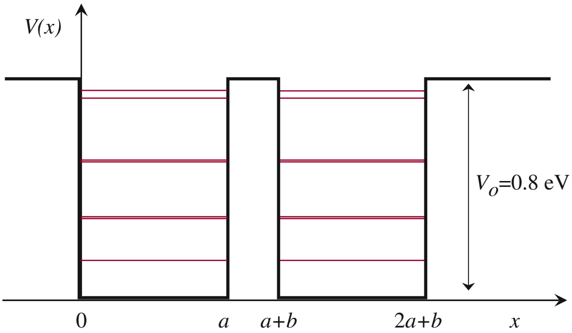


Fig. 5.6 Eigenvalues $E_{\mu,\nu}$ of the double quantum well with finite lateral walls. The levels $E_{\mu,1}$ and $E_{\mu,2}$, with the exception of $E_{5,1}$ and $E_{5,2}$, are so close to each other, that the levels splitting can hardly be seen

ficients a_i and b_i and evaluate the wave functions inside and outside the double quantum well.

5.3.2 *Eigenvalues and Eigenfunctions in a Double Quantum Well*

From the general properties mentioned before, we expect that the energy in the double quantum well will quantize only for energies below the barrier height, i.e. for $E < V_o$. The energy eigenvalues equation appears naturally when new physical conditions are imposed. In fact, when $E < V_o$ and $x \rightarrow \pm\infty$, some terms of the wave functions

$$\varphi_{\mathrm{I}}(x) = a_1 e^{qx} + b_1 e^{-qx} \qquad \text{and} \qquad \varphi_{\mathrm{V}}(x) = a_5 e^{qx} + b_5 e^{-qx}. \tag{5.43}$$

diverge. Thus, the finiteness condition requires that $b_1 = a_5 = 0$. If we introduce these coefficients into (5.34), we have

$$\begin{pmatrix} 0 \\ b_5 e^{-q(2b+a)} \end{pmatrix} = \begin{pmatrix} \alpha_D & \beta_D \\ -\beta_D & \delta_D \end{pmatrix} \begin{pmatrix} a_1 \\ 0 \end{pmatrix}, \tag{5.44}$$

which means that

$$\alpha_D = 0, \tag{5.45}$$

and

$$b_5 = -a_1 \beta_D e^{q(2a+b)}. \tag{5.46}$$

The first of these equations is the energy eigenvalues equation. If we substitute the explicit functions obtained before for α_a and β_a, the eigenvalues equation in the symmetric double quantum well takes the form:

$$\left(\cos ka + \frac{k^2 - q^2}{2qk}\sin ka\right)^2 e^{qb} - \left(\frac{k^2 + q^2}{2qk}\right)^2 \sin^2 ka\, e^{-qb} = 0. \tag{5.47}$$

In Fig. 5.6, we plot some energy eigenvalues $E_{\mu\nu}$ when the double well parameters are $V_o = 0.8\,\text{eV}$, $a = 3\,\text{nm}$ and $b = 1\,\text{nm}$. We have similar results to those of the double well with infinite walls. Though it is hard to distinguish the split, the energy levels come also by pairs. These energy eigenvalues are denoted also with two indices, i.e. as $E_{\mu,\nu}$.

Let us now study the eigenfunctions of the double quantum well. We want to know whether they are localized in one well or extended along the two wells and the barrier.[3] To obtain the double well eigenfunctions, we need first to express the state vectors as explicit functions of x and in terms of only one coefficient, for example the coefficient a_1, which might be later fixed through a normalization requirement. If we use the relations (5.31), (5.32) and (5.46), the transition matrices defined before and the coefficients $b_1 = a_5 = 0$ and $b_5 = -a_1\beta_D e^{q(2a+b)}$, we can obtain state vectors, wave functions and eigenfunctions of the double quantum well. The student can easily verify the following results: in region I, for $x < 0$, we have

$$\phi_{\text{I}}(x) = \begin{pmatrix} e^{qx} & 0 \\ 0 & e^{-qx} \end{pmatrix}\begin{pmatrix} a_1 \\ 0 \end{pmatrix} \quad\Rightarrow\quad \varphi_{\text{I}}(x) = a_1 e^{qx}; \tag{5.48}$$

in region II, which corresponds to the left-side well, with $0 < x < a$, we have

$$\begin{aligned} \phi_{\text{II}}(x) &= \begin{pmatrix} e^{ikx} & 0 \\ 0 & e^{-ikx} \end{pmatrix}\frac{1}{2k}\begin{pmatrix} k - iq & k + iq \\ k + iq & k - iq \end{pmatrix}\phi_{\text{I}}(0), \\ &\Downarrow \\ \varphi_{\text{II}}(x, E) &= a_1(\cos kx + \frac{q}{k}\sin kx); \end{aligned} \tag{5.49}$$

inside the barrier, for $a < x < a + b$, we have

$$\begin{aligned} \phi_{\text{III}}(x) &= \begin{pmatrix} e^{q(x-a)} & 0 \\ 0 & e^{-q(x-a)} \end{pmatrix}\frac{1}{2q}\begin{pmatrix} q + ik & q - ik \\ q - ik & q + ik \end{pmatrix}\phi_{\text{II}}(a), \\ &\Downarrow \\ \varphi_{\text{III}}(x, E) &= a_1\left[\cos ka\, e^{q(x-a)} + \sin ka\left(\frac{q}{k}\cosh q(x-a) - \frac{k}{q}\sinh q(x-a)\right)\right]; \end{aligned} \tag{5.50}$$

in the right well, with $a + b < x < 2a + b$, we have

---

[3] Quite frequently one finds, in the scientific literature, approximate double well eigenfunctions. They are generally built with the eigenfunctions of the single wells and, when the barrier width is large, as single well eigenfunctions.

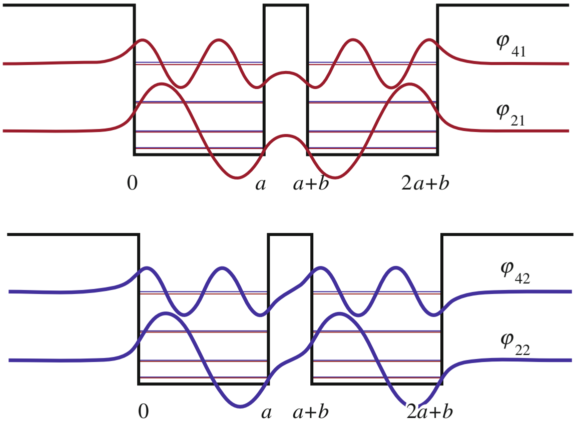


Fig. 5.7 Some eigenfunctions of the double well bounded by finite walls. We plot here the eigenfunctions $\varphi_{2,1}$ and $\varphi_{2,2}$, corresponding to the eigenvalues $E_{2,1}$ and $E_{2,2}$, and the eigenfunctions $\varphi_{4,1}$ and $\varphi_{4,2}$ corresponding to the energy levels $E_{4,1}$ and $E_{4,2}$. The eigenfunctions describe extended states with well defined parity and they penetrate in the lateral barriers

$$\phi_{\text{IV}}(x) = \begin{pmatrix} e^{ik(x-a-b)} & 0 \\ 0 & e^{-ik(x-a-b)} \end{pmatrix} \frac{1}{2k} \begin{pmatrix} k-iq & k+iq \\ k+iq & k-iq \end{pmatrix} \phi_{\text{III}}(a+b),$$
$$\Downarrow$$
$$\varphi_{\text{IV}}(x,E) = a_1 \left[(\alpha_a+\beta_a)e^{qb}\cos k(x-a-b) + \frac{k}{q}(\alpha_a-\beta_a)\sin k(x-a-b) \right]; \tag{5.51}$$

and finally, outside the double well on the right, for $x > 2a+b$, we have

$$\phi_{\text{V}}(x) = \begin{pmatrix} 0 \\ b_5 e^{-q(x-2a-b)} \end{pmatrix} \quad \text{with} \quad b_5 = -a_1\beta_D e^{q(2a+b)},$$
$$\Downarrow$$
$$\varphi_{\text{V}}(x,E) = -a_1\beta_D e^{-q(x-2(2a+b))}. \tag{5.52}$$

The eigenfunctions $\varphi_{\mu\nu}(z)$ are obtained when these functions are evaluated at the corresponding energy eigenvalues $E_{\mu\nu}$. In Fig. 5.7 we plot some of these eigenfunctions. The eigenfunctions of the split pairs, have different symmetries and different parities. They are extended states along the double well system. This means that a particle in the double well, can be found, in principle, everywhere. Even in the lateral barriers, with exponentially decreasing probability.

As noticed earlier, the split is a consequence of the coherent interference. If the well widths are not equal, one has an asymmetric double well. Also in this case one can solve the problem using the transfer matrix method. As the symmetry is lost the level splitting becomes rather irregular.

5.3.3 Transport Properties in the Double Quantum Well

When the energies are larger than the barrier's height, i.e. for $E > V_o$, the wave functions are everywhere oscillating functions and propagate in both directions. For these energies, it is possible to talk of reflection and transmission coefficients. If the incidence is only from the left, we have to take $b_5 = 0$. In this case

$$\begin{pmatrix} a_5 e^{ik_2(2a+b)} \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_D & \beta_D \\ \beta_D^* & \alpha_D^* \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}. \tag{5.53}$$

This relation is equivalent to

$$b_1 = -\frac{\beta_D^*}{\alpha_D^*} a_1 = r_D a_1, \tag{5.54}$$

$$\begin{aligned} a_5 &= \left(\alpha_D - \beta_D \frac{\beta_D^*}{\alpha_D^*}\right) a_1 e^{-ik_2(2a+b)} = \frac{1}{\alpha_D^*} a_1 e^{-ik_2(2a+b)} \\ &= t_D\, a_1\, e^{-ik_2(2a+b)}. \end{aligned} \tag{5.55}$$

If we replace these coefficients into $\varphi_{\rm I}(x)$ and $\varphi_{\rm V}(x)$, we have

$$\varphi_{\rm I}(x) = a_1 e^{ik_2 x} + a_1\, r_D\, e^{-ik_2 x}, \tag{5.56}$$

and

$$\varphi_{\rm V}(x) = a_1\, t_D\, e^{ik_2[x-(2a+b)]}. \tag{5.57}$$

The wave function $\varphi_{\rm I}(x)$, at the left, is the sum of the incident $e^{ik_2 x}$ and the reflected $r_D\, e^{-ik_2 x}$ wave functions. On the other hand, the transmitted wave function $\varphi_{\rm V}(x)$ is proportional to the transmission amplitude

$$\begin{aligned} t_D &= \frac{1}{\alpha_{DR} - i\alpha_{DI}}, \\ &= \frac{1}{\left(\cos ka + i\dfrac{k_2^2 + k^2}{2kk_2} \sin\, ka\right)^2 e^{ik_2 b} + \left(\dfrac{k^2 - k_2^2}{2kk_2}\right)^2 \sin^2 ka\; e^{-ik_2 b}}, \end{aligned} \tag{5.58}$$

with $k_2 = \sqrt{2m(E - V_o}/\hbar$ and $k = \sqrt{2mE}/\hbar$. Therefore, when $E > V_o$, the transmission coefficient through the double well potential is given by

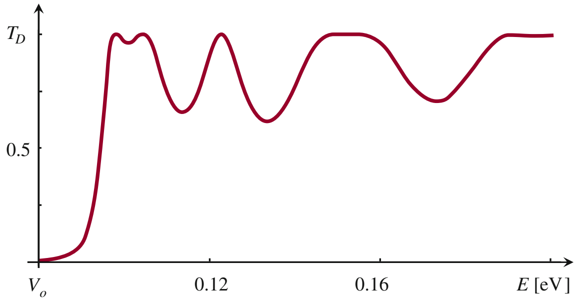


Fig. 5.8 Transmission coefficient of the double well when the energy is greater than the height of potential barriers $V_o = 0.8\,\text{eV}$. The parameters are the same as in Fig. 5.6

$$T_D = \frac{1}{\alpha_{DR}^2 + \alpha_{DI}^2}, \tag{5.59}$$

with

$$\alpha_{DR} = \cos 2ka\ \cos k_2 b - \frac{k^2 + k_2^2}{2kk_2} \sin 2ka\ \sin k_2 b, \tag{5.60}$$

$$\begin{aligned}\alpha_{DI} = &\frac{k^2 + k_2^2}{2kk_2} \sin 2ka\ \cos k_2 b - \left(\frac{k^2 - k_2^2}{2kk_2}\right)^2 \sin k_2 b \\ &+ \left(\frac{k^2 + k_2^2}{2kk_2}\right)^2 \cos 2ka\ \sin k_2 b.\end{aligned} \tag{5.61}$$

The complexity of the function $\alpha_{DR}^2 + \alpha_{DI}^2$ in the denominator of T_D is significantly higher than for a single well. In Fig. 5.8 we plot this transmission coefficient for $V_o = 0.8\,\text{eV}$, $a = 3\,\text{nm}$ and $b = 1\,\text{nm}$. The transmission probability increases with the energy and also has a resonant behavior.

So far we have studied an important set of soluble examples. In each of these examples we found new results, which altogether, configure a basic picture of the quantum phenomenology. In Chap. 4 we met the quantum tunneling and the resonant transmission. In this chapter, we found the quantum coherence phenomenon and the energy levels splitting. The physical systems which behavior comprises all this phenomenology, are the periodic systems. Systems such as metals, semiconductors and other crystalline structures, natural or artificial, are examples of periodic systems. To get an insight into these systems we will briefly discuss some results related with the one-dimensional finite periodic structures. We will study the finite Kronig–Penney model, with particular emphasis on the splitting phenomenon that builds up the energy bands.

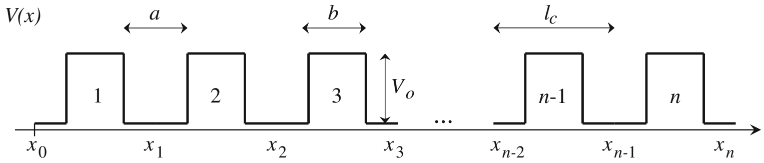


Fig. 5.9 A finite periodic system with rectangular barriers and wells

5.4 Finite Periodic Systems

Our interest in this section is to present an introductory application of the quantum formalism to periodic systems. The natural extension of the simple structures studied before, characterized by the level splitting phenomenon, are the one-dimensional periodic systems and their energy bands. We have seen that going from the one barrier system to the double barrier system, or from a single well to the double well, the complexity of the transfer matrices grows. It is natural to think that the transfer matrix of a periodic system, will necessarily imply higher levels of complexity and perhaps more involved expressions. that is however not the case. We will show here that, using properly the transfer matrix properties and symmetries, it is possible to obtain compact and universal expressions for the transfer matrix of a finite periodic system. These results, which are independent of the specific potential profile, can equally be used to obtain the transfer matrix of a piecewise constant periodic potential, as in Fig. 5.9, or the transfer matrix of a periodic system with an arbitrary potential profile, like the one shown in Fig. 5.10.

Having the transfer matrix of a periodic system, we will be able to derive general formulas for the transmission and reflection coefficients, in the same way as we did for the simple examples studied before. We will apply these formulas for the finite periodic system, with n rectangular barriers and $n-1$ rectangular wells, shown in Fig. 5.9, and we will see how the level splitting leads to one of the most significant quantum properties: the band structure of periodic systems. This system, in the limit $n = \infty$, is the well known Kronig–Penney model.[4] The Kronig–Penney model was one of the first simple models that revealed the existence of the allowed and forbidden energy bands. Periodicity is one of the most distinctive features of crystal structures.

The periodic potential in Figs. 5.9 and 5.10 are quite different but they share a common property: both are built by the repetition of a unit cell. The unit cells of these systems are shown in Fig. 5.11. If the length of a unit cell is l_c, the length of the n cells system is $L = nl_c$. To obtain the physical quantities of a finite periodic system, with n cells, we need the transfer matrix $M_n(x_r, x_l)$ that connects the state vectors at the left end $x_l = x_0$ with those at the right end $x_r = x_n = x_0 + nl_c$. It turns out and will be shown here that, to obtain this transfer matrix, it is necessary and sufficient to know the unit-cell transfer matrix M. We will show that assuming the existence of M, without even knowing the specific functions of $\alpha, \beta, \ldots$, we can

---

[4] Kronig, R. d. L. and Penney, W. G. Proc. Roy. Soc. A **130** 499 (1931) .

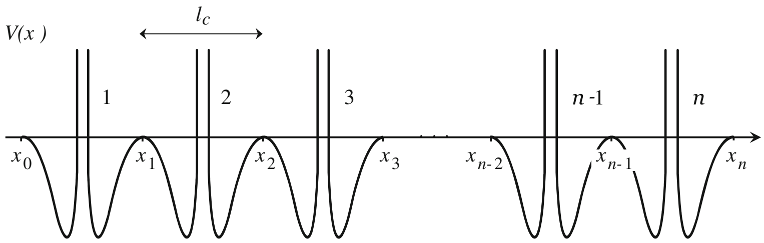
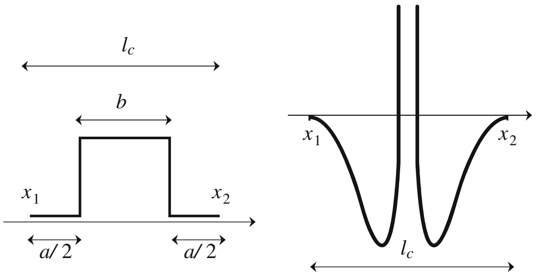


Fig. 5.10 A finite periodic system with an arbitrary potential profile

Fig. 5.11 Unit cells of the periodic systems in Figs. 5.9 and 5.10



derive general expressions for $M_n(x_r, x_l)$ in terms of α and β. The calculation of the unit-cell transfer matrix can be a simple or a complex task. When the potential is piecewise constant, as in Fig. 5.9, the solution of the Schrödinger equation of the unit cell is very simple. When the potential profile is as in Fig. 5.10, with a unit cell like the one shown on the right side of Fig. 5.11, the calculation of the unit cell transfer matrix may be a cumbersome problem. Suppose now that

$$M = \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix}, \tag{5.62}$$

is the transfer matrix of a unit cell.[5] If we use the multiplicative property of transfer matrices, we can write the transfer matrix of the sequence of n-cells as

$$M_n = \underbrace{M.M \dots M}_{n \text{ factors}} = M^n. \tag{5.63}$$

This product of matrices can be written in different ways, for example we can write M_n as the product of M_{n-1} with the matrix M, i.e. as

$$M_n = M \underbrace{M.M \dots M}_{n-1 \text{ factors}} = M M_{n-1}, \tag{5.64}$$

---

[5] Notice that to simplify the calculations we are also assuming that the unit cell is invariant under time reversal.

or, as the product

$$M_n = \underbrace{M.M \ldots M}_{n-1 \text{ factors}} M = M_{n-1}M. \tag{5.65}$$

In terms of the matrix elements, this product looks like

$$\begin{pmatrix} \alpha_n & \beta_n \\ \beta_n^* & \alpha_n^* \end{pmatrix} = \begin{pmatrix} \alpha_{n-1} & \beta_{n-1} \\ \beta_{n-1}^* & \alpha_{n-1}^* \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix}. \tag{5.66}$$

Our purpose is to obtain α_n and β_n, provided that α and β are known. From (5.66), we have[6]

$$\alpha_n = \alpha_{n-1}\alpha + \beta_{n-1}\beta^*, \tag{5.67}$$

and

$$\beta_n = \alpha_{n-1}\beta + \beta_{n-1}\alpha^*. \tag{5.68}$$

If we solve for α_{n-1}, in the last equation, we have

$$\alpha_{n-1} = \beta^{-1}\beta_n - \alpha^*\beta^{-1}\beta_{n-1}. \tag{5.69}$$

It is clear here that defining the function

$$p_{n-1} = \beta^{-1}\beta_n, \quad \rightarrow \quad \beta_n = \beta p_{n-1}, \tag{5.70}$$

(5.69) can be rewritten as

$$\alpha_{n-1} = p_{n-1} - \alpha^* p_{n-2} \quad \text{or} \quad \alpha_n = p_n - \alpha^* p_{n-1}. \tag{5.71}$$

In the last two equations we have α_n and β_n in terms of p_n and the matrix elements of M, i.e. of α and β. If we know the functions p_n we will formally solve the problem, and the transfer matrix M_n will be known. Notice that if we replace (5.70) and the $\alpha_n{}'s$ of the last equations in (5.67), we end up with the interesting three terms recurrence relation

$$p_n - (\alpha + \alpha_*)\, p_{n-1} + p_{n-2} = 0. \tag{5.72}$$

To solve this equation we need to define the initial conditions. Taking into account that $M^0 = I$, $M^1 = M$ and the relation (5.70), one can conclude that $p_0 = 1$ and $p_{-1} = 0$. Since $\alpha + \alpha_* = 2\alpha_{\rm R}$, with $\alpha_{\rm R}$ the real part of α, the recurrence relation takes the form

$$p_n - 2\alpha_{\rm R} p_{n-1} + p_{n-2} = 0 \quad \text{with} \quad p_0 = 1 \quad \text{and,} \quad p_{-1} = 0. \tag{5.73}$$

---

[6] In 3D systems or systems with more than one propagating mode, α and β are matrices. This kind of systems are beyond the purpose of this book.

It turns out that this recurrence relation is exactly the same as the recurrence relation

$$U_n - 2xU_{n-1} + U_{n-2} = 0, \tag{5.74}$$

of the Chebyshev polynomials of the second kind $U_n(x)$, with $U_0 = 1$ and $U_{-1} = 0$. Thus, the functions p_n that satisfy (5.73) are the Chebyshev polynomials evaluated at α_R.[7] It is important to notice that, to determine these polynomials and, consequently, the transfer matrix of the whole n-cell system, *it is enough to know the transfer matrix of the unit cell*. Therefore

$$M_n = \begin{pmatrix} p_n - \alpha^* p_{n-1} & \beta p_{n-1} \\ \beta^* p_{n-1} & p_n - \alpha p_{n-1} \end{pmatrix}, \tag{5.75}$$

with α and β the elements of the unit-cell transfer matrix M.

This is an important result that relieves us of multiplying matrices.[8] Since the transmission amplitude t_n is the inverse of α_n^*, the transmission coefficient of the whole system is given in general as

$$T_n = \frac{1}{|p_n - \alpha^* p_{n-1}|^2}. \tag{5.76}$$

If we use the fact that the transfer matrices are unimodular, with $|\alpha_n|^2 = 1 + |\beta_n|^2$, we can also write the transmission coefficient in the form

$$T_n = \frac{1}{1 + |\beta|^2\, p_{n-1}^2}. \tag{5.77}$$

These are the general expressions we wanted to derive. These formulas are valid for any one-dimensional periodic system. Notice that the specific potential profile comes in through the specific functions α and β, and the size of the system through the number of cells n.

5.4.1 Transport Properties in the Kronig–Penney Model

To apply the transport formulas obtained here, we will consider the Kronig–Penney model of rectangular barriers and wells, with a finite number of cells. The unit cell of this system, chosen as in Fig. 5.11, contains a rectangular barrier and two halves of a well (one at the left and one at the right). To obtain the unit cell transfer matrix $M(x_2, x_1)$, with $x_2 = x_1 + l_c$, we need the transfer matrices M_l, M_b and M_r, for the wells and the barrier. For the barrier we have the transfer matrix

---

[7] One can easily verify that $p_1(\alpha_R) = -2\alpha_R$, $p_2(\alpha_R) = 4\alpha_R^2 - 1$, and so on.

[8] See P. Pereyra, *Phys. Rev. Lett* **80** (1998) 2677.

$$M_b = \begin{pmatrix} \alpha_b & \beta_b \\ \beta_b^* & \alpha_b^* \end{pmatrix}, \tag{5.78}$$

with α_b and β_b given in equation (4.89). The transfer matrices M_l and M_r connect points in the same potential region. Since the distance between these points is $a/2$, these transfer matrices are of the form

$$\begin{pmatrix} e^{ika/2} & 0 \\ 0 & e^{-ika/2} \end{pmatrix}. \tag{5.79}$$

Thus, the transfer matrix of the unit cell is obtained from the product

$$M(x_2, x_1) = M_d M_b M_i = \begin{pmatrix} e^{ika/2} & 0 \\ 0 & e^{-ika/2} \end{pmatrix} \begin{pmatrix} \alpha_b & \beta_b \\ \beta_b^* & \alpha_b^* \end{pmatrix} \begin{pmatrix} e^{ika/2} & 0 \\ 0 & e^{-ika/2} \end{pmatrix}, \tag{5.80}$$

which leads us to

$$M(x_2, x_1) = \begin{pmatrix} e^{ika}\alpha_b & \beta_b \\ \beta_b^* & e^{-ika}\alpha_b^* \end{pmatrix}. \tag{5.81}$$

Given this matrix, we can evaluate the transmission coefficient of the Kronig–Penney model using the general formula

$$T_n = \frac{1}{1 + |\beta_b|^2 \, p_{n-1}^2(\alpha_R)}, \tag{5.82}$$

together with the specific functions

$$\beta_b = -i \frac{k^2 + q^2}{2qk} \sinh qb, \tag{5.83}$$

and

$$\alpha_R = \cosh qb \cos ka - \frac{k^2 - q^2}{2qk} \sinh qb \sin ka. \tag{5.84}$$

In Fig. 5.12 we plot the transmission coefficient for different values of n, when $V_o = 0.8\,\text{eV}$, $a = 3\,\text{nm}$ and $b = 1\,\text{nm}$. The most significant result in this series, related with the levels splitting phenomenon, is the appearance of forbidden and allowed energy regions (more sharply defined for larger values of n). This is an example of the energy band structure. It emerges as a result of coherent interference in periodic potentials.

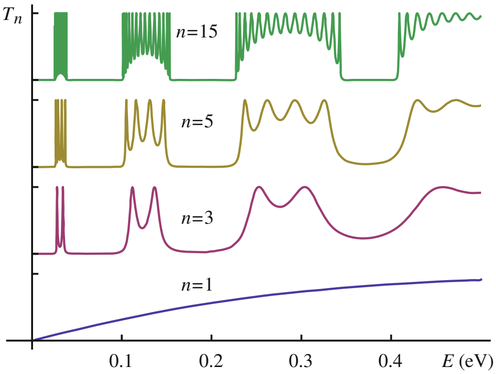


Fig. 5.12 The transmission coefficient T_n of a periodic potential for different number of cells n. We plot here for $n = 1, 3, 5$ and 15, and rectangular barrier parameters $a = 2\,\mathrm{nm}$, $b = 10\,\mathrm{nm}$ and $V_o = 0.23\,\mathrm{eV}$

5.5 Solved Problems

Exercise 14 Using the transfer matrix definition, obtain the wave functions of the double well bounded by infinite walls, and show that they coincide with (5.13) and (5.14). What are the eigenfunctions of this problem?

Solution To start with, let us recall that the wave functions vanish at the infinite walls. Based on this fact, we have

$$\varphi_I(0) = a_1 + b_1 = 0, \qquad \text{hence} \qquad \phi(0) = \begin{pmatrix} a_1 \\ -a_1 \end{pmatrix}. \tag{5.85}$$

From the transfer matrix definition, we know that we can find a transfer matrix $M_p(x, 0)$ such that

$$\phi(x) = \begin{pmatrix} \varphi^+(x) \\ \varphi^-(x) \end{pmatrix} = M_p(x, 0)\phi(0) = \begin{pmatrix} \alpha_p & \beta_p \\ \gamma_p & \delta_p \end{pmatrix} \phi(0), \tag{5.86}$$

with $\phi(x)$ the state vector at any point x within the double well. Thus, the wave function at x will be given by

$$\varphi(x, E) = \varphi^+(x) + \varphi^-(x) = \left(\alpha_p - \beta_p + \gamma_p - \delta_p\right) a_1. \tag{5.87}$$

This function is well defined if we know the matrix elements $\alpha_p, \beta_p, \gamma_p$ and δ_p. Our purpose is then to obtain these matrix elements. We will show now that knowing the solutions of the Schrödinger equation, in each region of the double well, we can obtain the matrix elements α_p, β_p, γ_p and δ_p, hence the wave function $\varphi(x, E)$. Once we know the function, it will be easy to obtain the eigenfunctions $\varphi_{\mu,\nu}(x)$. Let us start with the wave function.

When the point x is in the first region ($0 < x < a$), the transfer matrix $M_p(x, 0)$ connects two points in a constant potential region, thus $\alpha_p(x) = e^{ikx}$, $\beta_p(x) = 0$, $\gamma_p(x) = 0$ and $\delta_p(x) = e^{-ikx}$, and the wave function is

$$\varphi_{\mathrm{I}}(x, E) = 2ia_1 \sin kx. \tag{5.88}$$

This function coincides with the function $\varphi_{\mathrm{I}}(x, E)$ in (5.13). For x in the second region ($a < x < a + b$), we have

$$\begin{pmatrix} a_2 e^{qx} \\ b_2 e^{-qx} \end{pmatrix} = \begin{pmatrix} e^{q(x-a)} & 0 \\ 0 & e^{-q(x-a)} \end{pmatrix} \frac{1}{2q} \begin{pmatrix} q+ik & q-ik \\ q-ik & q+ik \end{pmatrix} \begin{pmatrix} e^{ika} & 0 \\ 0 & e^{-ika} \end{pmatrix} \begin{pmatrix} a_1 \\ -a_1 \end{pmatrix}. \tag{5.89}$$

After multiplying matrices, we obtain the expression

$$\begin{pmatrix} a_2 e^{qx} \\ b_2 e^{-qx} \end{pmatrix} = \frac{1}{2q} \begin{pmatrix} e^{q(x-a)}\,(q+ik)\,e^{ika} & e^{q(x-a)}\,(q-ik)\,e^{-ika} \\ e^{-q(x-a)}\,(q-ik)\,e^{ika} & e^{-q(x-a)}\,(q+ik)\,e^{-ika} \end{pmatrix} \begin{pmatrix} a_1 \\ -a_1 \end{pmatrix}, \tag{5.90}$$

which allows us to identify the matrix elements

$$\begin{aligned} \alpha_p &= \frac{1}{2} e^{q(x-a)} \left(1 + i\frac{k}{q}\right) e^{ika}, & \beta_p &= \frac{1}{2} e^{q(x-a)} \left(1 - i\frac{k}{q}\right) e^{-ika}, \\ \gamma_p &= \frac{1}{2} e^{-q(x-a)} \left(1 - i\frac{k}{q}\right) e^{ika}, & \delta_p &= \frac{1}{2} e^{-q(x-a)} \left(1 + i\frac{k}{q}\right) e^{-ika}. \end{aligned} \tag{5.91}$$

Using these matrix elements we obtain, straightforwardly, the wave function in the second region as

$$\varphi_{\mathrm{II}}(x, E) = 2ia_1 \left(\sin ka \cosh q(x-a) + \frac{k}{q} \cos ka \sinh q(x-a)\right). \tag{5.92}$$

Finally, for the wave function in the third region ($a + b < x < 2a + b$) we need to use the relation

$$\begin{pmatrix} \varphi^+(x) \\ \varphi^-(x) \end{pmatrix} = \begin{pmatrix} e^{ik(x-a-b)} & 0 \\ 0 & e^{-ik(x-a-b)} \end{pmatrix} \begin{pmatrix} \alpha_b & \beta_b \\ \beta_b^* & \alpha_b^* \end{pmatrix} \begin{pmatrix} e^{ika} & 0 \\ 0 & e^{-ika} \end{pmatrix} \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, \tag{5.93}$$

with α_b and β_b the elements of the transfer matrix of the barrier. These matrix elements were given in (4.89). After multiplying, we can also identify the transfer matrix elements

$$\begin{aligned} \alpha_p &= e^{ik(x-a-b)} \alpha_b e^{ika}, & \beta_p &= e^{ik(x-a-b)} \beta_b e^{-ika}, \\ \gamma_p &= e^{-ik(x-a-b)} \beta_b^* e^{ika}, & \delta_p &= e^{-ik(x-a-b)} \alpha_b^* e^{-ika}, \end{aligned} \tag{5.94}$$

which, replaced in (5.87), give the wave function

$$\varphi_{\mathrm{III}}(x, E) = 2ia_1 \Im m \left[e^{ik(x-a-b)} \left(\alpha_b e^{ika} - \beta_b e^{-ika} \right) \right]. \tag{5.95}$$

The notation $\Im m$ stands for the imaginary part. Therefore, the double quantum well wave function (with $A = 2ia_1$) is

$$\varphi(x,E) = \begin{cases} A \sin kx, & 0 < x \le a, \\ A\left(\sin ka \cosh q(x-a) + \frac{k}{q}\cos ka \sinh q(x-a)\right), & a < x \le a+b, \\ A\Im\mathrm{m}\left[e^{ik(x-a-b)}\left(\alpha_b e^{ika} - \beta_b e^{-ika}\right)\right], & a+b < x \le 2a+b. \end{cases} \tag{5.96}$$

We leave as a problem to show that these functions satisfy the continuity conditions at $x = a$ and at $x = a + b$.

We are now ready to obtain the eigenfunctions of the double well. All we have to do is to evaluate the wave $\varphi(x, E)$ at the corresponding energy eigenvalue. The eigenfunction that corresponds to the eigenvalue $E_{\mu,\nu}$ is then

$$\varphi_{\mu,\nu}(x) = \varphi(x,E)|_{E=E_{\mu,\nu}}. \tag{5.97}$$

Exercise 15 Obtain the eigenvalues $E_{\mu\nu}$ of the double quantum well bounded by infinite walls, for $a = 30\,\text{Å}$, $b = 7.5\,\text{Å}$ and $V_o = 0.6\,\text{eV}$.

Solution One way to find the eigenvalues is to use a program that finds the zeros of the implicit function in (5.17). Below we copy a program (written for Mathematica) where the eigenvalues Eq. (5.17) is defined, and the roots of $eigenv = 0$ are found. These roots are the eigenvalues $E_{\mu\nu}$ of the double quantum well with infinite walls. Some of the roots appear more than once. This depends on the numbers given as seeds to generate the table.

```
VO = 0.6;
me = 0.5109906 10^6;
c = 2.99792458 10^18;
ℏ = 6.5821220 10^-16;
```

$$k = \sqrt{\frac{2m_e}{\hbar^2 c^2}}\sqrt{e}\;;$$

$$q = \sqrt{\frac{2m_e}{\hbar^2 c^2}}\sqrt{V_O - e}\;;$$

$$eigenv = \mathrm{Cosh}[q\,b]\,\mathrm{Sin}[2\,k\,a] + \frac{k^2 - q^2}{2\,q\,k}\mathrm{Sinh}[q\,b]\,\mathrm{Cos}[2\,k\,a] + \frac{k^2 + q^2}{2\,q\,k}\mathrm{Sinh}[q\,b];$$

```
a = 30;
b = 7.5;
m = 18;
Teig = Table[FindRoot[eigenv == 0, {e , 0.0195 l}], {l, 1, m}]

{{e} 0.0348863}, {e→0.0348863}, {e→0.0360748}, {e→0.0348863}, 8e→0.138404},
 {e→0.138404}, {e→0.138404}, {e→0.138404}, {e→0.144012}, 8e→0.144012},
 {e→0.144012}, {e→0.138404}, {e→0.144012}, {e→0.322909}, 8e→0.306246},
 {e→0.306246}, {e→0.306246}, {e→0.306246}, {e→0.306246}, 8e→0.322909},
 {e→0.322909}, {e→0.322909}, {e→0.322909}, {e→0.322909}, 8e→0.322909},
 {e→0.322909}, {e→0.57112}, {e→0.528077}, {e→0.528077}, 8e→0.528077}}
```

Thus

$$\begin{aligned} E_{11} &= 0.034886, & E_{12} &= 0.036074,\\ E_{21} &= 0.138404, & E_{22} &= 0.144012,\\ E_{31} &= 0.306246, & E_{32} &= 0.322909,\\ E_{41} &= 0.528077, & E_{42} &= 0.571120. \end{aligned} \tag{5.98}$$

Exercise 16 Plot the eigenfunction $\varphi_{21}(x)$ of the double quantum well bounded by infinite walls when $a = 30\,\text{Å}$, $b = 7.5\,\text{Å}$ and $V_o = 0.6\,\text{eV}$.

Solution We have seen in the solved problem 5.5 that, in order to obtain the eigenfunctions $\varphi_{\mu\nu}(x)$, we need, first, to obtain the wave function $\varphi(x, E)$ for all values of x where the physical system is defined. Once we have this function, we evaluate for $E = E_{\mu\nu}$. Below we copy the program where, after introducing the parameters, V_o, m_e and $\hbar$ in eV, and the light velocity in Å$/s$, we define the wave vectors k and q, in units of Å^{-1}, and the wave functions $\varphi_{\text{I}}(x, E)$, $\varphi_{\text{II}}(x, E)$ and $\varphi_{\text{III}}(x, E)$. To plot the eigenfunction we need to evaluate $\varphi_i(x, E)$ for $E = E_{21}$. In the solved problem 5.2, we obtained, among others, the eigenvalue $E_{21} = 0.13840405001245132\,\text{eV}$. In the second part of the program, copied below, we evaluate the wave functions for

$$\mathbf{k} = \sqrt{\frac{2\,\mathbf{m_e}}{\hbar^2\,\mathbf{c}^2}}\,\sqrt{\mathbf{e}}\;;$$

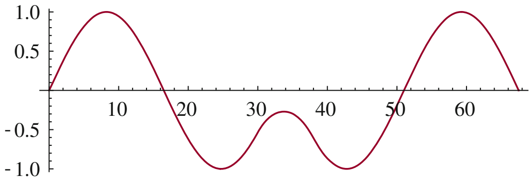
$$\mathbf{q} = \sqrt{\frac{2\,\mathbf{m_e}}{\hbar^2\,\mathbf{c}^2}}\,\sqrt{\mathbf{V_o - e}}\;;$$

$$\varphi_{\mathbf{I}} = \mathbf{Sin[k\,x]};$$

$$\varphi_{\mathbf{II}} = \mathbf{Sin[k\,a]\,Cosh[q\,(x-a)] + \frac{k}{q}\,Cos[k\,a]\,Sinh[q\,(x-a)]};$$

$$\varphi_{\mathbf{III}} = \mathbf{Sin[k\,(x-a-b)]\left(Cosh[q\,b]\,Cos[k\,a] + \frac{q}{k}\,Sinh[q\,b]\,Sin[k\,a]\right) + Cos[k\,(x-a-b)]\left(\frac{k}{q}\,Sinh[q\,b]\,Cos[k\,a] + Cosh[q\,b]\,Sin[k\,a]\right)};$$

```
a = 30;
b = 7.5;
e = 0.1384040500124513;
Eigenfuncion21 = Plot[{Which[x ≤ a, φI, x < (a + b), φII, x ≥ a + b, φIII]},
  {x, 0.000001, 2 a + b}, PlotStyle → {Red, Thickness[0.005]},
  TicksStyle → Directive[Black, 13], PlotPoints → 4000,
  AspectRatio → 1/ 3]
```



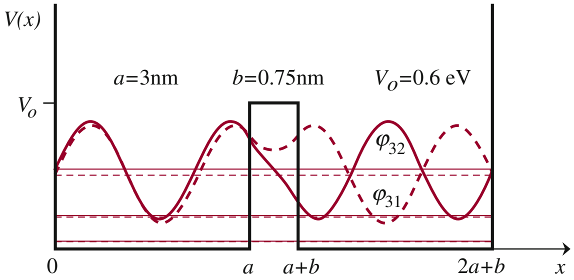


Fig. 5.13 The eigenfunctions $\varphi_{31}(x)$ and $\varphi_{32}(x)$ from the double well with potential parameters $a = 3\,\text{nm}$, $b = 0.75\,\text{nm}$ and $V_o = 0.6\,\text{eV}$. One of the eigenfunctions is symmetric while the other is antisymmetric with respect to the center of the double well

$E = E_{2,1}$ and plot it for x between $x = 0.0$ and $x = 2a + b$. The function is continuous, symmetric (with respect to the center of the double well) and vanishes at the infinite walls located at $x = 0$ and at $x = 2a + b = 67.5\,\text{Å}$.

The student can verify that plotting the function for a slightly different value of the energy, the evident symmetry of the eigenfunction disappears.

Exercise 17 Plot the eigenfunctions $\varphi_{\mu 1}(x)$ and $\varphi_{\mu 2}(x)$, for the same value of μ, and verify that taking the center of the double quantum well as the symmetry point, the eigenfunctions $\varphi_{\mu 1}(x)$ and $\varphi_{\mu 2}(x)$ possess different parity symmetries.

Solution In Fig. 5.13 we plot the eigenfunctions $\varphi_{31}(x)$ and $\varphi_{32}(x)$, corresponding to $E_{31} = 0.3062462125992447$ and $E_{32} = 0.32290870476454275$. To plot these functions we use basically the same program as in the previous exercise. It is clear from these graphs that while $\varphi_{31}(x)$ is even, $\varphi_{32}(x)$ is odd.

Exercise 18 Show that near the resonant energies E^*_μ the transmission coefficient T_D of the double barrier potential can be approximated by the Breit–Wigner formula

$$T_D \simeq \frac{\Gamma^2}{\Gamma^2 + \left(E - E^*_\mu\right)^2}, \tag{5.99}$$

with

$$\Gamma \propto \frac{\hbar}{\tau_\mu}, \tag{5.100}$$

and τ_μ the double-barrier tunneling time.

Solution The transmission coefficient for the double barrier potential was given as

$$T_D = \frac{1}{1 + 2\left|\beta_b\right|^2 \left(\left|\alpha_b\right|^2 + \Re e\ \alpha_b^2 e^{i2ka}\right)}, \tag{5.101}$$

with the resonance condition

$$\left|\alpha_b\right|^2 + \Re e\ \alpha_b^2 e^{i2ka} = 0. \tag{5.102}$$

If we write the complex function α_b as $|\alpha_b|e^{i\theta_\alpha}$ and define the phase

$$\theta = 2(\theta_\alpha + ka), \tag{5.103}$$

the transmission coefficient becomes

$$T_D(E) = \frac{1}{1 + 2\,|\beta_b|^2\,|\alpha_b|^2\,(1+\cos\theta)}, \tag{5.104}$$

and the resonant energies E_μ^* are such that

$$\theta|_{E=E_\mu^*} = \theta_\mu = (2\mu - 1)\pi, \qquad \text{with} \qquad \mu = 1, 2, \dots \tag{5.105}$$

Since $|\alpha_b|$ and $|\beta_b|$ vary slowly with E, we will approximate only the trigonometric function $\cos\theta(E)$ near the resonant points, where $\cos\theta_\mu = -1$, by

$$\begin{aligned} \cos\theta &\simeq \cos\theta_\mu - \sin\theta_\mu \frac{\partial\theta}{\partial E}\bigg|_{E_\mu^*} \left(E - E_\mu^*\right) - \frac{1}{2}\cos\theta_\mu \frac{\partial\theta}{\partial E}\bigg|_{E_\mu^*}^2 \left(E - E_\mu^*\right)^2, \\ &\simeq -1 + \frac{1}{2}\frac{\partial\theta}{\partial E}\bigg|_{E_\mu^*}^2 \left(E - E_\mu^*\right)^2. \end{aligned} \tag{5.106}$$

Replacing this in (5.104) and defining the energy

$$\Gamma = \frac{1}{|\beta_b|\,|\alpha_b|\,\dfrac{\partial\theta}{\partial E}\bigg|_{E_\mu^*}}, \tag{5.107}$$

that has relation with the resonance width, we obtain the approximate formula

$$T_D(E) \simeq \frac{\Gamma^2}{\Gamma^2 + \left(E - E_\mu^*\right)^2}, \tag{5.108}$$

known in nuclear reactions theory as the Breit–Wigner formula for symmetric potentials.

Notice that ka is the phase acquired between the barriers, and since $|t_b|e^{i\theta_t} = 1/\alpha_b^*$, the phase θ_t is equal to θ_α. Thus, the phase θ in (5.103) can be seen as the double barrier phase shift. We will see, in Chap. 7, that a possible definition of the tunneling time across a quantum or electromagnetic system, with transmission-amplitude phase θ_t, is the phase time defined as

$$\tau_{ph} = \hbar\frac{\partial\theta_t}{\partial E}. \tag{5.109}$$

Using this definition, and identifying the double barrier resonant tunneling time τ_μ with $\hbar \left.\frac{\partial \theta}{\partial E}\right|_{E_\mu^*}$, we have

$$\Gamma = \frac{T_b}{\sqrt{1 - T_b}} \frac{\hbar}{\tau_\mu}. \tag{5.110}$$

5.6 Problems

1. Find the eigenvalue equation for the double quantum well bounded by infinite walls when the energy is greater than V_o. Analyze the main differences with the energy levels below V_o.
2. Write the explicit functions of the matrix elements α_D and β_D for the double barrier potential when the energy is below V_o and when the energy is above V_o.
3. Write the explicit functions of the transfer matrix elements α_D and β_D of the double well, bounded by finite walls, when $E < V_o$ and when $E > V_o$.
4. In the finite double well potential, with $a = 30\,\text{Å}$, $b = 7.5\,\text{Å}$ and $V_o = 0.6\,\text{eV}$, obtain:
   a. the eigenvalues $E_{2,1}$, $E_{2,2}$, $E_{3,1}$ and $E_{3,2}$;
   b. the eigenfunctions $\phi_{2,1}$, $\phi_{2,2}$, $\phi_{3,1}$ and $\phi_{3,2}$;
   c. plot the eigenfunctions $\phi_{2,1}$, $\phi_{2,2}$, $\phi_{3,1}$ and $\phi_{3,2}$.
5. The Lorentzian function with amplitude $1/(\pi\gamma)$ is defined as

   $$F(x; x_o, \gamma) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x - x_o)^2}. \tag{5.111}$$

   Show that the Breit–Wigner formula (5.108) is a Lorentzian function with amplitude 1. Using this formula and the potential parameters $a = 2\,\text{nm}$, $b = 10\,\text{nm}$ and $V_o = 0.23$, determine the function Γ and plot the approximate transmission coefficient of the double barrier potential near the first resonance.
6. Obtain the condition that should be satisfied of a resonant value of the transmission coefficient of a double quantum well, bounded by finite walls.
7. Derive the eigenvalue equation for the double quantum well (Eqs. (5.43)–(5.47)) with finite walls (when the energy is less than V_o).
8. Obtain the transfer matrix of the asymmetric double barrier when the width of the second barrier is twice the width of the first one. Plot the transmission coefficient and compare with the transmission coefficient of the symmetric double barrier.
9. After imposing the physical conditions and assuming that the incidence is from the right, obtain the wave function for each region of the double well.
10. Use the general expression for the transmission coefficient, written in terms of the Chebyshev polynomials for $n = 2$, and show that it agrees with that obtained for the double barrier in (5.27).

11. Show that in the forbidden bands the energy is such that

$$\mathrm{Tr}M \geq 2. \tag{5.112}$$

12. Plot as functions of energy both the function α_R and the transmission coefficient T_n for a Kronig–Penney system with $V_o = 0.23\,\mathrm{eV}$, $a = b = 20\,\mathrm{nm}$, and different values of n.

Chapter 6
The WKB Approximation

Immediately after the birth of the quantum theory, Wentzel, Kramers and Brillouin introduced in 1926 one of the first methods to obtain approximate solutions to the Schrödinger equation. A method that might be used when rigorous analytical solutions are not possible. In most of the actual quantum systems we face this problem. It is known that only for exceptional systems like the harmonic oscillator, the hydrogen atom and the piecewise constant potentials, it is possible to solve the Schrödinger equation analytically. It is therefore important to study an approximate method to solve the Schrödinger equation. In this chapter we present the basic ideas behind the WKB approximation and we will learn how to deal with this approximation in the transfer matrix representation. For simplicity reasons, our discussion will be mostly devoted to one-dimensional systems. Another widely used approximation method, is the so-called perturbation theory. In Chap. 12 we will discuss basic aspects of this theory.

6.1 The Semi-Classical Approximation

This method aims to obtain quantum solutions based on the corresponding classical solutions. More precisely, the purpose is to obtain the quantum action S as a correction to the classical action S_c. In the classical theory it is known that the action S_c is the solution of the Hamilton–Jacobi equation

$$\frac{(\nabla S_c)^2}{2m} + V(\mathbf{r}) = E. \tag{6.1}$$

where $V(\mathbf{r})$ is an arbitrary smooth potential. For example the potential in Fig. 6.1. If we write the solution of the corresponding Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\varphi(\mathbf{r}) + V(\mathbf{r})\varphi(\mathbf{r}) = E\varphi(\mathbf{r}), \tag{6.2}$$

P. Pereyra, *Fundamentals of Quantum Physics*, Undergraduate Lecture Notes in Physics, DOI: 10.1007/978-3-642-29378-8_6, © Springer-Verlag Berlin Heidelberg 2012

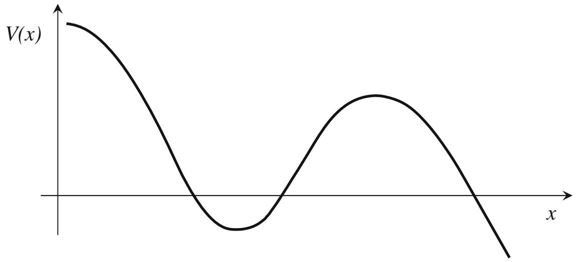


Fig. 6.1 An example of an arbitrary potential function

as

$$\varphi(\mathbf{r}) = e^{iS(\mathbf{r})/\hbar}, \tag{6.3}$$

the quantum action S should satisfy the differential equation

$$\frac{(\nabla S)^2}{2m} - \frac{i\hbar}{2m}\nabla^2 S + V(\mathbf{r}) = E. \tag{6.4}$$

The semiclassical method applies when

$$\hbar|\nabla^2 S| << |(\nabla S)^2|. \tag{6.5}$$

In this case, the quantum action S can be written as the classical action plus some corrections. To obtain these corrections, the prescription is to write the quantum action as the power series of $\hbar$

$$S = S_c + i\hbar S_1 + (i\hbar)^2 S_2 + \ldots, \tag{6.6}$$

with the quantum action S equal to S_c in the limit when $\hbar \to 0$. Using this action in φ and replacing this function into the Schrödinger equation, we get (to second order in $\hbar$) the following equation

$$\begin{aligned} \frac{(\nabla S_c)^2}{2m} + V + i\frac{\hbar}{m}(\nabla S_c \cdot \nabla S_1 - \frac{1}{2}\nabla^2 S_c) \\ -\frac{\hbar^2}{m}\big(2\nabla S_2 \cdot \nabla S_c + (\nabla S_1)^2 - \nabla^2 S_1\big) = E. \end{aligned} \tag{6.7}$$

We have grouped here all terms with the same power of $\hbar$. If we equate to zero the coefficients of $\hbar^0$, $\hbar$ and $\hbar^2$, we end up with the following system of equations

$$\frac{(\nabla S_c)^2}{2m} + V = E, \tag{6.8}$$

$$\nabla S_1 \cdot \nabla S_c - \frac{1}{2}\nabla^2 S_c = 0, \tag{6.9}$$

and

$$\nabla S_2 \cdot \nabla S_c - \frac{1}{2}\nabla^2 S_1 + \frac{1}{2}(\nabla S_1)^2 = 0, \tag{6.10}$$

whose solutions give us the classical action and the quantum corrections S_1 and S_2. To obtain these quantities, we recall the following relation for the classical action and the linear momentum

$$\nabla S_c = \mathbf{p} \quad \text{or equivalently} \quad S_c = \int \mathbf{p} \cdot d\mathbf{r}. \tag{6.11}$$

In one dimensional systems, the classical action will be given by

$$S_c = \pm \int^x p(x)dx \quad \text{con} \quad p(x) = \pm\sqrt{2m\big(E - V(x)\big)}, \tag{6.12}$$

and the first-order correction from the differential equation

$$p(x)\frac{dS_1}{dx} = \frac{1}{2}\frac{dp(x)}{dx}, \tag{6.13}$$

with solution

$$S_1 = \ln\sqrt{p(x)}. \tag{6.14}$$

As a consequence, we have that to first order in $\hbar$, the action will be given by

$$S = \pm \int^x p(x)dx + i\hbar \ln\sqrt{p(x)}. \tag{6.15}$$

Thus, the solution of the one-dimensional Schrödinger equation is

$$\varphi(x) = a\, e^{\left(\frac{i}{\hbar}\int^x p\,dx - \ln\sqrt{p}\right)} + b\, e^{\left(-\frac{i}{\hbar}\int^x p\,dx - \ln\sqrt{p}\right)}. \tag{6.16}$$

This function will be written as

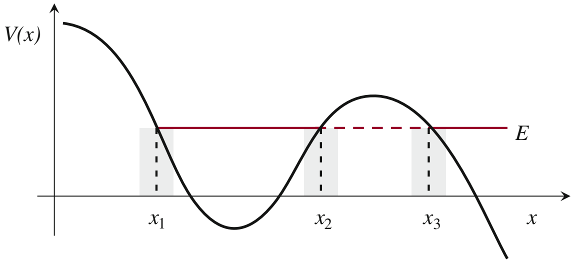
$$\varphi(x) = a\,\frac{1}{\sqrt{p}}e^{\frac{i}{\hbar}\int^x p(x)\,dx} + b\,\frac{1}{\sqrt{p}}e^{-\frac{i}{\hbar}\int^x p(x)\,dx}. \tag{6.17}$$

In the following sections we will consider this type of wave functions.[1] In regions where $E > V(x)$, the wave function can be written in the form

---

[1] Notice that the factor $1/\sqrt{p}$ diverges at the classical turning points. A more detailed analysis can be found, for example, in *Quantum Mechanics* by L. Schiff (McGraw-Hill, NY 1949). See also Sukhatme U and Pagnamenta A *Am. J. Phys.* **59** 944 (1991).

Fig. 6.2 Return points x_i in an arbitrary potential function



$$\varphi(x) = a\,\frac{1}{\sqrt{k}}e^{i\int^x k(x)\,dx} + b\,\frac{1}{\sqrt{k}}e^{-i\int^x k(x)\,dx}, \tag{6.18}$$

while for $E < V(x)$, the momentum $\hbar k(x)$ should be replaced by $i\hbar q(x)$ and the solution will be written in the form

$$\varphi(x) = a\,\frac{1}{\sqrt{q}}e^{\int^x q\,dx} + b\,\frac{1}{\sqrt{q}}e^{-\int^x q\,dx}. \tag{6.19}$$

Although there is some similarity with the functions we had for piecewise constant potentials, one should not forget that k and q are functions of x. It is clear that for constant potentials, the integral will become kx or qx, and we will have the solutions that we know already.

6.2 The Scope of the WKB Approximation

In the motivation of the WKB approximation we have seen that the Schrödinger equation, written in terms of the action S, becomes the classical Hamilton–Jacobi equation when

$$\hbar|\nabla^2 S| << (\nabla S)^2 \qquad \text{or} \qquad \hbar|\nabla\cdot\mathbf{p}| << p^2. \tag{6.20}$$

We want to explore now the meaning of this condition. To simplify the analysis, let us restrict ourselves to one-dimensional systems. In this case, the last condition becomes

$$\hbar\frac{dp}{dx} << p^2(x), \tag{6.21}$$

and can be expressed in different ways. For example, as

$$-\frac{\hbar\,m}{p(x)}\,\frac{dV}{dx} << p^2(x) \qquad \text{or} \qquad \hbar m\frac{F(x)}{p(x)} << p^2(x). \tag{6.22}$$

To fulfill these conditions we require a potential $V(x)$ which is a smooth function of x. Alternatively, if we approximate $V(x)$ by the first order term of its Taylor expansion, around a return point x_i (with x_i any of the return points x_1, x_2 or x_3 in Fig. 6.2), we have

$$\begin{aligned} V(x) &= V(x_i) + \left.\frac{dV}{dx}\right|_{x=x_i} (x - x_i) \\ &= E - F(x_i)(x - x_i). \end{aligned} \tag{6.23}$$

If we now write the momentum $p(x) = \sqrt{2m(E - V(x))}$ as

$$p \simeq \sqrt{2mF(x_i)|x - x_i|}\ , \tag{6.24}$$

we can express (6.21) in the form

$$|x - x_i| >> \frac{1}{2}\left(\frac{\hbar^2}{mF(x_i)}\right)^{1/3}, \quad \text{with} \quad x_i = x_1, x_2, \ldots \tag{6.25}$$

These relations and those in (6.22), define the scope and the physical conditions that the physical system should meet to apply the WKB approximation. Notice that, while (6.22) implies a smooth variation of the potential function, (6.25) tells us that the semiclassical method provides a good description at points x far enough from the return points x_i. This condition will not prevent us from searching the continuity conditions in the neighborhood of or at the return points. The fulfillment of the continuity conditions will allow us to extend the use of the transfer matrix method to the WKB approximation.

6.2.1 *Continuity Conditions and the Connection Formulas for the WKB Approximation*

A detailed analysis of the mathematical continuity conditions leading to the so-called connection formulas at the return points, goes beyond the scope of this book. We shall just use them. The connection formulas that are known for points like the point a of Fig. 6.3 are

$$\frac{1}{\sqrt{q}} e^{-\int_x^a q\,dx} \longleftarrow\!\!\!\!\twoheadrightarrow \frac{2}{\sqrt{k}} \cos\Bigl(\int_a^x k\,dx - \frac{\pi}{4}\Bigr), \tag{6.26}$$

$$-\frac{1}{\sqrt{q}} e^{\int_x^a q\,dx} \twoheadleftarrow\!\!\!\!\longrightarrow \frac{1}{\sqrt{k}} \sin\Bigl(\int_a^x k\,dx - \frac{\pi}{4}\Bigr); \tag{6.27}$$

while for a point like b in Fig. 6.4, the connection formulas are

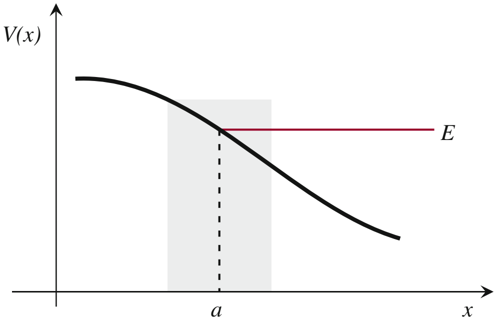


Fig. 6.3 Return point a where the connection formulas (6.26) and (6.27) should be used

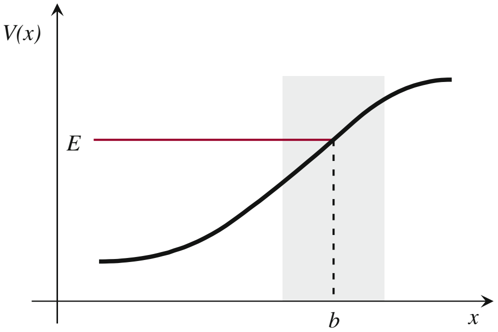


Fig. 6.4 Return point b with connection formulas (6.28) and (6.29)

$$\frac{2}{\sqrt{k}}\cos\Bigl(\int\limits_x^b k\,dx-\frac{\pi}{4}\Bigr)\twoheadleftarrow\!\!\!\!\longrightarrow\frac{1}{\sqrt{q}}e^{-\int_b^x q\,dx}, \tag{6.28}$$

$$\frac{2}{\sqrt{k}}\sin\Bigl(\int_x^b k\,dx-\frac{\pi}{4}\Bigr)\longleftarrow\!\!\!\!\twoheadrightarrow-\frac{1}{\sqrt{q}}e^{\int_b^x q\,dx}. \tag{6.29}$$

These relations should be used in the manner indicated by the double arrows. If we relax the double arrow and we write the connection formulas in a matrix representation, we have

$$\frac{1}{\sqrt{q}}\begin{pmatrix}e^{-\int_x^a q\,dx}\\ e^{\int_x^a q\,dx}\end{pmatrix}\longleftrightarrow\frac{1}{\sqrt{2}}\begin{pmatrix}1-i & 1+i\\ \dfrac{1+i}{2} & \dfrac{1-i}{2}\end{pmatrix}\frac{1}{\sqrt{k}}\begin{pmatrix}e^{i\int_a^x k\,dx}\\ e^{-i\int_a^x k\,dx}\end{pmatrix}, \tag{6.30}$$

for state vectors in the neighborhood of a point like a, and

$$\frac{1}{\sqrt{k}}\begin{pmatrix}e^{-i\int_x^b k\,dx}\\ e^{i\int_x^b k\,dx}\end{pmatrix}\longleftrightarrow\frac{1}{\sqrt{2}}\begin{pmatrix}1+i & \dfrac{1-i}{2}\\ 1-i & \dfrac{1+i}{2}\end{pmatrix}\frac{1}{\sqrt{q}}\begin{pmatrix}e^{\int_b^x q\,dx}\\ e^{-\int_b^x q\,dx}\end{pmatrix}, \tag{6.31}$$

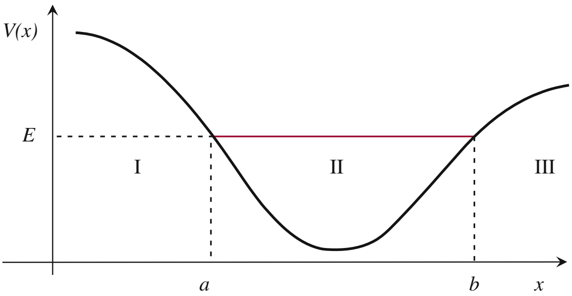


Fig. 6.5 An arbitrary smooth potential well. In a potential like this, the energy eigenvalues are obtained from the Sommerfeld-Wilson-Ishiwara formula

for state vectors in the neighborhood of a point like b. The connection matrices

$$C_a = \frac{1}{\sqrt{2}}\begin{pmatrix} 1-i & 1+i \\ \frac{1+i}{2} & \frac{1-i}{2} \end{pmatrix} \equiv \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix}, \tag{6.32}$$

and

$$C_b = \frac{1}{\sqrt{2}}\begin{pmatrix} 1+i & \dfrac{1-i}{2} \\ 1-i & \dfrac{1+i}{2} \end{pmatrix} = \begin{pmatrix} c_{11}^* & c_{21}^* \\ c_{12}^* & c_{22}^* \end{pmatrix} = C_a^\dagger, \tag{6.33}$$

differ slightly from the transition matrices that will be obtained in the next sections. It is easy to verify that C_b is not the inverse of C_a. C_b is the adjoint of C_a.[2]

6.2.2 *Energy Quantization in the Potential Well*

We know that if we have a potential function as in Fig. 6.5, the energy must quantize. Let us now see how these energies come out in the semiclassical approximation. In this system, as in the case of the square well, it is important to guarantee the finiteness of the Schrödinger solutions as the variable $x \to \pm\infty$. The exponential functions in regions I and III are, certainly, functions with divergency problems in these limits. Therefore, we are left with

$$\varphi_{\rm I}(x) = A_1 \frac{1}{\sqrt{q(x)}} e^{-\int_x^a q(x)dx}, \quad \text{when} \quad x < a, \tag{6.34}$$

and

$$\varphi_{\rm III}(x) = A_3 \frac{1}{\sqrt{q(x)}} e^{-\int_b^x q(x)dx}, \quad \text{when} \quad x > b. \tag{6.35}$$

---

[2] The adjoint of a matrix is defined as its transpose conjugate, i.e. $C^\dagger = (C^T)^*$.

In this case the connection formulas that we are going to use are: the connection formula (6.26), which helps us define the function at the right hand side of the point a, and the connection formula (6.28), that helps us define the function at the left of b. Using these formulas, we have for the wave function in region II either the stationary solution in the form

$$\varphi_{\text{II}}(x) = \frac{A}{\sqrt{k(x)}} \cos\left(\int_a^x k(x)dx - \frac{\pi}{4}\right), \tag{6.36}$$

or the stationary function in the form

$$\widetilde{\varphi}_{\text{II}}(x) = \frac{B}{\sqrt{k(x)}} \cos\left(\int_x^b k(x)dx - \frac{\pi}{4}\right). \tag{6.37}$$

Both expressions should coincide. Hence, they should transform one into the other. If we notice that the cosine, appearing in (6.36), can be rewritten as

$$\cos\left(\int_a^x kdx - \frac{\pi}{4}\right) = \cos\left(\int_a^b kdx - \int_x^b kdx + \frac{\pi}{4} - \frac{\pi}{2}\right). \tag{6.38}$$

and we define the phase

$$\eta = \int_a^b k(x)dx - \frac{\pi}{2}, \tag{6.39}$$

we have

$$\cos\left(\int_a^x kdx - \frac{\pi}{4}\right) = \cos\left(\eta - \int_x^b kdx + \frac{\pi}{4}\right). \tag{6.40}$$

Using the identity

$$\cos\left(n\pi - \theta\right) = (-1)^n \cos\left(\theta\right), \quad n = 0, 1, 2, \ldots, \tag{6.41}$$

it is easy to verify that $\varphi_{\text{II}}(x)$ equal to $\widetilde{\varphi}_{\text{II}}(x)$, only when the phase η fulfills the relation

$$\eta = \int_a^b k(x)dx - \frac{\pi}{2} = n\pi, \quad \text{and} \quad A = (-1)^n B. \tag{6.42}$$

Taking into account that the momentum $p(x)$ changes sign when the integration is from b to a, we have

$$\int_a^b k(x)dx = \int_b^a k(x)dx = \pi\Big(n + \frac{1}{2}\Big). \tag{6.43}$$

Therefore

$$\oint p(x)dx = h\Big(n + \frac{1}{2}\Big), \tag{6.44}$$

which is equivalent to the quantization rule of Sommerfeld, Wilson and Ishiwara.

6.3 Transfer Matrices in the WKB Approximation

We are interested in applying the semiclassical method for arbitrary potential wells and barriers, such that the above mentioned conditions are met. We know already that the transport properties and the energy eigenvalues of a quantum system can easily be evaluated when the transfer matrices are known. Our aim is to obtain the transfer matrices of arbitrary potential wells and barriers, within the WKB method so that they can then be used in specific applications.

6.3.1 The Transfer Matrix of a Quantum Well

In this problem and in the potential barrier, some integrals will appear repeatedly. To simplify the notation let us define the following functions

$$\xi_1(x) = \int_{x_1}^{x} k(x)dx, \qquad \xi_2(x) = \int_{x}^{x_2} k(x)dx, \tag{6.45}$$

$$\vartheta_1(x) = \int_{x}^{x_1} q(x)dx, \qquad \vartheta_2(x) = \int_{x_2}^{x} q(x)dx. \tag{6.46}$$

with

$$k(x) = \sqrt{2m\big(E - V(x)\big)}/\hbar \qquad \text{and} \qquad q(x) = \sqrt{2m\big(V(x) - E\big)}/\hbar. \tag{6.47}$$

Using these functions, the WKB solutions in regions I, II and III, for the potential well shown in Fig. 6.6, can be written, respectively, as

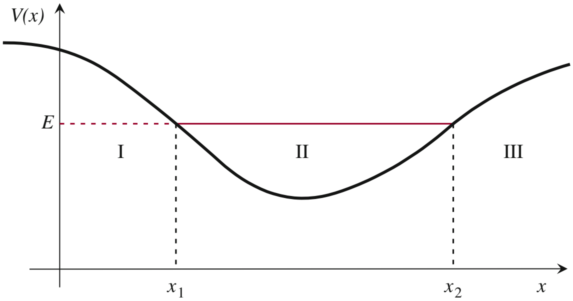


Fig. 6.6 An arbitrary potential well

$$\varphi_{\rm I}(x) = \frac{A_1}{\sqrt{q(x)}}e^{-\vartheta_1(x)} + \frac{B_1}{\sqrt{q(x)}}e^{\vartheta_1(x)}, \qquad \text{for} \quad x < x_1, \tag{6.48}$$

$$\varphi_{\rm II}(x) = \frac{A_2}{\sqrt{k(x)}}e^{i\xi_1(x)} + \frac{B_2}{\sqrt{k(x)}}e^{-i\xi_1(x)}, \qquad \text{for} \quad x_1 < x < x_2,$$

$$\widetilde{\varphi}_{\rm II}(x) = \frac{A_2'}{\sqrt{k(x)}}e^{-i\xi_2(x)} + \frac{B_2'}{\sqrt{k(x)}}e^{i\xi_2(x)}, \tag{6.49}$$

$$\varphi_{\rm III}(x) = \frac{A_3}{\sqrt{q(x)}}e^{\vartheta_2(x)} + \frac{B_3}{\sqrt{q(x)}}e^{-\vartheta_2(x)}, \qquad \text{for} \quad x > x_2. \tag{6.50}$$

For region II we have written two representations of the wave function. As in the piecewise constant potential problems, the functions should satisfy the continuity conditions at x_1 and at x_2. But we can not evaluate at the return points. At these points $p(x)$ and $q(x)$ become zero and the wave functions diverge. We will evaluate the functions near x_1 and near x_2, and we will adjust their coefficients so that the functions and their derivatives, on both sides of the return points, become equal. At x_1 we want that

$$\varphi_1(x_1 - \epsilon) \simeq \varphi_2(x_1 + \epsilon) \quad \text{and} \quad \varphi_1'(x_1 - \epsilon) = \varphi_2'(x_1 + \epsilon), \tag{6.51}$$

where $\epsilon = |x - x_i|_{x \to x_i}$, such that the condition (6.25) will be satisfied. The first continuity condition gives rise to

$$\frac{A_1}{\sqrt{q(x_1^-)}} + \frac{B_1}{\sqrt{q(x_1^-)}} \simeq \frac{A_2}{\sqrt{k(x_1^+)}} + \frac{B_2}{\sqrt{k(x_1^+)}}, \tag{6.52}$$

and the second to

$$\frac{A_1}{q^{3/2}}\left(\frac{dq}{dx} + 2q^2\right)\bigg|_{x=x_1} + \frac{B_1}{q^{3/2}}\left(\frac{dq}{dx} - 2q^2\right)\bigg|_{x=x_1} \simeq \frac{A_2}{k^{3/2}}\left(\frac{dk}{dx} + i2k^2\right)\bigg|_{x=x_1} + \frac{B_2}{k^{3/2}}\left(\frac{dk}{dx} - i2k^2\right)\bigg|_{x=x_1}. \tag{6.53}$$

If we assume that the potential function meets the conditions

$$\left|\frac{dq}{dx}\right| << q^2 \quad \text{and} \quad \left|\frac{dk}{dx}\right| << k^2, \tag{6.54}$$

for the WKB approximation, the continuity condition for the wave function derivatives takes the simpler form

$$q(x_1^-)\left(\frac{A_1}{\sqrt{q(x_1{}^-)}} - \frac{B_1}{\sqrt{q(x_1{}^-)}}\right) \simeq ik(x_1{}^+)\left(\frac{A_2}{\sqrt{k(x_1{}^+)}} - \frac{B_2}{\sqrt{k(x_1{}^+)}}\right). \tag{6.55}$$

To simplify the notation we will use $q_1 = q(x_1^-)$ and $k_1 = k(x_1^+)$. Therefore, Eqs. (6.52) and (6.55) can be expressed in compact form as follows:

$$2i\frac{k_1}{\sqrt{k_1}}\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \simeq \begin{pmatrix} ik_1 + q_1 & ik_1 - q_1 \\ ik_1 - q_1 & ik_1 + q_1 \end{pmatrix}\begin{pmatrix} A_1 \\ B_1 \end{pmatrix}\frac{1}{\sqrt{q_1}}, \tag{6.56}$$

or even better as

$$\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \simeq \frac{1}{2\sqrt{k_1 q_1}}\begin{pmatrix} k_1 - iq_1 & k_1 + iq_1 \\ k_1 + iq_1 & k_1 - iq_1 \end{pmatrix}\begin{pmatrix} A_1 \\ B_1 \end{pmatrix}, \tag{6.57}$$

with the transition matrix

$$M_{ta} = \frac{1}{2\sqrt{k_1 q_1}}\begin{pmatrix} k_1 - iq_1 & k_1 + iq_1 \\ k_1 + iq_1 & k_1 - iq_1 \end{pmatrix}. \tag{6.58}$$

It is worth noticing the similarity of this transition matrix with the corresponding transition matrix for the rectangular potential well. In the same way we obtain the relation

$$\begin{pmatrix} A_3 \\ B_3 \end{pmatrix} \simeq \frac{1}{2\sqrt{q_2 k_2}}\begin{pmatrix} q_2 + ik_2 & q_2 - ik_2 \\ q_2 - ik_2 & q_2 + ik_2 \end{pmatrix}\begin{pmatrix} A_2' \\ B_2' \end{pmatrix}. \tag{6.59}$$

with the transition matrix

$$M_{tb} = \frac{1}{2\sqrt{q_2 k_2}}\begin{pmatrix} q_2 + ik_2 & q_2 - ik_2 \\ q_2 - ik_2 & q_2 + ik_2 \end{pmatrix}. \tag{6.60}$$

With the transition matrices just obtained, we have an important part of our main goal. Notice that M_{tb} is the inverse of M_{ta}. To complete the calculation we still need the relation between the coefficients A_2', B_2' and the coefficients A_2, B_2.

As mentioned before, the functions $\varphi_{\rm II}(x)$ and $\widetilde{\varphi}_{\rm II}(x)$ in (6.49) are equivalent. Therefore, the state vectors

$$\phi_2(x) = \begin{pmatrix} e^{i\int_{x_1}^{x} k(x)dx} & 0 \\ 0 & e^{-i\int_{x_1}^{x} k(x)dx} \end{pmatrix} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix}, \tag{6.61}$$

and

$$\widetilde{\phi}_2(x) = \begin{pmatrix} e^{-i\int_{x}^{x_2} k(x)dx} & 0 \\ 0 & e^{i\int_{x}^{x_2} k(x)dx} \end{pmatrix} \begin{pmatrix} A_2^{'} \\ B_2^{'} \end{pmatrix}, \tag{6.62}$$

should also be equal. This allows writing

$$\begin{pmatrix} A_2^{'} \\ B_2^{'} \end{pmatrix} = \begin{pmatrix} e^{i\int_{x_1}^{x_2} k\,dx} & 0 \\ 0 & e^{-i\int_{x_1}^{x_2} k\,dx} \end{pmatrix} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix}. \tag{6.63}$$

The propagating matrix that comes out in this equation connects state vectors at two points of the well. In this case from x_1^+ to x_2^-. If we use the notation

$$\xi(x_2, x_1) = \int_{x_1^+}^{x_2^-} k(x)dx, \tag{6.64}$$

and combine (6.58), (6.63) and (6.60) we obtain

$$\begin{pmatrix} A_3 \\ B_3 \end{pmatrix} \simeq \frac{1}{4(q_1k_1q_2k_2)^{1/2}} \begin{pmatrix} q_2+ik_2 & q_2-ik_2 \\ q_2-ik_2 & q_2+ik_2 \end{pmatrix} \begin{pmatrix} e^{i\xi} & 0 \\ 0 & e^{-i\xi} \end{pmatrix} \begin{pmatrix} k_1-iq_1 & k_1+iq_1 \\ k_1+iq_1 & k_1-iq_1 \end{pmatrix} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix}. \tag{6.65}$$

At this point we can use the potential defined in (6.23)

$$V(x) = E - F(x_i)(x - x_i), \tag{6.66}$$

with

$$F(x_1^-) \cong F(x_1^+) = F_1 \geq 0, \tag{6.67}$$

and also with

$$F(x_2^-) \cong F(x_2^+) = F_2 \leq 0, \tag{6.68}$$

to write[3]

---

[3] One must be especially careful with these equations when at both sides of the return point one has an abrupt change in the potential function or different effective masses.

$$q_1k_2 \simeq \frac{2m}{\hbar^2}\sqrt{F_1|F_2||x_1^{\pm} - x_1||x_2^{\pm} - x_2|} \simeq q_2k_1, \tag{6.69}$$

$$q_1q_2 \simeq \frac{2m}{\hbar^2}\sqrt{F_1|F_2||x_1^{\pm} - x_1||x_2^{\mp} - x_2|} \simeq k_1k_2. \tag{6.70}$$

It is easy to verify that, after multiplying and using these equivalences, the relation (6.65) is greatly simplified and takes the form

$$\begin{pmatrix} A_3 \\ B_3 \end{pmatrix} = \begin{pmatrix} \cos\xi & \sin\xi \\ -\sin\xi & \cos\xi \end{pmatrix} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix}, \tag{6.71}$$

with

$$M_w(x_2^+, x_1^-) = \begin{pmatrix} \cos\xi & \sin\xi \\ -\sin\xi & \cos\xi \end{pmatrix}, \tag{6.72}$$

the transfer matrix of a quantum well with arbitrary potential profile in the WKB approximation. This matrix has an extremely compact representation. Using connection formulas one obtains a matrix with numerical factors 2 and 1/2 in the diagonal terms that may be questionable (see solved problem 24). The transfer matrix (6.72) is compatible with the Sommerfeld–Wilson–Ishiwara quantization (see the solved problem 19). We must not forget, however, that the argument ξ of the trigonometric functions is an integral.

6.3.2 *Transfer Matrix and Tunneling Through a Barrier*

In this part we will use the definitions

$$\xi_1(x) = \int_x^{x_1} k(x)dx, \qquad \xi_2(x) = \int_{x_2}^{x} k(x)dx, \tag{6.73}$$

$$\vartheta_1(x) = \int_{x_1}^{x} q(x)dx, \qquad \vartheta_2(x) = \int_{x}^{x_2} q(x)dx. \tag{6.74}$$

with

$$k(x) = \sqrt{2m(E - V(x))}/\hbar \quad \text{and} \quad q(x) = \sqrt{2m(V(x) - E)}/\hbar. \tag{6.75}$$

The calculation of the transfer matrix and the reflection and transmission coefficients of an arbitrary potential barrier, as the one shown in Fig. 6.7, is an important goal. We will obtain these reflection and transmission coefficients using the WKB method. In

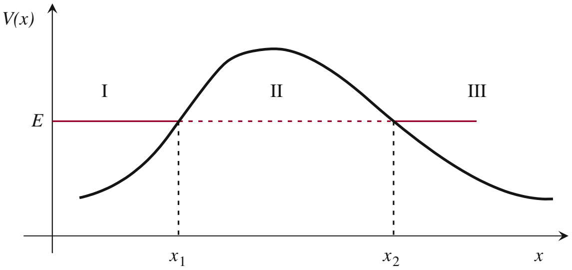


Fig. 6.7 An arbitrary potential barrier

regions I, II and III we have the following solutions:

$$\varphi_{\rm I}(x) = \frac{A_1}{\sqrt{k(x)}} e^{-i\xi_1(x)+i\frac{\pi}{4}} + \frac{B_1}{\sqrt{k(x)}} e^{i\xi_1(x)-i\frac{\pi}{4}}; \tag{6.76}$$

$$\varphi_{\rm II}(x) = \frac{A_2}{\sqrt{q(x)}} e^{-\vartheta_1(x)} + \frac{B_2}{\sqrt{q(x)}} e^{\vartheta_1(x)}; \tag{6.77}$$

$$\widetilde{\varphi}_{\rm II}(x) = \frac{A_2^{'}}{\sqrt{q(x)}} e^{-\vartheta_2(x)} + \frac{B_2^{'}}{\sqrt{q(x)}} e^{\vartheta_2(x)} \tag{6.78}$$

and

$$\varphi_{\rm III}(x) = \frac{A_3}{\sqrt{k(x)}} e^{i\xi_2(x)-i\frac{\pi}{4}} + \frac{B_3}{\sqrt{k(x)}} e^{-i\xi_2(x)+i\frac{\pi}{4}}. \tag{6.79}$$

From the continuity conditions on x_1 and x_2 we have, in matrix representation, the relationships

$$\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} = \frac{1}{2\sqrt{q_1 k_1}} \begin{pmatrix} q_1 + ik_1 & q_1 - ik_1 \\ q_1 - ik_1 & q_1 + ik_1 \end{pmatrix} \begin{pmatrix} A_1 e^{i\frac{\pi}{4}} \\ B_1 e^{-i\frac{\pi}{4}} \end{pmatrix}, \tag{6.80}$$

$$\begin{pmatrix} A_3 e^{-i\frac{\pi}{4}} \\ B_3 e^{i\frac{\pi}{4}} \end{pmatrix} = \frac{1}{2\sqrt{q_2 k_2}} \begin{pmatrix} k_2 - iq_2 & k_2 + iq_2 \\ k_2 + iq_2 & k_2 - iq_2 \end{pmatrix} \begin{pmatrix} A_2^{'} \\ B_2^{'} \end{pmatrix}. \tag{6.81}$$

Since $\varphi_{\rm II}(x)$ and $\widetilde{\varphi}_{\rm II}(x)$ are equal, we have

$$\begin{pmatrix} e^{-\int_x^{x_2} q(x)dx} & 0 \\ 0 & e^{\int_x^{x_2} q(x)dx} \end{pmatrix} \begin{pmatrix} A_2^{'} \\ B_2^{'} \end{pmatrix} = \begin{pmatrix} e^{\int_{x_1}^{x} q(x)dx} & 0 \\ 0 & e^{-\int_{x_1}^{x} q(x)dx} \end{pmatrix} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix}, \tag{6.82}$$

that leads to

$$\begin{pmatrix} A_2^{'} \\ B_2^{'} \end{pmatrix} = \begin{pmatrix} e^{\int_{x_1}^{x_2} q(x)dx} & 0 \\ 0 & e^{-\int_{x_1}^{x_2} q(x)dx} \end{pmatrix} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix}. \tag{6.83}$$

In this equation we have the propagation matrix through the barrier, from x_1 to x_2. To simplify the notation we define the function

$$\vartheta(x_1^-, x_2^+) = \int_{x_1^-}^{x_2^+} q(x)dx. \tag{6.84}$$

With all results obtained so far, we are able to connect the state vector on the right with the state vector on the left of the barrier. Substituting we have

$$\begin{pmatrix} A_3 e^{-i\frac{\pi}{4}} \\ B_3 e^{i\frac{\pi}{4}} \end{pmatrix} = \frac{1}{4\sqrt{q_1 k_1 q_2 k_2}} \begin{pmatrix} k_2 - iq_2 & k_2 + iq_2 \\ k_2 + iq_2 & k_2 - iq_2 \end{pmatrix} \begin{pmatrix} e^{\vartheta} & 0 \\ 0 & e^{-\vartheta} \end{pmatrix} \\ \times \begin{pmatrix} q_1 + ik_1 & q_1 - ik_1 \\ q_1 - ik_1 & q_1 + ik_1 \end{pmatrix} \begin{pmatrix} A_1 e^{i\frac{\pi}{4}} \\ B_1 e^{-i\frac{\pi}{4}} \end{pmatrix}. \tag{6.85}$$

After multiplying and using the equalities in (6.69), we have:

$$\begin{pmatrix} A_3 e^{-i\frac{\pi}{4}} \\ B_3 e^{i\frac{\pi}{4}} \end{pmatrix} = \begin{pmatrix} \cosh\vartheta & -i\sinh\vartheta \\ i\sinh\vartheta & \cosh\vartheta \end{pmatrix} \begin{pmatrix} A_1 e^{i\frac{\pi}{4}} \\ B_1 e^{-i\frac{\pi}{4}} \end{pmatrix}. \tag{6.86}$$

Also in this case the transfer matrix has a simple and compact form. The transfer matrix for an arbitrary barrier in the WKB approximation is

$$M_b(x_2^+, x_1^-) = \begin{pmatrix} \cosh\vartheta & -i\sinh\vartheta \\ i\sinh\vartheta & \cosh\vartheta \end{pmatrix}. \tag{6.87}$$

This matrix differs also in numerical factors with the matrix that is obtained using the connection formulas (see the solved problem 24). The transfer matrix in (6.87) has the form

$$M_b = \begin{pmatrix} \alpha_b & \beta_b \\ \beta_b^* & \alpha_b^* \end{pmatrix}, \tag{6.88}$$

characteristic of time reversal invariant systems. The calculation of transmission and reflection coefficients is one of the important applications of this result. To calculate these coefficients we have to use again the relations between the transfer matrix elements and the scattering amplitudes. Given

$$\alpha_b = \cosh\vartheta = \frac{1}{2}\left(e^{\vartheta} + e^{-\vartheta}\right), \tag{6.89}$$

and

$$\beta_b = -i\sinh\vartheta = -\frac{i}{2}\left(e^{\xi} - e^{-\vartheta}\right), \tag{6.90}$$

we obtain the transmission coefficient

$$T_b = \frac{1}{|\alpha_b|^2} = \frac{4}{\left(e^{\vartheta} + e^{-\vartheta}\right)^2}. \tag{6.91}$$

Using the connection formulas, the transmission coefficient is[4]

$$T_b = \frac{4}{\left(2e^{\vartheta} + \frac{1}{2}e^{-\vartheta}\right)^2}. \tag{6.92}$$

These expressions coincide in the limit ϑ large, but they have important differences for small ϑ. While the transmission coefficient obtained here takes values in the range [0, 1], the transmission coefficient obtained with the connection formulas take values in the range [0, 0.64]! If $\vartheta >> 1$, in both cases, we have

$$T_b \simeq \frac{1}{e^{2\vartheta}} \propto e^{-2\int_{x_1}^{x_2} q(x)dx}. \tag{6.93}$$

This is a very well known and widely used result. For example, in the explanation of the nuclear α decay. The transfer matrices obtained in these previous sections are fundamental in the solution of various problems in the WKB approximation, which in many cases provide excellent results.

6.4 Solved Problems

Exercise 19 Using the relation (6.71) between the coefficients of $\varphi_{\rm III}(x_2^+)$ and $\varphi_{\rm I}(x_1^-)$ for a quantum well in the WKB approximation, obtain the quantization rule that has been found in Sect. 6.2.2.

Solution The relation that was obtained in Sect. 6.3.1 between the coefficients of the wave functions $\varphi_{\rm III}(x_2^+)$ and $\varphi_{\rm I}(x_1^-)$ in a quantum well, with an arbitrary profile, is

$$\begin{pmatrix} A_3 \\ B_3 \end{pmatrix} = \begin{pmatrix} \cos\xi & \sin\xi \\ -\sin\xi & \cos\xi \end{pmatrix} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix}. \tag{6.94}$$

The finiteness condition of functions in regions I and III, requires that the coefficients B_1 and A_3 vanish. The above equation is then written in the form

$$\begin{pmatrix} 0 \\ B_3 \end{pmatrix} = \begin{pmatrix} \cos\xi & \sin\xi \\ -\sin\xi & \cos\xi \end{pmatrix} \begin{pmatrix} A_1 \\ 0 \end{pmatrix} = \begin{pmatrix} A_1\cos\xi \\ -\,A_1\sin\xi \end{pmatrix}, \tag{6.95}$$

to be satisfied when

---

[4] See for example: E Merzbacher, *Quantum Mechanics*, (John Wiley & sons, NY 1961).

$$\cos\xi = 0 \quad \text{and} \quad B_3 = -iA_1 \sin\xi. \tag{6.96}$$

The first of these equations is satisfied if $\xi = \left(n + \frac{1}{2}\right)\pi$, i.e. if

$$\frac{1}{\hbar}\int\limits_{x_1}^{x_2} p(x)dx = \left(n + \frac{1}{2}\right)\pi, \tag{6.97}$$

which is compatible with (6.44) and the quantization rule of Sommerfel–Wilson–Ishiwara.

Exercise 20 Plot the Morse potential

$$V(x) = A(e^{-2ax} - 2e^{-ax}), \tag{6.98}$$

and using the quantization rule, determine the energy levels for a particle that moves in this potential.

Solution We leave the student to plot the Morse potential for different values of the parameters a and A. To obtain the energy eigenvalues in this attractive potential, we start defining the variable $y = e^{-ax}$. In terms of this variable, the quantization rule becomes

$$-\frac{\sqrt{2m}}{a\hbar}\int\limits_{y_1}^{y_2} \frac{\sqrt{E_n + 2Ay - Ay^2}}{y}\, dy = \left(n + \frac{1}{2}\right)\pi. \tag{6.99}$$

with y_1 and y_2 the roots of

$$E_n + 2Ay - Ay^2 = 0. \tag{6.100}$$

Since

$$\begin{aligned}\int \frac{\sqrt{E + 2Ay - Ay^2}}{y}\, dy &= \sqrt{E + 2Ay - Ay^2} + \sqrt{A}\tan^{-1}\frac{\sqrt{A}(y-1)}{\sqrt{E + 2Ay - Ay^2}} \\ &\quad + \sqrt{E}\ln\frac{y}{E + Ay + \sqrt{E}\sqrt{E + 2Ay - Ay^2}},\end{aligned} \tag{6.101}$$

we have

$$\int\limits_{y_1}^{y_2} \frac{\sqrt{E_n + 2Ay - Ay^2}}{y}\, dy = \sqrt{A}\pi + \sqrt{E_n}\ln\frac{y_2(E_n + Ay_1)}{y_1(E_n + Ay_2)}. \tag{6.102}$$

Hence the quantization rule takes in this case the form

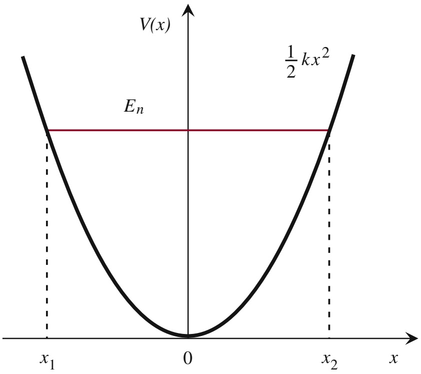


Fig. 6.8 Harmonic potential. In the return points $E_n = kx_i^2/2$

$$
\begin{aligned}
-\frac{\sqrt{2m}}{a\hbar}\int_{y_1}^{y_2}\frac{\sqrt{E_n+2Ay-Ay^2}}{y}\,dy &= -\frac{\sqrt{2m}}{a\hbar}\left(\pi\sqrt{A}+i\pi\sqrt{E_n}\right),\\
&= \left(n+\frac{1}{2}\right)\pi. \qquad (6.103)
\end{aligned}
$$

Thus, for the eigenvalues of the Morse potential we obtain

$$E_n = -A\left(1-\frac{\hbar a}{\sqrt{2mA}}\left(n+\frac{1}{2}\right)\right)^2. \qquad (6.104)$$

It is a good exercise to verify that the same energy eigenvalues are obtained when the corresponding Schrödinger equation is solved.

Exercise 21 Using the quantization rule, show that the energy levels of a particle, in the harmonic oscillator $V(x) = kx^2/2$, are

$$E_n = \hbar\omega\left(n+\frac{1}{2}\right). \qquad (6.105)$$

Where ω is the classical frequency $\sqrt{k/m}$. [Note: the constant k here should not be confused with the wave number k(x)].

Solution The quantization rule (6.97) applied to the harmonic oscillator is

$$\frac{1}{\hbar}\int_{x_1}^{x_2}\sqrt{2m\left(E_n-\frac{1}{2}kx^2\right)}\,dx = \left(n+\frac{1}{2}\right)\pi. \qquad (6.106)$$

If we factor $\sqrt{mk}$, from the square root, we have (Fig. 6.8)

$$\frac{\sqrt{mk}}{\hbar}\int\limits_{x_1}^{x_2}\sqrt{\frac{2E_n}{k}-x^2}\,dx=(n+\frac{1}{2})\pi. \tag{6.107}$$

Let us recall that, at the return point, $2E_n/k=x_1^2=x_2^2$. It is easy to see that the integral of this equation is of the type $\int_{-a}^{a}\sqrt{a^2-x^2}dx$. This integral is known and has the following result

$$\int\limits_{-a}^{a}\sqrt{a^2-x^2}dx=\left[\frac{x\sqrt{a^2-x^2}}{2}+\frac{a^2}{2}\sin^{-1}\frac{x}{a}\right]_{-a}^{a}=2\frac{a^2}{2}\frac{\pi}{2}. \tag{6.108}$$

If we substitute a^2 by $2E_n/k$ and $\sqrt{k/m}$ by ω, we have

$$\frac{\pi\,E_n}{\hbar\,\omega}=\left(n+\frac{1}{2}\right)\pi, \tag{6.109}$$

that obviously can be rewritten in the well-known formula for the eigenvalues of the harmonic oscillator

$$E_n=\hbar\,\omega\left(n+\frac{1}{2}\right). \tag{6.110}$$

This formula is obtained also after solving, rigorously, the Schrödinger equation of the harmonic oscillator.

Exercise 22 Consider the potential well in Fig. 6.6. Using the connection formulas (6.30) and (6.31), show that the transition matrices, defined by the relations

$$\begin{pmatrix}A_2\\B_2\end{pmatrix}=M_{t1}\begin{pmatrix}A_1\\B_1\end{pmatrix}\qquad\text{and}\qquad\begin{pmatrix}A_3\\B_3\end{pmatrix}=M_{t2}\begin{pmatrix}A_2'\\B_2'\end{pmatrix}, \tag{6.111}$$

are $M_{t1}=C_a^T$ and $M_{t2}=C_b^T$, with C_a and C_b, defined in Sect. 6.2.1.

Solution We will assume that the WKB solutions in regions I, II and III are as we write in (6.48), (6.49) and (6.50), i.e.:

$$\varphi_{\text{I}}(x)=\frac{A_1}{\sqrt{q(x)}}e^{-\vartheta_1(x)}+\frac{B_1}{\sqrt{q(x)}}e^{\vartheta_1(x)},\quad\text{for}\quad x<x_1, \tag{6.112}$$

$$\varphi_{\text{II}}(x)=\frac{A_2}{\sqrt{k(x)}}e^{i\xi_1(x)}+\frac{B_2}{\sqrt{k(x)}}e^{-i\xi_1(x)}, \tag{6.113}$$

$$\text{for}\quad x_1<x<x_2,$$

$$\widetilde{\varphi}_{\text{II}}(x)=\frac{A_2'}{\sqrt{k(x)}}e^{-i\xi_2(x)}+\frac{B_2'}{\sqrt{k(x)}}e^{i\xi_2(x)}, \tag{6.114}$$

$$\varphi_{\text{III}}(x)=\frac{A_3}{\sqrt{q(x)}}e^{\vartheta_2(x)}+\frac{B_3}{\sqrt{q(x)}}e^{-\vartheta_2(x)},\quad\text{for}\quad x>x_2. \tag{6.115}$$

And the connection formulas at the points x_1 and x_2, are respectively:

$$\frac{1}{\sqrt{q}}\begin{pmatrix} e^{-\vartheta_1(x)} \\ e^{\vartheta_1(x)} \end{pmatrix} \longleftrightarrow \frac{1}{\sqrt{2}}\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \frac{1}{\sqrt{k}}\begin{pmatrix} e^{i\xi_1(x)} \\ e^{-i\xi_1(x)} \end{pmatrix}, \tag{6.116}$$

and

$$\frac{1}{\sqrt{k}}\begin{pmatrix} e^{-i\xi_2(x)} \\ e^{i\xi_2(x)} \end{pmatrix} \longleftrightarrow \frac{1}{\sqrt{2}}\begin{pmatrix} c_{11}^* & c_{21}^* \\ c_{12}^* & c_{22}^* \end{pmatrix} \frac{1}{\sqrt{q}}\begin{pmatrix} e^{\vartheta_2(x)} \\ e^{-\vartheta_2(x)} \end{pmatrix}, \tag{6.117}$$

with the elements c_{ij} defined in (6.32). In (6.116) we have two scalar equations. If we multiply the first by A_1, the second by B_1 an then add the results, we have

$$A_1\frac{e^{-\vartheta_1(x)}}{\sqrt{q(x)}} + B_1\frac{e^{\vartheta_1(x)}}{\sqrt{q(x)}} \longleftrightarrow (c_{11}A_1 + c_{21}B_1)\frac{e^{i\xi_1(x)}}{\sqrt{k(x)}} + (c_{12}A_1 + c_{22}B_1)\frac{e^{-i\xi_1(x)}}{\sqrt{k(x)}}. \tag{6.118}$$

At the left side of this relation we have $\varphi_{\rm I}(x)$ and on the right side we must have the function $\varphi_{\rm II}(x)$ to which it connects at x_1. This implies the following relations

$$\begin{aligned} A_2 &= c_{11}A_1 + c_{21}B_1, \\ B_2 &= c_{12}A_1 + c_{22}B_1, \end{aligned} \tag{6.119}$$

that can be written as

$$\begin{pmatrix} A_2 \\ B_2 \end{pmatrix} \longleftrightarrow \begin{pmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \end{pmatrix}\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = M_{t1}\begin{pmatrix} A_1 \\ B_1 \end{pmatrix}. \tag{6.120}$$

Comparing the transition matrix M_{t1}, defined in this equation, with the matrix C_a of Eq. (6.32), we end up with

$$M_{t1} = C_a^T = \frac{1}{\sqrt{2}}\begin{pmatrix} 1-i & \dfrac{1+i}{2} \\ 1+i & \dfrac{1-i}{2} \end{pmatrix}. \tag{6.121}$$

Proceeding in the same way with the connection formula at x_2 we obtain

$$\begin{pmatrix} A_3 \\ B_3 \end{pmatrix} \longleftrightarrow \begin{pmatrix} c_{11}^* & c_{12}^* \\ c_{21}^* & c_{22}^* \end{pmatrix}\begin{pmatrix} A_2^{'} \\ B_2^{'} \end{pmatrix} = M_{t2}\begin{pmatrix} A_2^{'} \\ B_2^{'} \end{pmatrix}. \tag{6.122}$$

Thus

$$M_{t2} = C_b^T = \frac{1}{\sqrt{2}}\begin{pmatrix} 1+i & 1-i \\ \dfrac{1-i}{2} & \dfrac{1+i}{2} \end{pmatrix}. \tag{6.123}$$

Notice that M_{t2}, it is not the inverse of M_{t1} but its adjoint. This is possibly one of the reasons why the expression obtained for transfer matrices, using the WKB connection formulas, has a structure that is not completely satisfactory.

Exercise 23 Obtain the transfer matrices $M_w(x_2^+, x_1^-)$ and $M_b(x_2^+, x_1^-)$, for the quantum well and the potential barrier in the WKB approximation, using the connection formulas at points x_1 and x_2.

Solution To obtain the required transfer matrices, we have almost all we need. We just got (see exercise 22), the matrices M_{t1} and M_{t2} to connect the coefficients on either side of the return points. In Sect. 6.3.1, we had

$$\begin{pmatrix} A_2^{'} \\ B_2^{'} \end{pmatrix} = \begin{pmatrix} e^{i\int_{x_1}^{x_2} k\,dx} & 0 \\ 0 & e^{-i\int_{x_1}^{x_2} k\,dx} \end{pmatrix} \begin{pmatrix} A_2 \\ B_2 \end{pmatrix}. \tag{6.124}$$

Hence, the connection between the state vector $(A_3, B_3)^T$ and the state vector $(A_1, B_1)^T$ can be written as follows

$$\begin{pmatrix} A_3 \\ B_3 \end{pmatrix} = M_{t2} \begin{pmatrix} e^{i\xi} & 0 \\ 0 & e^{-i\xi} \end{pmatrix} M_{t1} \begin{pmatrix} A_1 \\ B_1 \end{pmatrix}, \tag{6.125}$$

where the notation $\xi = \int_{x_1}^{x_2} k\,dx$ has been used. Therefore, the transfer matrix of the quantum well becomes

$$\begin{aligned} M_w &= M_{t2} \begin{pmatrix} e^{i\xi} & 0 \\ 0 & e^{-i\xi} \end{pmatrix} M_{t1}, \\ &= \frac{1}{2} \begin{pmatrix} 1+i & 1-i \\ \dfrac{1-i}{2} & \dfrac{1+i}{2} \end{pmatrix} \begin{pmatrix} e^{i\xi} & 0 \\ 0 & e^{-i\xi} \end{pmatrix} \begin{pmatrix} 1-i & \dfrac{1+i}{2} \\ 1+i & \dfrac{1-i}{2} \end{pmatrix}, \\ M_w &= \begin{pmatrix} 2\cos\xi & -\sin\xi \\ \sin\xi & \frac{1}{2}\cos\xi \end{pmatrix}. \end{aligned} \tag{6.126}$$

In a similar way, the transfer matrix for the potential barrier reads

$$\begin{aligned} M_b &= M_{t1} \begin{pmatrix} e^{\vartheta} & 0 \\ 0 & e^{-\vartheta} \end{pmatrix} M_{t2}, \\ &= \frac{1}{2} \begin{pmatrix} 1-i & \dfrac{1+i}{2} \\ 1+i & \dfrac{1-i}{2} \end{pmatrix} \begin{pmatrix} e^{\vartheta} & 0 \\ 0 & e^{-\vartheta} \end{pmatrix} \begin{pmatrix} 1+i & 1-i \\ \dfrac{1-i}{2} & \dfrac{1+i}{2} \end{pmatrix}, \\ M_b &= \begin{pmatrix} e^{\vartheta} + \frac{1}{4}e^{-\vartheta} & -i\left(e^{\vartheta} - \frac{1}{4}e^{-\vartheta}\right) \\ i\left(e^{\vartheta} - \frac{1}{4}e^{-\vartheta}\right) & e^{\vartheta} + \frac{1}{4}e^{-\vartheta} \end{pmatrix}. \end{aligned} \tag{6.127}$$

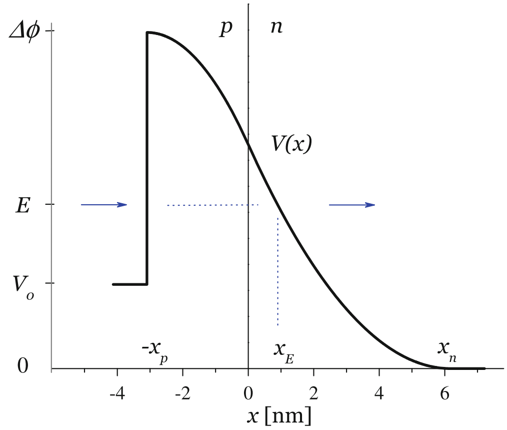


Fig. 6.9 A potential barrier characteristic of tunneling diodes at the interface of strongly doped semiconductors

Exercise 24 Suppose a system has a potential profile like the one shown in Fig. 6.9, and electrons are approaching from the left side. To define this potential we can think of it as the potential in a real *np* diode, or Esaki diode.[5] Thus, let us define, the potential profile as follows.

$$V(x) = \begin{cases} V_0 = -eV_b, & \text{for} \quad x < -x_p, \\ V_1(x) = \Delta\phi - a(x + x_p)^2, & \text{for} \quad -x_p \le x \le 0, \\ V_2(x) = b(x - x_n)^2, & \text{for} \quad 0 \le x \le x_n, \\ 0, & \text{for} \quad x > x_n, \end{cases} \tag{6.128}$$

with V_b a bias potential, and $\Delta\phi$, a, b, x_n and x_n device dependent parameters.

Given this potential, and using the WKB approximation, calculate the transmission and reflection coefficients as functions of the incoming electron energies E, for a bias $V_b = -0.5\,\mathrm{V}$ and fixed potential parameters.

Though in real systems there is a gap of forbidden energies, we will solve here for all values of E.[6] The parameters $\Delta\phi$, a, b, x_p and x_n, are defined as follows:

---

[5] In tunneling diodes, or Esaki diodes, electrons from the valence band, on the p side, move towards the n side under small reverse bias.

[6] In tunnel or Esaki diodes, the impurity concentrations are so large that the Fermi levels μ_p and μ_n move inside the valence and the conduction bands, respectively. On the p side, between the Fermi level μ_p (that corresponds to V_0 in the figure) and the conduction band (at $\Delta\phi$), there is a gap and no electrons can be found there.

$$\begin{aligned}
\Delta\phi &= \mu_n - \mu_p - eV_b = E_g - k_B T \ln \frac{N_c N_v}{n^+ p^+} - eV_b,\\
a &= \frac{e^2 p^+}{2\epsilon_p \epsilon_0}, \quad b = \frac{e^2 n^+}{2\epsilon_n \epsilon_0},\\
x_n &= \left(\frac{\Delta\phi}{b}\frac{p^+}{n^+ + p^+}\right) \quad x_p = \left(\frac{\Delta\phi}{a}\frac{n^+}{n^+ + p^+}\right).
\end{aligned} \tag{6.129}$$

with E_g the energy gap; e the electron charge, k_B the Boltzmann constant; T the temperature; n^+ and p^+ the donors and acceptors (high) concentration on the left and right sides, respectively; ϵ_n and ϵ_p the relative electric permittivities; and N_c and N_v, the intrinsic electron and hole concentrations, in the conduction and valence bands.[7]

Solution All we need is the transfer matrix to connect state vectors at two points outside the potential barrier, say, $-x_p$ and x_n. For a given value of E, we need first the transition matrix M_l to account for the continuity requirements of the wave functions at $-x_p$. We then need to propagate the state vector inside the potential barrier, from $-x_p$ to the turning point x_E, where $V(x) = E$. For this purpose we will determine a transfer matrix M_1. At this point we use the connection matrix C_a, and finally a transfer matrix M_2 to reach the point x_n. These matrices are the following:

$$M_l = \frac{1}{2q_p}\begin{pmatrix} q_p + ik & q_p - ik \\ q_p - ik & q_p + ik \end{pmatrix}, \tag{6.130}$$

with

$$k = \sqrt{\frac{2m_e^*}{\hbar^2}\left(E - V_0\right)} \quad \text{and} \quad q_p = \sqrt{\frac{2m_e^*}{\hbar^2}\Big(V(-x_p) - E\Big)}. \tag{6.131}$$

The transfer matrix that propagates the state vector inside the barrier is

$$M_l = \begin{pmatrix} e^{\vartheta} & 0 \\ 0 & e^{-\vartheta} \end{pmatrix}. \tag{6.132}$$

Assuming that the return point is x_E on the n side, we have here

$$\vartheta = \int_{-x_p}^{0} \sqrt{\frac{2m_e^*}{\hbar^2}\Big(V_1(x) - E\Big)dx} + \int_{0}^{x_E} \sqrt{\frac{2m_e^*}{\hbar^2}\Big(V_2(x) - E\Big)dx}, \tag{6.133}$$

which, after integrating, becomes

---

[7] These are well known quantities in the semiconductor physics, and they read $N_c = 2\left(\frac{2\pi m_e^* k_B T}{\hbar^2}\right)^{3/2}$ and $N_v = 2\left(\frac{2\pi m_h^* k_B T}{\hbar^2}\right)^{3/2}$, with m_e^* and m_h^* the effective electron and hole masses.

$$\vartheta = \frac{c}{2b}\sin^{-1}(bx_p) + \frac{c\,x_p}{2}\sqrt{1-b^2x_p^2} + \frac{d}{2}\left(x_n\sqrt{x_n^2-z^2} + z^2\ln\frac{z}{x_n+\sqrt{x_n^2-z^2}}\right), \tag{6.134}$$

with

$$c = \sqrt{\frac{2m_e^*}{\hbar^2}(\Delta\phi - E)}, \quad d = \sqrt{\frac{m_e^*}{\hbar^2}\frac{e^2n^+}{\epsilon_n\epsilon_0}} \tag{6.135}$$

and

$$z = \sqrt{\frac{2\epsilon_n\epsilon_0}{e^2n^+}E}. \tag{6.136}$$

As mentioned before, to connect at x_E we use the matrix

$$C_a = \frac{1}{\sqrt{2}}\begin{pmatrix} 1-i & 1+i \\ \frac{1+i}{2} & \frac{1-i}{2} \end{pmatrix}, \tag{6.137}$$

defined in (6.32). Finally, the transfer matrix that propagates the physics from x_E to x_n is

$$M_2 = \begin{pmatrix} e^{i\xi} & 0 \\ 0 & e^{-i\xi} \end{pmatrix}, \tag{6.138}$$

with

$$\xi = \int_{x_E}^{x_n} \sqrt{\frac{2m_e^*}{\hbar^2}\Big(E - V_2(x)\Big)dx} = d\frac{\pi z^2}{4}. \tag{6.139}$$

With these matrices we can obtain the transfer matrix to connect state vectors across the potential barrier. This matrix is given by the product

$$M_b = M_2C_aM_1M_l = \begin{pmatrix} \alpha_b & \beta_b \\ \beta_b^* & \alpha_b^* \end{pmatrix}. \tag{6.140}$$

with

$$\alpha_b = \frac{e^{i\xi}}{\sqrt{2}q_p}\Big(\big(k+q_p\big)\cosh\vartheta + i\big(k-q_p\big)\sinh\vartheta\Big) \tag{6.141}$$

and

$$\beta_b = -\frac{e^{i\xi}}{\sqrt{2}q_p}\Big(\big(k-q_p\big)\cosh\vartheta + i\big(k+q_p\big)\sinh\vartheta\Big) \tag{6.142}$$

It is now easy to evaluate the transmission and reflection coefficients:

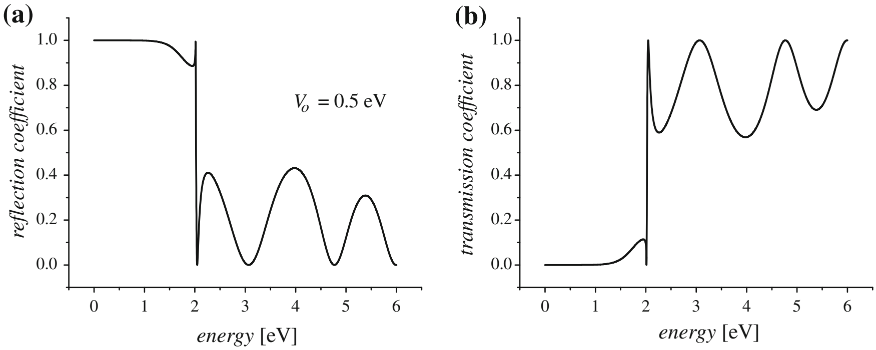


Fig. 6.10 The reflection and transmission coefficients by and through the potential barrier of Fig. 6.9. At the transition to the state of propagating modes a nice Fano–Majorana resonance is observed

$$T_b = \frac{1}{|\alpha_b|^2} \quad \text{and} \quad R = 1 - T. \tag{6.143}$$

In Fig. 6.10a, b we plot these quantities using parameters that correspond to GaAs (i.e. $E_g = 1.51\,\text{eV}$, $m_e^* = 0.069\,m_e$, $m_h^* = 0.1\,m_e$), concentrations $n^+ = 6\times10^{-25}\,\text{m}^{-3}$ and $p^+ = 1\times10^{-26}\,\text{m}^{-3}$ and bias $V_b = -0.5\,\text{V}$. For specific systems, one has to choose carefully the parameter values. As the electron energy approaches the barrier height, where the potential varies slowly and the transition to the propagating mode starts, we see another characteristic property of the scattering phenomena: The interference between the exponential and the propagating modes in the continuum, that manifest through the asymmetric Fano–Majorana resonance at $E \simeq 2\,\text{eV}$.[8]

6.5 Problems

1. Show that the transition matrix at a point like x_2, in Fig. 6.6, is certainly given by the product of matrices in Eq. (6.65), and show also that it reduces to the matrix given in (6.72).
2. Show that the transition matrix at the point b of Fig. 6.4 is given by the matrix M_{tb} in (6.60). Show also that $M_{tb} = (M_{ta})^{-1}$
3. Obtain the transfer matrix of the barrier in Fig. 6.7, assuming that the transition matrix M_{t2} at point x_2 is the inverse of M_{t1} at point x_1.
4. Show that the eigenvalue equation of a double symmetric quantum well, like the one shown in Fig. 6.11, is

$$\tan\xi = \pm e^{\vartheta}. \tag{6.144}$$

---

[8] After the discovery and explanation of these kind of resonances by Ettore Majorana and Ugo Fano, respectively.

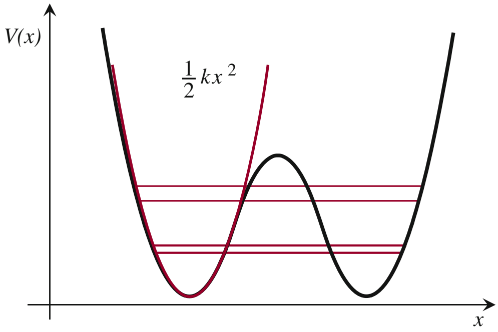


Fig. 6.11 A double symmetric quantum well (*black*) and the parabolic approximation. The energy levels of this system split up

5. Plot the function $V(x) = -U_o/\cosh^2 ax$ for different values of the potential parameters U_o and a. Using the quantization rule, obtain the eigenvalues for a particle with this potential energy.
6. Plot the function $V(x) = U_o/\cosh^2 ax$ for different values of the potential parameters U_o and a, and using the semiclassical approximation, obtain the transmission and reflection coefficients for particles coming from the left side of this potential barrier.
7. Obtain analytically the first two eigenvalues of a quantum well if, for sufficiently low energies, the potential can be approximated by a parabolic function, as shown in Fig. 6.11.
8. Show from the graphical solution of the eigenvalue equation of problem 4 that:

    a. the energies split up;
    b. the lowest energies match with those obtained analytically in the approximation $V(x) \simeq kx^2/2$.

Chapter 7
Operators and Dynamical Variables

In the last chapters we studied some solvable one-dimensional systems described by the stationary Schrödinger equation. In the solutions of these systems we had a first approach to the quantum phenomena. We found that for systems with confining potentials the physical variables generally quantize. For some of this kind of systems, we could determine the momentum and energy eigenvalues as well as their corresponding eigenfunctions. We found also the tunneling effect, the energy levels splitting and we were able to evaluate the particle current density and the reflection and transmission coefficients.

A major aim of the quantum theory is the prediction of values for the relevant dynamical variables. The possibility or not, of determining such values, depends to a great extent on the physical conditions and the physical variables themselves. We will see here that an important property of a physical variable is related with the commutation relation of its associated operator with other operators like the energy operator, the Hamiltonian. When we are dealing with transport of particles through a quantum system, one can fix, through the experimental set up, the energy of the incoming particles. In that case the quantum theory will be able to predict a number of energy-dependent quantities like the transmission coefficients, but the theory will not be able to predict, with the same accuracy, the arrival time of the transmitted particles. We will see in this chapter that the product of the statistical dispersions of canonical conjugate variables, like energy and time or momentum and position, have restricted values in quantum theory.

The theory predicts a set of possible states, but it does not have, generally, the ability to determine which state the system is actually in. We will see that in cases like this, there is no other way than to calculate the expected or average values. This indeterminacy should not be confused with an inaccuracy in the calculation of physical quantities. We have seen in previous chapters that the eigenvalues, eigenfunctions and other physical variables, such as reflection and transmission coefficients can be calculated accurately. The precise evaluation of wave functions allowed us to see, for example, the energy levels splitting in double quantum wells (where energies can be evaluated with the desired accuracy) and the way in which the allowed and forbidden

P. Pereyra, *Fundamentals of Quantum Physics*, Undergraduate Lecture Notes in Physics, DOI: 10.1007/978-3-642-29378-8_7, © Springer-Verlag Berlin Heidelberg 2012

bands appear in metals, semiconductors and other periodic systems. It is important, to notice that it is quite normal that in comparing theoretical predictions with the experimental results, may result in large differences. There are many reasons for such differences to appear. Usually real systems are affected by other variables, such as temperature, impurities, etc., which are not taken into account and are naturally absent in simplified models. This kind of limitations should be explained or corrected using more realistic approaches.

Since the theory can not generally tell us which of the possible states the particle is actually in, it was natural to introduce the concept of the *expectation value* of a physical variable. We will comment on this issue and we will talk about operators, the Hermitian operators, their properties and consequences. We will show one of various derivations of the Heisenberg uncertainty relation and we will make some comments on the time evolution and equations of motion of dynamical variables and their expected values. Closely related to this subject is the widely spread and repeated assertion that concepts like position and trajectory of a particle, as conceived in the classical physics, become, in the quantum theory, meaningless. We will make a brief reference to this matter. In this chapter we will also introduce the Heisenberg picture.

7.1 Wave Packets, Group Velocity and Tunneling Time

In Chap. 3 we addressed the free particle description where we studied basically the normalization problem of the linearly independent solutions $\varphi^+(x) = Ae^{ikx}$ and $\varphi^-(x) = Ae^{-ikx}$. We know that if we include the temporal component, the wave functions $\psi^{\pm}(x,t) = A\varphi^{\pm}(x)e^{-i\omega t}$ describe particles that move to the right or to the left. Since the particle is not confined, there is no quantization of energy or momentum. If the particle's energy is fixed, the wave number $k = \sqrt{2mE/\hbar^2}$, the momentum $p = \hbar k$ and the frequency $\omega = E/\hbar$ are also fixed. The quest now is, whether we can determine the position and the speed of the particle at the same time. Using the previous relations, we can easily show that the speed of the free particle is

$$v = \frac{\hbar k}{m} = \frac{d\omega}{dk}, \tag{7.1}$$

but, where is the particle? If we describe the free particle with the wave function $\psi^{\pm}(x,t) = A\varphi^{\pm}(x)e^{-i\omega t}$, practically, we ignore where the particle is. In the previous chapters we mentioned the probabilistic interpretation of the wave function and we have seen that the probability of finding a particle at the position x, at time t, is given by $|\psi(x,t)|^2$. For a free particle moving to the right or to the left, this probability is

$$|\psi(x,t)|^2 = |A|^2. \tag{7.2}$$

Independent of x and time, therefore the same for all points from $-\infty$ to ∞.

If we describe the particle by the superposition of a right-moving and a left-moving wave functions like

$$\psi(x,t) = a\, e^{i(kx-\omega t)} + b\, e^{-i(kx+\omega t)}, \tag{7.3}$$

the probability of finding the particle at the position x, at time t, will be

$$|\psi(x,t)|^2 = |a|^2 + |b|^2 + 2\Re e\; a\, b^*\, e^{i2kx}, \tag{7.4}$$

which is no longer a constant but an oscillatory function. In the particular case of $a = b$, the wave function becomes

$$\psi(x,t) = 2a \cos kx e^{-i\omega t}. \tag{7.5}$$

The real part of this function oscillates in space with wave number k and in time with angular frequency ω. The probability of finding the particle at point x, at time t, is

$$|\psi(x,t)|^2 = 4|a|^2 \cos^2 kx, \tag{7.6}$$

independent of time but periodic as function of x.

It is well known that with an appropriate superposition of functions, the probability $|\psi(x,t)|^2$ can eventually be localized. If we consider the superposition of two waves[1] describing free particles moving in the same direction like

$$\psi_2(x,t) = a\, e^{i(kx-\omega t)} + a\, e^{i(k'x-\omega' t)}, \tag{7.7}$$

such that

$$\begin{aligned} k &= k_o - \delta k \qquad & k' &= k_o + \delta k \\ \omega &\simeq \omega_o - \delta\omega \qquad & \omega' &\simeq \omega_o + \delta\omega, \end{aligned} \tag{7.8}$$

with $\delta\omega = \hbar k_o \delta k/m$, the two-component wave function in (7.7) can be written as

$$\psi_2(x,t) = 2a \cos(\delta k\, x - \delta\omega t) e^{i(k_o x - \omega_o t)}, \tag{7.9}$$

This is also a product of two wave functions: one with frequency ω_o and wave number k_o and the other a wave function with frequency $\delta\omega_o$ and wave number δk_o. These functions are plotted in Fig. 7.1. When $\delta k_o \ll k_o$, the wave function $\cos(\delta k x - \delta\omega t)$ behaves as a modulation amplitude of the internal oscillations described by $e^{i(k_o x - \omega_o t)}$. These waves have different physical meaning and move with different velocities. The constant-phase condition[2] applied to $(k_o x - \omega_o t)$ gives the phase velocity

---

[1] This simple example is frequently used in optical physics to introduce phase and group velocities.

[2] The constant-phase condition implies

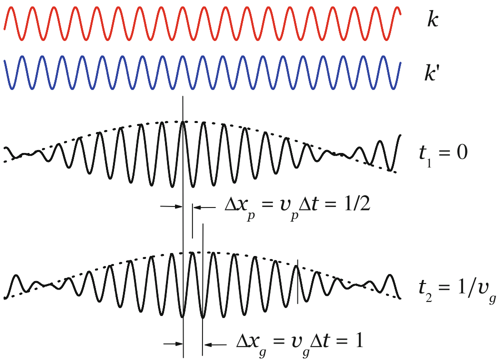


Fig. 7.1 The real parts of two wave functions, with wave numbers $k = k_o - \delta k$ and $k' = k_o + \delta k$, and snapshots of their superposition at $t_1 = 0$ and at $t_2 = 1/v_g$. During the time interval $\Delta t = t_2 - t_1 = 1/v_g$, a constant-phase point of the internal waves moves a distance Δx_p and the maximum of the modulation envelope moves a distance Δx_g

$$v_p = \frac{dx}{dt} = \frac{\omega_o}{k_o}, \tag{7.11}$$

while applied to $(\delta k x - \delta\omega t)$ gives the group velocity[3]

$$v_g = \frac{dx}{dt} = \frac{\delta\omega}{\delta k} \simeq \frac{\hbar k_o}{m}. \tag{7.13}$$

This result reiterates the relation $p = mv = \hbar k$. In Fig. 7.1 we plot the real parts of the wave components together with snapshots of their superposition at $t_1 = 0$ and at $t_2 = 1/v_g = m/\hbar k_o$. During the period of time $\Delta t = t_2 - t_1$, a point of constant pase of the internal oscillations moves with the phase velocity a distance $\Delta x_p = v_p \Delta t = 1/2$ while the modulation envelope moves a distance $\Delta x_g = v_g \Delta t$. This shows that the modulating function moves with the group velocity $v_g = \hbar k_o/m$. The same thing happens in the superposition of more than two components, with wave numbers $k_o \pm k$ symmetric with respect to k_o, to build a wave packet. The overall wave envelope will move with the group velocity. The phase velocity $\omega_o/k_o = \hbar k_o/2m$ of the internal oscillations is approximately the same as the phase velocities ω/k and ω'/k' of the wave components.

We will consider now a superposition of functions like

(Footnote 2 continued)

$$\frac{d(kx - \omega t)}{dt} = 0, \tag{7.10}$$

and can be used to obtain the phase and group velocities.

[3] In the limit of $\delta k \to 0$ we have

$$v_g = \frac{d\omega}{dk}. \tag{7.12}$$

$$\phi(x,t) = \int_{k_p-k_a}^{k_p+k_a} A(k) e^{i(kx-\omega t)} dk. \tag{7.14}$$

In this superposition, pairs of wave functions with wave numbers k in the neighborhood of k_p, and amplitude $A(k)$ each, add to form a wave packet. The shape of the packet relies strongly on $A(k)$. One of the most frequently used amplitude is the Gaussian function centered at k_p, with dispersion (or standard deviation) σ_k. In this case, we have

$$\phi(x,t) = \frac{1}{\sqrt{2\pi}\sigma_k} \int_{-\infty}^{\infty} e^{-(k-k_p)^2/2\sigma_k^2} e^{i(kx-\omega t)} dk. \tag{7.15}$$

Let us see if a superposition like this can be used to describe a free particle moving with wave number k_p and speed $v = \hbar k_p/m$. To visualize this, it is convenient to perform the integral in the variable k. Rewriting the exponent as follows

$$-\frac{(k-k_p)^2}{2\sigma_k^2} + i\left(kx - \frac{\hbar k^2 t}{2m}\right) = -a\,(k-\kappa)^2 + a\kappa^2 - \frac{k_p^2}{2\sigma_k^2}, \tag{7.16}$$

with

$$a = \frac{m + i\hbar\sigma_k^2 t}{2m\sigma_k^2} \quad \text{and} \quad \kappa = \frac{im\sigma_k^2(x - ik_p/\sigma_k^2)}{m + i\hbar\sigma_k^2 t}, \tag{7.17}$$

the integration on the variable k is direct and the packet takes the form

$$\psi(x,t) = \frac{1}{\sigma_x\sigma_k} e^{\dfrac{i(k_p x - w_p t)}{\sigma_k^2\sigma_x^2}} e^{-\dfrac{x^2}{2\sigma_x^2}}. \tag{7.18}$$

This is also a Gaussian function with

$$\sigma_x^2 = \frac{m + i\hbar\sigma_k^2 t}{m\sigma_k^2} = \frac{1 + i\frac{\hbar t}{m}\sigma_k^2}{\sigma_k^2}. \tag{7.19}$$

As expected, the procedure of building up a wave packet, effectively concentrates the probability $|\psi(x,t)|^2$ of finding the particle around the Gaussian centroid.

Does the wave packet move with the speed of the free particle? This question can be answered if we follow the space-time evolution of the wave-packet centroid, i.e. of the wave-packet peak. If the wave-packet peak moves with constant speed, we can determine the peak speed knowing its position at any two times, t_1 and t_2. In Fig. 7.2 we present results for this "experiment" where the wave function was

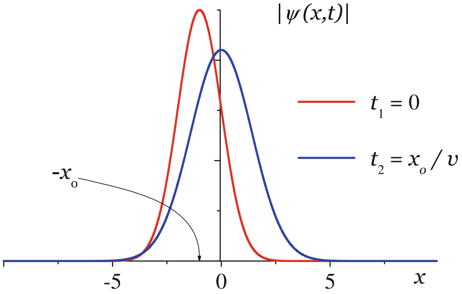


Fig. 7.2 Snapshots of a Gaussian wave packet at $t_1 = 0$ and at $t_2 = x_o/v$, with $v = \hbar k_p/m$. Since the wave-packet peak moves from $x_1 = -x_o$ to $x_2 = 0$, the wave-packet velocity v_g coincides with the free particle speed v

$$\psi(x,t) = \frac{1}{\sigma_x \sigma_k} e^{\frac{i(k_p(x+x_o) - w_p t)}{\sigma_k^2 \sigma_x^2}} e^{-\frac{(x+x_o)^2}{2\sigma_x^2}}. \tag{7.20}$$

At $t = 0$ the peak is at $x = -x_o$. We choose $t_1 = 0$ and $t_2 = x_o/v$, with $v = \hbar k_p/m$ the speed of the free particle. If the wave-packet peak moves with the speed of the free particle (i.e. with $v_g = v$), the peak position at t_2 must be $x_2 = -x_o + vt_2 = 0$. The snapshots show that this is precisely what happens. Hence, the Gaussian wave packet moves with the speed of the free particle. If a Gaussian wave packet describes a free particle, the peak position will coincide on the average with the particle position, and the standard deviation will describe the particle-position dispersion Δx. In Fig. 7.3 we plot the functions $A(k)$ and $|\psi(x,t)|$ for the standard deviations indicated in the figure. It is evident from these plots that $\sigma_x = \Delta x \simeq 1/\Delta k = 1/\sigma_k$. This relation is consistent with (7.19). We will return to this subject some lines below.

A widely discussed problem in quantum physics, is the time that a particle takes to pass through a potential region, such as the potential barrier. In classical physics the procedure to calculate the transit time is well defined. In quantum physics, the tunneling time has been a controversial topic. In recent years there have been accurate measurements of transit time of photons and electromagnetic pulses. It has been also shown that the tunneling or delay time in the quantum theory and in the electromagnetic theory are correctly described by the so-called phase time.

How is the phase time defined? We have seen in different examples that the transmitted wave function $\varphi_t(x_2)$ and the incident wave function $\varphi_i(x_1)$ have a very simple relation. In fact, if the points x_1 and x_2 are just at the left and right of a potential region, say a potential barrier like in Fig. 7.4, there exist a barrier transfer matrix such that

$$\begin{pmatrix} a_t e^{ikx_2} \\ b_3 e^{-ikx_2} \end{pmatrix} = \begin{pmatrix} \alpha_b & \beta_b \\ \beta_b^* & \alpha_b^* \end{pmatrix} \begin{pmatrix} a_i e^{ikx_1} \\ b_r e^{-ikx_1} \end{pmatrix}. \tag{7.21}$$

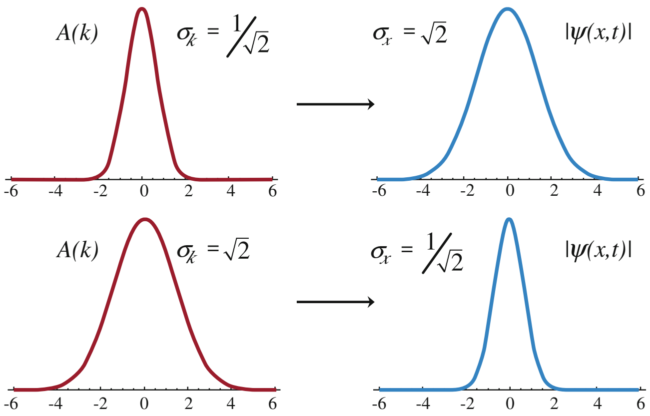


Fig. 7.3 The Gaussian amplitude $A(k)$ in the superposition (7.15), produces a Gaussian function whose variance σ_x^2 is roughly the inverse of the variance σ_k^2 in $A(k)$. This is usually referred to as the uncertainty relation of position and linear momentum $\Delta x \simeq 1/\Delta k$

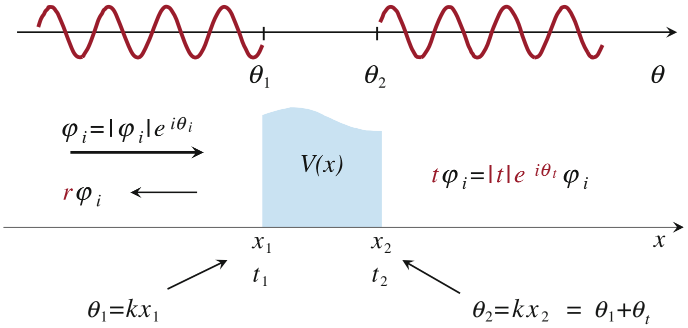


Fig. 7.4 The transit time through a potential region and its relation with the phases of the incoming and transmitted wave functions

In terms of the reflection and transmission amplitudes, r_b and t_b, this relation reads

$$\begin{pmatrix} a_t e^{i\theta_2} \\ b_3 e^{-i\theta_2} \end{pmatrix} = \begin{pmatrix} 1/t_b^* & -r_b^*/t_b^* \\ -r_b/t_b & 1/t_b \end{pmatrix} \begin{pmatrix} a_i e^{i\theta_1} \\ b_r e^{-i\theta_1} \end{pmatrix}. \tag{7.22}$$

If the incident particle comes only from the left, $b_3 = 0$ and we have

$$b_r e^{-ikx_1} = r_b\, a_i e^{ikx_1} \tag{7.23}$$

$$a_t e^{ikx_2} = \frac{1}{t_b^*} a_i e^{ikx_1} - \frac{r_b^*}{t_b^*} b_r e^{-ikx_1}. \tag{7.24}$$

Combining these equations and taking into account that $1 - r_b r_b^* = t_b t_b^*$, we obtain

$$a_t e^{ikx_2} = t_b a_i e^{ikx_1} = |t_b| e^{i\theta_t} a_i e^{ikx_1}, \tag{7.25}$$

which is a known result that we have used to obtain the transmission coefficient and the transmitted wave function. In the last equation we have written the transmission amplitude in terms of its absolute value $|t_b|$ and its phase θ_t. This allows us to observe the following important relation between phases

$$kx_2 = kx_1 + \theta_t, \quad \theta_t = k(x_2 - x_1). \tag{7.26}$$

If we take the frequency derivative of the transmission phase θ_t, we have

$$\frac{\partial \theta_t}{\partial \omega} = \frac{\partial k}{\partial \omega}(x_2 - x_1) = \frac{1}{v_g}(x_2 - x_1). \tag{7.27}$$

If the elapsed time while the particle moves from x_1 to x_2 is $(x_2 - x_1)/v_g$, the time we are seeking, known as the phase time, is

$$\tau = \frac{\partial \theta_t}{\partial \omega}. \tag{7.28}$$

This time has been evaluated for photon transmission problems and electromagnetic pulses through optical media. The calculations describe correctly the experimental results.[4] When the transmission is by quantum tunneling, the transit time is known as the tunneling time.

7.2 Operators and Expectation Values

In contrast to classical mechanics where, given the forces, momenta and initial conditions, it is possible to determine the time evolution of dynamical variables, in the quantum theory the calculation of physical variables proceeds in a different way. Even though the potential energy, and some times also the kinetic term,[5] specify the classical and quantum system, in the quantum theory we seek primarily for the solutions $\psi(\mathbf{r}, \mathbf{t})$ of the Schrödinger equation

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right)\psi(\mathbf{r}, \mathbf{t}) = i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r}, \mathbf{t}), \tag{7.29}$$

while in classical physics we look for the equations of motion for the physical variables. The wave function contains all the available information about the physical state of the system, and the dynamical operators, acting on the wave functions, suck

---

[4] See for example P. Pereyra and H. Simanjuntak, *Phys Rev. E* **75**, 056604 (2007).

[5] For example for systems with central forces, systems in the presence of magnetic fields, systems with more than one particle, etc.

information about the physical variables. In the Schrödinger equation we have two important operators: the Hamiltonian $\widehat{H} = -\hbar^2\nabla^2/2m + V(\mathbf{r})$ and $i\hbar\partial/\partial t$. Both represent the energy. They are equal since $\widehat{H}\psi(\mathbf{r}, t)$ is equal to $i\hbar\partial/\partial t\psi(\mathbf{r}, t)$. Indeed, as mentioned earlier, when the potential energy is independent of time, and we consider the function

$$\psi(\mathbf{r}, \mathbf{t}) = \varphi(\mathbf{r})\tau(t), \tag{7.30}$$

we have

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right)\varphi(\mathbf{r}) = E\varphi(\mathbf{r}); \quad \text{and} \quad i\hbar\frac{\partial}{\partial t}\tau(t) = E\tau(t). \tag{7.31}$$

In these equations the parameter E plays the role of the separation constant, but it is also the physical quantity that comes out when the Hamiltonian, acts on $\varphi(\mathbf{r})$, and comes out also when the operator $i\hbar\partial/\partial t$ acts on $\tau(t)$.

If we now consider the kinetic energy term $\widehat{K} = -\hbar^2\nabla^2/2m$, it is clear that the operator that represents the linear momentum is

$$\widehat{\mathbf{p}} = -i\hbar\nabla. \tag{7.32}$$

We will see later other operators such as the angular momentum

$$\widehat{\mathbf{L}} = \mathbf{r} \times \widehat{\mathbf{p}} = -i\hbar\nabla. \tag{7.33}$$

In short, the physical variables appear in the quantum physics formalism as operators. The replacement of the physical variables by operators is one of the most significant and formal characteristics of the quantum theory. We have already worked out with some operators, in the following we will see what other properties they possess.

At different points of this text, especially in the interpretation of the wave function, some contact was made with statistical concepts. In fact, as will be seen here, some analogies exist between the quantum and the statistical mechanics. In statistical physics, we deal with systems containing a large number of particles, and even though each particle has a definite position and momentum, it is practically impossible to describe the position and momentum for each of them. An individual description becomes meaningless. As a consequence, it is only the statistical description that turns viable, in terms of collective variables. In this theory, the distribution of positions and momenta, $f(\mathbf{r}, p)$, is essential to calculate average values of the physical variables of interest. The average value of a physical variable $A(\mathbf{r}, p)$ is obtained as

$$\overline{A} = \int A(\mathbf{r}, p) f(\mathbf{r}, p) d\mathbf{r} dp. \tag{7.34}$$

In the quantum theory, which in some ways is also a statistical theory, the *expectation values*[6] will be obtained using equivalent methods.

The evaluation of averages is something that happens frequently in our daily life. We know that whenever a variable v, that can take any of a set of values $\{v_j\}$ is measured N times, we find at the end that the output v_1 occurred N_1 times, v_2, N_2 times, and so on. With these results we not only define the relative frequencies $f_j = N_j/N$, that fulfill the relation $\sum_j f_j = 1$, but also evaluate the average value of v, as

$$\overline{v} = \frac{\sum_j v_j N_j}{\sum_j N_j} = \frac{\sum_j v_j f_j}{\sum_j f_j}, \tag{7.35}$$

where the sums extend over the whole set $\{v_j\}$ of possible values.

On the other hand, when we solve the Schrödinger equation (analytically or numerically) and find that a set of eigenstates φ_n and eigenvalues E_n are allowed for the physical system, we have a similar situation. Let us do a simple exercise that will shed light on the issue of the expectation values. Suppose that the set of eigenvalues E_n and (normalized) eigenfunctions φ_n satisfy the equation

$$\widehat{H}\varphi_n(\mathbf{r}) = E_n\varphi_n(\mathbf{r}). \tag{7.36}$$

The most general solution is the superposition

$$\varphi(\mathbf{r}) = \sum_j c_j\varphi_j(\mathbf{r}). \tag{7.37}$$

The action of the Hamiltonian on this function produces the following equality

$$\widehat{H}\varphi(\mathbf{r}) = \sum_j c_j E_j\varphi_j(\mathbf{r}). \tag{7.38}$$

If we multiply this equation by $\varphi^*(\mathbf{r})$ and integrate over the whole configuration space, we have

$$\int \varphi^*(\mathbf{r})\widehat{H}\varphi(\mathbf{r})d\mathbf{r} = \sum_{ij} c_i^* c_j E_j \langle\varphi_i|\varphi_j\rangle = \sum_{ij} c_i^* c_j E_j \delta_{ij}. \tag{7.39}$$

The Kronecker delta δ_{ij} allows us to sum on one of the two indices. If we sum on the index i, we get

$$\int \varphi^*(\mathbf{r})\widehat{H}\varphi(\mathbf{r})d\mathbf{r} = \sum_j |c_j|^2 E_j. \tag{7.40}$$

---

[6] We will use indistinctly the terms expected and expectation value.

This relation provides the meaning of the integral on the left side. We will add one more ingredient. In Chap. 3 we saw that the coefficients $|c_j|^2$ of the superposition give the weight of the state φ_j that participates in $\varphi(\mathbf{r})$. We have shown also that

$$\sum_i |c_i|^2 = 1. \tag{7.41}$$

Using this identity we can rewrite the Eq. (7.40) in the form

$$\langle E\rangle = \frac{\sum_j |c_j|^2 E_j}{\sum_i |c_i|^2} = \int \varphi^*(\mathbf{r})\widehat{H}\varphi(\mathbf{r})d\mathbf{r}. \tag{7.42}$$

The quotient at the center of this triple equation is a mean value; in this case the mean value of the energy. Therefore, the integral on the right defines also the average value of the energy. It contains the operator $\widehat{H}$ that represents the energy, i.e.

$$\langle E\rangle = \langle\widehat{H}\rangle = \int \varphi^*(\mathbf{r})\widehat{H}\varphi(\mathbf{r})d\mathbf{r}. \tag{7.43}$$

This result can be generalized and we can say that the expected value $\langle\widehat{A}\rangle$ of the physical variable A in the state ψ is obtained as follows

$$\langle\widehat{A}\rangle = \int \psi^*(\mathbf{r})\widehat{A}\psi(\mathbf{r})d\mathbf{r}. \tag{7.44}$$

Here $\widehat{A}$ is the operator that represents the physical variable A. If A_n and φ_n fulfill the eigenvalue equation $\widehat{A}\phi_n$=$A_n\varphi_n$, and $\psi(\mathbf{r})$=$\sum_n c_n\varphi_n$, we have in general that

$$\langle A\rangle = \sum_n |c_n|^2 A_n. \tag{7.45}$$

In the next section we will define Hermitian operators. We will introduce some theorems concerning these operators and we will derive the Heisenberg inequality. At the end of this chapter, in the illustrative problems section, we will discuss a little more about the expectation values and the diagonalization method to obtain numerically eigenvalues and eigenfunctions.

7.3 Hermiticity

Operators $\widehat{F}$ whose expectation values are real, are particularly important. These are the so-called Hermitian operators. An operator $\widehat{F}$ is Hermitian if, for arbitrary ψ and φ, it satisfies the relation

$$\int \psi^* \widehat{F} \varphi d\mathbf{r} = \int \varphi \widehat{F}^* \psi^* d\mathbf{r}. \tag{7.46}$$

We will show below that a Hermitian operator, in the sense of this definition, has certainly real eigenvalues. We will also show that the matrix representation of a Hermitian operator is a self-adjoint matrix, i.e.[7]

$$F^\dagger = F. \tag{7.47}$$

1. *The eigenvalues of a Hermitian operator are real*. Suppose we have a Hermitian operator $\widehat{F}$ that satisfies the eigenvalue equation

$$\widehat{F} \varphi_n = f_n \varphi_n, \tag{7.48}$$

with

$$\int \varphi_m^* \varphi_n d\mathbf{r} = \delta_{nm}, \tag{7.49}$$

and

$$\widehat{F}^* \varphi_n^* = f_n^* \varphi_n^*. \tag{7.50}$$

We want to see what follows when $\widehat{F}$ fulfills the relation

$$\int \varphi_m^* \widehat{F} \varphi_n d\mathbf{r} = \int \varphi_n \widehat{F}^* \varphi_m^* d\mathbf{r}. \tag{7.51}$$

Using the eigenvalue Eqs. (7.48) and (7.50), we have

$$\int \varphi_m^* f_n \varphi_n d\mathbf{r} = \int \varphi_n f_m^* \varphi_m^* d\mathbf{r}, \tag{7.52}$$

that can be written as

$$f_n \delta_{nm} = f_m^* \delta_{nm}. \tag{7.53}$$

Therefore

$$f_n = f_n^*. \tag{7.54}$$

The eigenvalues of an operator that fulfill Eq. (7.46), i.e. the eigenvalues of a Hermitian operator, are real.

---

[7] The superscript † represents the joint operations of transposition and complex conjugation ^{T*}.

2. *The matrix representation of a Hermitian operator is self-adjoint*. Suppose we have a basis of functions ϕ_n, which are not eigenfunctions of $\widehat{F}$. Regardless of this, if $\widehat{F}$ is a Hermitian operator in the sense of (7.46), it must satisfy the following relation

$$\int \phi_m^* \widehat{F} \phi_n d\mathbf{r} = \int \phi_n \widehat{F}^* \phi_m^* d\mathbf{r}. \tag{7.55}$$

These integrals depend on the indices m and n and can be treated as matrix elements. Thus, on the left hand side we have the matrix element

$$F_{mn} = \int \phi_m^* \widehat{F} \phi_n d\mathbf{r}, \tag{7.56}$$

and on the right hand side also a matrix element that we can write as follows

$$\int \phi_n \widehat{F}^* \phi_m^* d\mathbf{r} = \left(\int \phi_n^* \widehat{F} \phi_m d\mathbf{r} \right)^* = (F_{nm})^*. \tag{7.57}$$

If we replace these matrix elements in (7.55), we have

$$F_{mn} = (F_{nm})^*, \tag{7.58}$$

or simply

$$F = (F^T)^* = F^\dagger. \tag{7.59}$$

A matrix that is equal to its transpose-conjugate, is a self-adjoint matrix.

7.3.1 Commutation Relations and Fundamental Theorems

The commutation relation of two operators may have important implications. Some operators commute, others not. To express this property in a compact way, it is useful to define the commutator of $\widehat{F}$ and $\widehat{G}$ as

$$\left[\widehat{F}, \widehat{G}\right] = \widehat{F}\widehat{G} - \widehat{G}\widehat{F}. \tag{7.60}$$

This commutator tells us whether the operators $\widehat{F}$ and $\widehat{G}$ commute or not. Clearly, if the operators commute, their commutator is zero, otherwise it is different from zero. Let us give here some simple examples of commutators. To determine the commutator of two operators, it is important to recognize that a commutator is also an operator.

1. $\widehat{x}_i$ and $\widehat{x}_j$ commute. In fact, their commutator is[8]

$$[\widehat{x}_i, \widehat{x}_j] = x_i x_j - x_j x_i = 0 \tag{7.61}$$

2. $\widehat{p}_i$ and $\widehat{p}_j$ commute; their commutator is

$$\begin{aligned}[\widehat{p}_i, \widehat{p}_j] &= \widehat{p}_i \widehat{p}_j - \widehat{p}_j \widehat{p}_i \\ &= -\hbar^2 \left(\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} - \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} \right) \\ &= 0. \end{aligned} \tag{7.62}$$

3. $\widehat{x}_i$ and $\widehat{p}_i$ do not commute; their commutator is

$$\begin{aligned}[\widehat{x}_i, \widehat{p}_i] &= x_i \widehat{p}_i - \widehat{p}_i x_i \\ &= -i\hbar \left(x_i \frac{\partial}{\partial x_i} - 1 - x_i \frac{\partial}{\partial x_i} \right) \\ &= i\hbar. \end{aligned} \tag{7.63}$$

Notice that in evaluating the commutators, the goal is to move the operator to the right, for example $\widehat{p}_i$ in $\widehat{p}_i x_i$ to the right of x_i. The rule is to imagine that on the right of $\widehat{p}_i x_i$ we have a function $f(x_i)$, thus the operator $\widehat{p}_i$, which is a derivative operator, acts on the product of functions at its right based on the well known rules for derivatives, as follows

$$\begin{aligned} -i\hbar \frac{\partial}{\partial x_i} x_i f(x_i) &= -i\hbar \frac{\partial x_i}{\partial x_i} f(x_i) - i\hbar x_i \frac{\partial}{\partial x_i} f(x_i) \\ &= -i\hbar f(x_i) + x_i \widehat{p}_i f(x_i) = [-i\hbar + x_i \widehat{p}_i] f(x_i). \end{aligned} \tag{7.64}$$

Hence,

$$\widehat{p}_i x_i = -i\hbar + x_i \widehat{p}_i, \tag{7.65}$$

which makes the result in (7.63) possible.

4. $\widehat{x}_i$ and $\widehat{p}_j$ commute when $i \neq j$; their commutator is

$$\begin{aligned}[\widehat{x}_i, \widehat{p}_j] &= x_i \widehat{p}_j - \widehat{p}_j x_i \\ &= -i\hbar \left(x_i \frac{\partial}{\partial x_j} - \frac{\partial}{\partial x_j} x_i \right) \\ &= 0. \end{aligned} \tag{7.66}$$

---

[8] The indices i, j, k can take the values 1, 2 and 3. The variables x_i refer to the coordinates x, y and z, such that: $x_1 = x$, $x_2 = y$ and $x_3 = z$. In the same way, $p_1 = p_x$, $p_2 = p_y$ and $p_3 = p_z$.

Theorem 1 *If $\widehat{F}$ and $\widehat{G}$ are any two Hermitian operators, their product $\widehat{F}\widehat{G}$ will be Hermitian only if $\widehat{F}$ and $\widehat{G}$ commute.*

Demonstration: Since $\widehat{F}$ and $\widehat{G}$ are Hermitian operators, we have

$$\widehat{F}^{\dagger} = \widehat{F} \quad \widehat{G}^{\dagger} = \widehat{G}. \tag{7.67}$$

The product $\widehat{F}\widehat{G}$ will be Hermitian if

$$\left(\widehat{F}\widehat{G}\right)^{\dagger} = \widehat{F}\widehat{G}. \tag{7.68}$$

Since

$$\left(\widehat{F}\widehat{G}\right)^{\dagger} = \widehat{G}^{\dagger}\widehat{F}^{\dagger} = \widehat{G}\widehat{F}. \tag{7.69}$$

The product $\left(\widehat{F}\widehat{G}\right)^{\dagger}$ can not be equal to $\widehat{F}\widehat{G}$, unless the operators commute.

Theorem 2 *If the operators $\widehat{F}$ and $\widehat{G}$ are Hermitian and do not commute, their commutator is of the form*

$$\left[\widehat{F}, \widehat{G}\right] = i\widehat{C}, \tag{7.70}$$

with $\widehat{C}$ Hermitian.

Demonstration: We know that

$$\left[\widehat{F}, \widehat{G}\right] = \widehat{F}\widehat{G} - \widehat{G}\widehat{F} \tag{7.71}$$

If we take the transpose conjugate of this commutator, we have

$$\left[\widehat{F}, \widehat{G}\right]^{\dagger} = \left(\widehat{F}\widehat{G} - \widehat{G}\widehat{F}\right)^{\dagger} = \widehat{G}^{\dagger}\widehat{F}^{\dagger} - \widehat{F}^{\dagger}\widehat{G}^{\dagger}. \tag{7.72}$$

Since $\widehat{F}$ and $\widehat{G}$ are Hermitian, the last equation can be written as

$$\left[\widehat{F}, \widehat{G}\right]^{\dagger} = \widehat{G}\widehat{F} - \widehat{F}\widehat{G} = -\left[\widehat{F}, \widehat{G}\right]. \tag{7.73}$$

This shows that the commutator is an anti-Hermitian operator. Hence the commutator $\left[\widehat{F}, \widehat{G}\right]$ can be written as the product $i\widehat{C}$ with $\widehat{C}$ a Hermitian operator. In that case

$$\left[\widehat{F}, \widehat{G}\right]^{\dagger} = \left(i\widehat{C}\right)^{\dagger} = -i\widehat{C} = -\left[\widehat{F}, \widehat{G}\right]. \tag{7.74}$$

As a consequence, *the commutator* of two non-commuting Hermitian operators is an anti-Hermitian operator. An anti-Hermitian operator can always be written as a Hermitian operator multiplied by the imaginary number i. Therefore

$$[\widehat{F}, \widehat{G}] = i\widehat{C}, \tag{7.75}$$

with $\widehat{C}$ Hermitian.

7.4 Deviation, Variance and Dispersion of a Physical Variable

Besides the expectation values, we need to introduce other concepts and relations that are characteristic of the statistical theory. For example, the variance and the standard deviation (or dispersion) of a physical quantity. We begin by introducing the deviation operator

$$\widehat{\delta F} = \widehat{F} - \langle \widehat{F} \rangle, \tag{7.76}$$

which is nothing else than the operator $\widehat{F}$ measured from its expected value. It is clear that

$$\langle \widehat{\delta F} \rangle = \langle \widehat{F} - \langle \widehat{F} \rangle \rangle = \langle \widehat{F} \rangle - \langle \widehat{F} \rangle = 0. \tag{7.77}$$

A quantity closely related with $\widehat{\delta F}$, is the variance of F, denoted as $var F$ or as ΔF^2. The variance of F is defined as

$$var F = \langle (\widehat{F} - \langle \widehat{F} \rangle)^2 \rangle = \langle \widehat{\delta F}^2 \rangle. \tag{7.78}$$

This positive-definite quantity, allows us to calculate another relevant statistical variable: the standard deviation. This quantity tells us how far the actual values of the statistical variable deviate from the mean value. The standard deviation or dispersion ΔF is defined as

$$\Delta F = \sqrt{var F}. \tag{7.79}$$

Notice that ΔF should not be confused with $\langle \widehat{\delta F} \rangle$, which is zero. The standard deviation of a physical quantity is also an important physical variable, useful to analyze the dispersion of the physical quantities. Its counterpart in the experimental physics is the uncertainty or measurement error. Unavoidable variations that should, necessarily, be taken into account when experimental results are compared with theoretical predictions. It is important to distinguish the standard deviation from the expected value $\langle \widehat{\delta F} \rangle$ of the deviation.

We will see now some theorems related to these statistical variables in the context of quantum theory.

Theorem 3 *When a system is precisely in one of the eigenstates of an operator $\widehat{F}$, the variance of F is zero.*

Demonstration: Suppose that $\widehat{F}\varphi_n = F_n\varphi_n$. When the system is in one of the eigenstates, say in the eigenstate φ_n, the expected value of $\widehat{F}$ will be given by

$$\langle\widehat{F}\rangle = \int \varphi_n^* \widehat{F}\varphi_n dx = F_n. \tag{7.80}$$

and F_n is the eigenvalue of $\widehat{F}$, corresponding to the eigenfunction φ_n. Similarly,

$$\langle\widehat{F}^2\rangle = \int \varphi_n^* \widehat{F}\widehat{F}\varphi_n dx = F_n \int \varphi_n^* \widehat{F}\varphi_n dx = F_n^2. \tag{7.81}$$

Substituting these expected values in the variance of $\widehat{F}$, we have

$$(\Delta F)^2 = \langle\, \widehat{F}^2\,\rangle - \langle \widehat{F}\,\rangle^2 = F_n^2 - F_n^2 = 0. \tag{7.82}$$

This proves the theorem.

Theorem 4 *If a system is in an eigenstate of $\widehat{F}$, the dispersion of this variable or of any other that commutes with $\widehat{F}$, is zero.*

Demonstration: It has been shown that if the system is in an eigenstate of $\widehat{F}$, there is no dispersion for F, i.e.

$$(\Delta F)^2 = 0. \tag{7.83}$$

Let $\widehat{G}$ be an operator that commutes with $\widehat{F}$. One can easily verify that if

$$\widehat{F}\varphi_n = f_n\varphi_n, \tag{7.84}$$

then

$$\widehat{G}\widehat{F}\varphi_n = f_n\widehat{G}\varphi_n. \tag{7.85}$$

Since we assume that $\widehat{G}$ commutes with $\widehat{F}$, the last equation can be written in the form

$$\widehat{F}(\widehat{G}\varphi_n) = f_n(\widehat{G}\varphi_n). \tag{7.86}$$

This equation has the appearance of an eigenvalue equation where $\widehat{G}\varphi_n$ behaves like the eigenfunction corresponding to the eigenvalue f_n, thus, the function $\widehat{G}\varphi_n$ can be written as

$$\widehat{G}\varphi_n = c\,\varphi_n. \tag{7.87}$$

This is an eigenvalue equation where the constant c is nothing else than the eigenvalue g_n of $\widehat{G}$, with eigenfunction φ_n, as stated in the theorem. Thus

$$\widehat{F}\widehat{G}\varphi_n = \widehat{G}\widehat{F}\varphi_n = g_n f_n \varphi_n, \tag{7.88}$$

If we apply the previous theorem for $\widehat{G}$, it is also possible to conclude that the dispersion of the operator $\widehat{G}$, that commutes and shares eigenfunctions with $\widehat{F}$, is also zero, i.e.

$$(\varDelta G)^2 = 0. \tag{7.89}$$

Corollary *If a system is in an eigenstate of the Hamiltonian*, the energy is a variable without dispersion, and all variables whose operators commute with $\widehat{H}$ are also dispersionless variables.

7.5 Heisenberg's Inequality

The Heisenberg inequality, also known as Heisenberg's Uncertainty Principle, is a relation for the dispersions of two physical variables, one the canonical conjugate of the other, whose Hermitian operators do not commute. Let the operators $\widehat{f}$ and $\widehat{g}$ be with this feature, whose commutator is

$$[\widehat{f}, \widehat{g}] = i\widehat{h}. \tag{7.90}$$

If we define the operators $\widehat{\delta f} = \widehat{f} - \langle\widehat{f}\rangle$ and $\widehat{\delta g} = \widehat{g} - \langle\widehat{g}\rangle$, it is easy to show that their commutator satisfies also the relation

$$[\widehat{\delta f}, \widehat{\delta g}] = i\widehat{h}. \tag{7.91}$$

with $\widehat{h}$ a Hermitian operator. So far we have seen that the commutator $\widehat{h}$ of canonically conjugated physical quantities, like $\widehat{x}_i$ and $\widehat{p}_j$, is just the Planck constant $\hbar$. Therefore, in what follows, we will replace $\widehat{h}$ by $\hbar$.

It is easy to show that the product of a pair of operators $\widehat{\delta f}$ and $\widehat{\delta g}$ can be expressed in the form

$$\widehat{\delta f}\ \widehat{\delta g} = \frac{1}{2}[\widehat{\delta f}, \widehat{\delta g}] + \frac{1}{2}\{\widehat{\delta f}, \widehat{\delta g}\} \tag{7.92}$$

where $\{a, b\}$ is an anti-commutator defined as $ab + ba$. On the other hand, the variances $\langle\widehat{\delta f}^2\rangle$ and $\langle\widehat{\delta g}^2\rangle$ satisfy the Schwartz inequality

$$\langle\widehat{\delta f}^2\rangle\langle\widehat{\delta g}^2\rangle \geq |\langle\widehat{\delta f}\ \widehat{\delta g}\rangle|^2. \tag{7.93}$$

Therefore

$$\langle\widehat{\delta f}^{\,2}\rangle\langle\widehat{\delta g}^{\,2}\rangle \geq \frac{1}{4}|\langle[\widehat{\delta f}, \widehat{\delta g}] + \{\widehat{\delta f}, \widehat{\delta g}\}\rangle|^2. \tag{7.94}$$

To obtain the Heisenberg inequality we need to simplify even more the previous expression. One can show that the anti-commutator of two Hermitian operators is also Hermitian. Suppose then that $\{\delta\widehat{f}, \delta\widehat{g}\}\rangle = \widehat{k}$. Since the expectation values of Hermitian operators are real, Eq. (7.94) takes the form

$$\langle\widehat{\delta f}^{2}\rangle\langle\widehat{\delta g}^{2}\rangle \geq \frac{1}{4}|\langle i\hbar + \widehat{k}\rangle|^2 = \frac{1}{4}|i\hbar + \langle\widehat{k}\rangle|^2. \tag{7.95}$$

which obviously can be written as

$$\langle\widehat{\delta f}^{2}\rangle\langle\widehat{\delta g}^{2}\rangle \geq \frac{1}{4}\hbar^2. \tag{7.96}$$

This uncertainty is a relation of variances. If we extract the square root, we have the uncertainty relation

$$\Delta f \Delta g \geq \frac{\hbar}{2}, \tag{7.97}$$

of standard deviations, or the uncertainty relation of dispersions Δf and Δg. This is popularly known as the *Heisenberg's Uncertainty Principle*. This relation shows that the product of dispersions of two variables, one the canonical conjugate of the other, can not be less than $\hbar/2$. The physical interpretation of this inequality has been controversial, diverse and a source of exotic ideas. The most widespread interpretation suggests the impossibility of determining simultaneously and accurately, with zero dispersion, the values of two canonical conjugate variables, such as position and momentum, energy and time, etc. Clearly, if the dispersion of one of the variables is zero, the inequality suggests that the dispersion of its canonical conjugate will be ∞. This is an overstated interpretation. If we accept that the counterpart of the theoretical dispersion could be the experimental error, the inequality leaves us a large margin for an accurate determination of the canonical conjugate variables. Indeed, if the error determining one variable, say the energy is as small as $\sqrt{\hbar/2} \sim 10^{-18}$ J, the uncertainty relation says that the accuracy of determining its canonical conjugate, in this case time, must be at least of the order of $\sqrt{\hbar/2} \sim 10^{-16}$ s. Both quantities are very, very small. Of course, if the experimental methods were so precise that the dispersion of one variable tends to zero, its canonical conjugate will, according to the uncertainty relation, have a dispersion that tends to ∞. With the accuracy of the current experimental methods the dispersions product is orders of magnitude greater than the lower limit $\hbar/2 \approx 10^{-34}$ Js.

7.6 Time Evolution; The Schrödinger and Heisenberg Pictures

7.6.1 The Schrödinger Picture and the Unitary Time Evolution Operator

Physical systems generally evolve in space and time. The solution of the time dependent Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{r},t) = \widehat{H}\psi(\mathbf{r},t). \tag{7.98}$$

is a function of the spatial coordinates and time. When the Hamiltonian is independent of time, the space and time evolution of the wave function is given by

$$\psi(\mathbf{r},t) = e^{-i\widehat{H}t/\hbar}\psi(\mathbf{r},0). \tag{7.99}$$

In this representation the wave function depends on time and the operators are independent of time, except of course the *time evolution operator* $e^{-i\widehat{H}t/\hbar}$ that is part of the wave function $\psi(\mathbf{r},t)$). This is called *the Schrödinger picture*. The time evolution operator is an important quantity of this representation. We will derive this operator based on the infinitesimal generators of a unitary transformation. If we write the Schrödinger equation as

$$\lim_{\delta t\to 0}\frac{\psi(\mathbf{r},\delta t)-\psi(\mathbf{r},0)}{\delta t} = -\frac{i\widehat{H}}{\hbar}\psi(\mathbf{r},0), \tag{7.100}$$

the wave function $\psi(\mathbf{r},\delta t)$, for δt infinitesimal, can be written as

$$\psi(\mathbf{r},\delta t) = \left(1-\tfrac{i\widehat{H}\delta t}{\hbar}\right)\psi(\mathbf{r},0). \tag{7.101}$$

Which means that the Hamiltonian $\widehat{H}$ is the generator of the infinitesimal time evolution operator

$$\widehat{U}(\delta t) = \left(1-\frac{i\widehat{H}\delta t}{\hbar}\right). \tag{7.102}$$

Since

$$1-\frac{i\widehat{H}\delta t}{\hbar}+\mathcal{O}(\delta t^2) \simeq e^{-i\hat{H}\delta t/\hbar}, \tag{7.103}$$

the infinitesimal time evolution operator can be written also as

$$\widehat{U}(\delta t) = e^{-i\hat{H}\delta t/\hbar}. \tag{7.104}$$

Based on this operator, we can obtain the time evolution operator $\widehat{U}(t)$ for any time t. If $t = \lim_{N\to\infty} N\delta t$, the time evolution operator is then given by

$$\widehat{U}(t) = \lim_{N\to\infty} \widehat{U}(\delta t)^N = \lim_{N\to\infty}\left(e^{-i\hat{H}t/\hbar N}\right)^N = e^{-i\hat{H}t/\hbar}. \tag{7.105}$$

For a Hermitian Hamiltonian $\widehat{H}^\dagger = \widehat{H}$, the time evolution operator has the property

$$\widehat{U}^\dagger(t) = e^{i\hat{H}t/\hbar} = \widehat{U}^{-1}(t) \tag{7.106}$$

of any unitary operator. Because of this property, the evolution operator transforms the wave function keeping its norm constant. This means that the probability density evaluated with

$$\psi(\mathbf{r}, t) = \widehat{U}(t)\psi(\mathbf{r}, 0) = e^{-i\hat{H}t/\hbar}\psi(\mathbf{r}, 0) \tag{7.107}$$

is the same as the probability density evaluated with $\psi(\mathbf{r}, 0)$. Indeed

$$|\psi(\mathbf{r}, t)|^2 = |e^{-i\hat{H}t/\hbar}\psi(\mathbf{r}, 0)|^2 = |\psi(\mathbf{r}, 0)|^2. \tag{7.108}$$

When the potential function is independent of time, the differential equation is separable. In that case, and using the notation of Sect. 7.2 with $\tau(0) = 1$ and $\widehat{H}\varphi(\mathbf{r}) = E\varphi(\mathbf{r})$, the wave function $\psi(\mathbf{r}, t)$ can be written as

$$\psi(\mathbf{r}, t) = e^{-i\widehat{H}t/\hbar}\tau(0)\varphi(\mathbf{r}) = e^{-iEt/\hbar}\varphi(\mathbf{r}), \tag{7.109}$$

a function that we met before, with $\varphi(\mathbf{r})$ the solution of the stationary Schrödinger equation.

If we have a time dependent wave function, the expectation values will also, in principle, evolve in time. To represent the expectation values in time-dependent states, we will use when necessary the notation $\langle\ \rangle_t$. The expected value of a physical quantity F, in the state $\psi(\mathbf{r}, t)$, is given by

$$\langle F\rangle_t = \langle\psi(\mathbf{r}, t)|\widehat{F}|\psi(\mathbf{r}, t)\rangle = \int \psi^*(\mathbf{r}, t)\widehat{F}\psi(\mathbf{r}, t)d^3\mathbf{r}. \tag{7.110}$$

When the operator $\widehat{F}$ is independent of time, the time dependence of the expectation value comes entirely from the wave function $\psi(\mathbf{r}, t)$.

7.6.2 Heisenberg's Picture

An alternative and useful picture for time-dependent calculations is *the Heisenberg picture* that *transfers the time dependence from the wave functions to the operators*. If we consider the expected value in (7.110), and write it as

$$\langle F\rangle_t = \langle\psi(\mathbf{r},0)\,|e^{i\widehat{H}t/\hbar}\widehat{F}e^{-i\widehat{H}t/\hbar}\,|\psi(\mathbf{r},0) = \langle\widehat{F}(t)\rangle_0. \tag{7.111}$$

We can see this quantity from a different perspective: as the expected value of the time dependent operator $\widehat{F}(t)$ in the time independent state $\psi(\mathbf{r},0)$. This operator, can also be obtained if we write the expected value in (7.110) as

$$\langle F\rangle_t = \langle\psi(\mathbf{r},t)|\widehat{F}\,|\psi(\mathbf{r},t)\rangle = \langle\psi(\mathbf{r},t)|\widehat{U}\widehat{U}^\dagger\widehat{F}\widehat{U}\widehat{U}^\dagger|\psi(\mathbf{r},t)\rangle, \tag{7.112}$$

where we have introduced the unit operator $\widehat{U}\widehat{U}^\dagger$. Since

$$\langle\psi(\mathbf{r},0)| = \langle\psi(\mathbf{r},t)|\widehat{U} \qquad \text{and} \qquad \widehat{U}^\dagger|\psi(\mathbf{r},t)\rangle = |\psi(\mathbf{r},0)\rangle, \tag{7.113}$$

the expected value can also be written as

$$\langle F\rangle_t = \langle\psi(\mathbf{r},0)|\widehat{U}^\dagger(t)\widehat{F}\widehat{U}(t)|\psi(\mathbf{r},0)\rangle = \langle\widehat{U}^\dagger(t)\widehat{F}\widehat{U}(t)\rangle_0, \tag{7.114}$$

which is the expected value of the time dependent operator

$$\widehat{F}(t) = \widehat{U}^\dagger(t)\widehat{F}\widehat{U}(t). \tag{7.115}$$

This leads us to *the Heisenberg picture*. In this representation, the expected value $\langle\ \rangle_0$ of a time dependent operator defined as

$$\widehat{F}(t) = e^{i\widehat{H}t/\hbar}\widehat{F}e^{-i\widehat{H}t/\hbar}, \tag{7.116}$$

is evaluated in the state $\psi(\mathbf{r},0)$, that remains fixed. Time dependent operators are sometimes denoted in a more compact notation as $\widehat{F}_{\mathrm{H}}$. Before discussing the behavior of time dependent operators, let us consider the expected value (7.111) to derive a relation that we will use later. When the potential is time independent and $\tau(0) = 1$, we can write the following equality (when we write ψ, we understand that we have $\psi(\mathbf{r},t)$, when we write ψ_0, it means $\psi(\mathbf{r},0)$ and φ for $\varphi(\mathbf{r})$ independent of time)

$$\langle\psi\,|\widehat{F}\,|\psi\rangle = \langle\varphi|\widehat{F}(t)|\varphi\rangle. \tag{7.117}$$

On the other hand, the time derivative of $\widehat{F}_{\mathrm{H}}$, gives us

$$\frac{d\widehat{F}(t)}{dt} = \frac{i}{\hbar}e^{i\widehat{H}t/\hbar}\widehat{H}\widehat{F}e^{-i\widehat{H}t/\hbar} + e^{i\widehat{H}t/\hbar}\frac{\partial\widehat{F}}{\partial t}e^{-i\widehat{H}t/\hbar} - \frac{i}{\hbar}e^{i\widehat{H}t/\hbar}\widehat{F}\widehat{H}e^{-i\widehat{H}t/\hbar}. \tag{7.118}$$

Since $e^{i\widehat{H}t/\hbar}$ commutes with $\widehat{H}$, the previous equation can be written also as

$$\frac{d\widehat{F}(t)}{dt} = e^{i\widehat{H}t/\hbar}\frac{\partial\widehat{F}}{\partial t}e^{-i\widehat{H}t/\hbar} + \frac{i}{\hbar}\left[\widehat{H}, \widehat{F}(t)\right]. \tag{7.119}$$

This equation of motion is an important result. According to this equation, quantities that have operators not explicitly dependent on time and commute with the Hamiltonian $\widehat{H}$, are constants of motion. Let us see if this is also valid for the expectation values. If we consider the expected value in (7.110), its time derivative

$$\frac{d}{dt}\langle F\rangle_t = \frac{d}{dt}\langle\psi(\mathbf{r}, t)|\widehat{F}\,|\psi(\mathbf{r}, t)\rangle = \frac{d}{dt}\langle\psi|\widehat{F}\,|\psi\rangle, \tag{7.120}$$

can be written as follows:

$$\frac{d}{dt}\langle F\rangle_t = \frac{\partial\langle\psi|}{\partial t}\widehat{F}\,|\psi\rangle + \langle\psi|\frac{\partial\widehat{F}}{\partial t}|\psi\rangle + \langle\psi|\widehat{F}\frac{\partial|\psi\rangle}{\partial t}. \tag{7.121}$$

Using the time-dependent Schrödinger equation, we have

$$\frac{d}{dt}\langle F\rangle_t = \langle\psi|\frac{\partial\widehat{F}}{\partial t}|\psi\rangle + \frac{i}{\hbar}\langle\psi|\left[\widehat{H}, \widehat{F}\right]|\psi\rangle. \tag{7.122}$$

This is the equation of motion of the expectation value of F. It is clear here that the expected value $\langle F\rangle_t$ is constant when $\widehat{F}$ does not depend explicitly on time and commutes with $\widehat{H}$. But, what happens if we have written the operator in the Heisenberg picture? To close the frame and avoid confusion we will obtain this equation of motion in the Heisenberg representation. If we use the expectation value in (7.111), the time derivative can be written as

$$\frac{d}{dt}\langle F\rangle_t = \langle\psi_0|\,\frac{d}{dt}\widehat{F}(t)\,|\psi_0\rangle, \tag{7.123}$$

which, with the help of Eq. (7.119), takes the form

$$\frac{d}{dt}\langle F\rangle_t = \langle\psi_0|\,e^{i\widehat{H}t/\hbar}\frac{\partial\widehat{F}}{\partial t}e^{-i\widehat{H}t/\hbar}|\psi_0\rangle + \frac{i}{\hbar}\langle\psi_0|\left[\widehat{H}, \widehat{F}(t)\right]|\psi_0\rangle, \tag{7.124}$$

or simply

$$\frac{d}{dt}\langle F\rangle_t = \langle\psi|\frac{\partial\widehat{F}}{\partial t}|\psi\rangle + \frac{i}{\hbar}\langle\psi_0|\left[\widehat{H}, \widehat{F}(t)\right]|\psi_0\rangle. \tag{7.125}$$

This equation is exactly like the equation of motion (7.122). Therefore, the equation of motion of the expected value of a physical quantity $\widehat{F}$, that does not depend explicitly on time, can be written as

$$\frac{d}{dt}\langle F\rangle_t = \frac{i}{\hbar}\langle\varphi|[\widehat{H},\widehat{F}(t)]|\varphi\rangle, \tag{7.126}$$

or just as

$$\frac{d}{dt}\langle F\rangle_t = \frac{i}{\hbar}\langle[\widehat{H},\widehat{F}]\rangle_t. \tag{7.127}$$

Notice that we can also write

$$\frac{d}{dt}\langle\psi|\widehat{F}|\psi\rangle = \langle\varphi|\frac{d\widehat{F}(t)}{dt}|\varphi\rangle. \tag{7.128}$$

Theorem 5 The Ehrenfest theorem. *The equations of motion for the expectation values* $\langle\mathbf{r}\rangle_t$ *and* $\langle\widehat{\mathbf{p}}\rangle_t$ *are similar to the classical equations of motion of* **r** *and* **p**.

Demonstration: We will obtain here the equation of motion for x and p_x. The extension to 3D is straightforward. If we apply the equation of motion (7.126) to x, we have

$$\frac{d}{dt}\langle x\rangle_t = \frac{i}{\hbar}\langle\psi|[\widehat{H},x]|\psi\rangle. \tag{7.129}$$

In Sect. 7.3 we obtained the conmutators[9]

$$[x_i,p_j] = i\hbar\delta_{i,j} \quad i,j=1,2,3 \tag{7.130}$$

which we will use to evaluate $[\widehat{H},x]$. Since $V(\mathbf{r})$ commutes with x, the commutator takes the form

$$[\widehat{H},x] = \frac{1}{2m}(\widehat{p}_x^2 x - x\widehat{p}_x^2). \tag{7.131}$$

If in the last expression, we subtract and add the operator $x\widehat{p}_x x$, it becomes

$$\begin{aligned}[\widehat{H},x] &= \frac{1}{2m}(\widehat{p}_x[\widehat{p}_x,x]-[x,\widehat{p}_x]\widehat{p}_x)\\ &= -i\frac{\hbar}{m}\widehat{p}_x\end{aligned} \tag{7.132}$$

If we replace this in (7.129), we have the equation of motion

$$m\frac{d}{dt}\langle x\rangle_t = \langle\widehat{p}_x\rangle_t \tag{7.133}$$

which corresponds to the classical definition $p_x = m\frac{dx}{dt}$.

---

[9] We write $x_1=x,\ldots,x_3=z$ and $p_1=p_x,\ldots,p_3=p_z$.

Let us now consider the equation of motion

$$\frac{d}{dt}\langle\widehat{p}_x\rangle_t = \frac{i}{\hbar}\langle\psi|[\widehat{H}, \widehat{p}_x]|\psi\rangle. \tag{7.134}$$

Here we need also to transform the commutator $[\widehat{H}, \widehat{p}_x]$. If we write explicitly the Hamiltonian and take into account that $\widehat{p}^2 = \widehat{p}_x^2 + \widehat{p}_y^2 + \widehat{p}_z^2$ commutes with $\widehat{p}_x$, we have

$$\begin{aligned}[\widehat{H}, \widehat{p}_x] &= V(\mathbf{r})\widehat{p}_x - \widehat{p}_x V(\mathbf{r}) \\ &= V(\mathbf{r})\widehat{p}_x + i\hbar\frac{\partial V(\mathbf{r})}{\partial x} - V(\mathbf{r})\widehat{p}_x. \end{aligned} \tag{7.135}$$

Therefore

$$\frac{d}{dt}\langle\widehat{p}_x\rangle_t = -\left\langle\frac{\partial V}{\partial x}\right\rangle_t = \langle F_x\rangle_t, \tag{7.136}$$

which corresponds to the second law of Newton.

7.7 Position and Momentum in the Momentum Representation

At the end of Chap. 3, we briefly recalled the Fourier transformation as a specific example to change from the coordinates to the momentum representation. In this section we will see how the momentum and position operators are expressed in the momentum representation.

Suppose that we have the basis $\{|x\rangle\}$, in coordinate representation, and the basis $\{|p_x\rangle\}$ in the momentum representation. To keep a light notation, we will use p instead of p_x. Let us see now how the operators $\widehat{x}$ and $\widehat{p}$ look like in the space and momentum representation. Since an operator is diagonal in its own representation, we have

$$\langle x'|\widehat{x}|x\rangle = x\delta(x'-x) \qquad \text{and} \qquad \langle p'|\widehat{p}|p\rangle = p\delta(p'-p). \tag{7.137}$$

The first of these equalities can be written as

$$\langle x'|\widehat{x}|x\rangle = x\delta(x'-x) = x\int\langle x'|p\rangle\langle p|x\rangle dp. \tag{7.138}$$

If we recall that $\int|p\rangle\langle p|dp = 1$, the last equation can be written in the form

$$\langle x'|\widehat{x}|x\rangle = x\langle x'|x\rangle, \tag{7.139}$$

thus

$$\widehat{x}|x\rangle = x|x\rangle. \tag{7.140}$$

$|x\rangle$ is the eigenfunction of $\widehat{x}$ with eigenvalue x. In the same way

$$\langle p'|\widehat{p}|p\rangle = p\delta(p'-p) = p\int \langle p'|x\rangle\langle x|p\rangle dx. \tag{7.141}$$

We now use $\int |x\rangle\langle x|dx = 1$, and we have

$$\widehat{p}|p\rangle = p|p\rangle. \tag{7.142}$$

$|p\rangle$ is the eigenfunction of $\widehat{p}$ with eigenvalue p. To make clear how one operates with this type of transformations, let us see how one obtains the operator $\widehat{p}$ in the coordinate representation. Using the matrix element $\langle x'|\widehat{p}|x\rangle$, and the unit operator, we easily obtain the identity

$$\begin{aligned}\langle x'|\widehat{p}|x\rangle &= \int dp'dp\langle x'|p'\rangle\langle p'|\widehat{p}|p\rangle\langle p|x\rangle \\ &= \int dp'dp\langle x'|p'\rangle\langle p'|p|p\rangle\langle p|x\rangle.\end{aligned} \tag{7.143}$$

If we take into account that p, in the last equation, is a scalar and can change its position inside the integral, and we remember that $\delta(p'-p) = \langle p'|p\rangle$, the previous equation becomes

$$\langle x'|\widehat{p}|x\rangle = \int dp'dp\langle x'|p'\rangle p\delta(p'-p)\langle p|x\rangle. \tag{7.144}$$

After integration on p', we have

$$\langle x'|\widehat{p}|x\rangle = \int dp\langle x'|p\rangle p\langle p|x\rangle. \tag{7.145}$$

In this integral we have the factor $p\langle p|x\rangle$ that can be written as follows

$$p\langle p|x\rangle = p\frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}} = -i\hbar\frac{\partial}{\partial x}\frac{e^{ipx/\hbar}}{\sqrt{2\pi\hbar}} = -i\hbar\frac{\partial}{\partial x}\langle p|x\rangle. \tag{7.146}$$

If we substitute this into (7.145), we have

$$\langle x'|\widehat{p}|x\rangle = -i\hbar\frac{\partial}{\partial x}\int dp\langle x'|p\rangle\langle p|x\rangle, \tag{7.147}$$

that, finally, we write in the form

$$\langle x'|\widehat{p}|x\rangle = \langle x'|\left(-i\hbar\frac{\partial}{\partial x}\right)|x\rangle. \tag{7.148}$$

This gives us the momentum operator that we know already. Let us see the operator $\widehat{x}$ in the momentum representation. Here also, we can start with the matrix element

$$\begin{aligned}\langle p'|\widehat{x}|p\rangle &= \int dx'dx\langle p'|x'\rangle\langle x'|\widehat{x}|x\rangle\langle x|p\rangle \\ &= \int dx'dx\langle p'|x'\rangle\langle x'|x|x\rangle\langle x|p\rangle,\end{aligned} \tag{7.149}$$

that using the identity $\langle x'|x\rangle = \delta(x' - x)$, transforms into

$$\begin{aligned}\langle p'|\widehat{x}|p\rangle &= \int dx'dx\langle p'|x'\rangle x\delta(x' - x)\langle x|p\rangle \\ &= \int dx\langle p'|x\rangle x\langle x|p\rangle.\end{aligned} \tag{7.150}$$

Now, it is the factor $x\langle x|p\rangle$ which we need to write as

$$x\langle x|p\rangle = x\frac{e^{-ipx/\hbar}}{\sqrt{2\pi\hbar}} = i\hbar\frac{\partial}{\partial p}\frac{e^{-ipx/\hbar}}{\sqrt{2\pi\hbar}} = i\hbar\frac{\partial}{\partial p}\langle x|p\rangle. \tag{7.151}$$

If we substitute this into (7.150), we have

$$\langle p'|\widehat{x}|p\rangle = i\hbar\frac{\partial}{\partial p}\int dx\langle p'|x\rangle\langle x|p\rangle, \tag{7.152}$$

which tells us that the position operator in the momentum representation is

$$\widehat{x} = i\hbar\frac{\partial}{\partial p}. \tag{7.153}$$

Therefore, the position operator $\widehat{\mathbf{r}}$ in the three-dimensional space is

$$\widehat{\mathbf{r}} = i\hbar\nabla_p, \tag{7.154}$$

similar to the momentum operator in the coordinates space.

7.8 Solved Problems

Exercise 25 Let the set of functions $\{\psi_n(\mathbf{r}, t)\}$ be such that $\psi_n(\mathbf{r}, t) = e^{-i\omega_n t}\varphi_n(\mathbf{r})$, and let $\widehat{f}$ be a Hermitian operator. Prove that

$$\langle\varphi_m|\frac{d\widehat{f}}{dt}|\varphi_n\rangle = i\frac{E_m - E_n}{\hbar}\langle\varphi_m|\widehat{f}(t)|\varphi_n\rangle. \tag{7.155}$$

Solution To prove this equality we recall that the time derivative

$$\frac{d}{dt}\langle\psi_m|\widehat{f}|\psi_n\rangle, \tag{7.156}$$

written in the form

$$\frac{d}{dt}\langle\psi_m|\widehat{f}|\psi_n\rangle = \frac{d}{dt}\langle\varphi_m|e^{i\omega_m t}\widehat{f}e^{-i\omega_n t}|\varphi_n\rangle, \tag{7.157}$$

can be written in the Heisenberg representation as

$$\frac{d}{dt}\langle\psi_m|\widehat{f}|\psi_n\rangle = \langle\varphi_m|\frac{d\widehat{f}(t)}{dt}|\varphi_n\rangle. \tag{7.158}$$

Since

$$\frac{d\widehat{f}(t)}{dt} = \frac{d}{dt}e^{i(\omega_m-\omega_n)t}\widehat{f} = i(\omega_m-\omega_n)e^{i\omega_m t}\widehat{f}e^{-i\omega_n t}, \tag{7.159}$$

we have the relation we are looking for:

$$\langle\varphi_m|\frac{d\widehat{f}(t)}{dt}|\varphi_n\rangle = i\frac{E_m-E_n}{\hbar}\langle\varphi_m|\widehat{f}(t)|\varphi_n\rangle. \tag{7.160}$$

Exercise 26 Apply the result of the previous exercise to x and $\widehat{p_x}$ and show that

$$\langle\varphi_m|\ddot{x}|\varphi_n\rangle = -\left(\frac{E_m-E_n}{\hbar}\right)^2\langle\varphi_m|x(t)|\varphi_n\rangle. \tag{7.161}$$

Solution If we apply (7.160) to the operator x, we have

$$\langle\varphi_m|\dot{x}|\varphi_n\rangle = i\frac{E_m-E_n}{\hbar}\langle\varphi_m|x(t)|\varphi_n\rangle, \tag{7.162}$$

while applied to the operator $\widehat{p}_x$, it gives

$$\langle\varphi_m|\dot{\widehat{p}}_x|\varphi_n\rangle = i\frac{E_m-E_n}{\hbar}\langle\varphi_m|\widehat{p}_x(t)|\varphi_n\rangle. \tag{7.163}$$

With the last two equations, and with the equation of motion for $\langle x\rangle_t$, we easily obtain the relation

$$\langle\varphi_m|\ddot{x}|\varphi_n\rangle = -\left(\frac{E_m-E_n}{\hbar}\right)^2\langle\varphi_m|x(t)|\varphi_n\rangle. \tag{7.164}$$

Since $E_m - E_n = \hbar\omega_{mn}$, the previous equations can be written also in terms of w_{mn} as

$$\langle \varphi_m | \ddot{x} | \varphi_n \rangle = -\omega_{mn}^2 \langle \varphi_m | x(t) | \varphi_n \rangle. \tag{7.165}$$

Exercise 27 How do we evaluate the energy eigenvalues and the corresponding eigenfunctions when the Schrödinger equation $\widehat{H}\psi(\mathbf{r}) = E\psi(\mathbf{r})$ cannot be solved analytically?

Solution If the eigenvalues and eigenfunctions can not be obtained analytically, one can proceed as follows. We consider first a complete set of orthonormal functions $\{\phi_i\}$, and use this basis to express the unknown solution $\psi(\mathbf{r})$ as the superposition

$$\psi(\mathbf{r}) = \sum_i c_i \phi_i(\mathbf{r}), \tag{7.166}$$

where the coefficients c_i do not depend on $\mathbf{r}$. The expectation value for the energy in the quantum state $\psi(\mathbf{r})$ can, formally, be written as

$$\langle E \rangle = \int \psi^* H \psi d\mathbf{r} = \sum_{ij} c_i^* c_j \int \phi_i^* \widehat{H} \phi_j d\mathbf{r}. \tag{7.167}$$

Since the functions ϕ_j are not eigenfunctions of $\widehat{H}$, we can not replace $\widehat{H}\phi_j$ by $E_j\phi_j$. The integrals $\int \phi_i^* \widehat{H} \phi_j d\mathbf{r}$ are numbers with values that depend on the functions ϕ_i and ϕ_j. Hence we can represent these integrals as the matrix elements

$$H_{ij} = \int \phi_i^* \widehat{H} \phi_j d\mathbf{r}. \tag{7.168}$$

These numbers define the *matrix representation H of the Hamiltonian $\widehat{H}$*. The specific matrix H depends on the basis $\{\phi_i\}$, but the eigenvalues that will be obtained should be independent of the basis $\{\phi_i\}$ that we have chosen. The matrix H possesses, generally, well defined properties and symmetries, related to the presence or not of basic physical symmetries such as time reversibility, spin rotation, and flux conservation. If we define the vector V_c as

$$V_c = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} \quad V_c^\dagger = (c_1^*, c_2^*, \ldots, c_N^*), \tag{7.169}$$

we can express the expected energy in (7.167) as

$$\langle E \rangle = \sum_{i,j} c_i^* c_j H_{ij} = V_c^\dagger H V_c. \tag{7.170}$$

Depending on the physical system and the symmetries present, the matrix that diagonalize H might be unitary, orthogonal or symplectic. Let us suppose that a unitary matrix U diagonalizes H, i.e.

$$U^{\dagger} H U = \varepsilon. \tag{7.171}$$

This diagonalization is an important operation, equivalent to deriving the energy eigenvalues. The elements of ε are the eigenvalues and the columns of U the corresponding eigenvectors. To visualize this statement let us observe with some detail the previous equation. If we multiply it with matrix U we have (using the convention of sums over repeated indices) that

$$HU = U\varepsilon \quad \longrightarrow \quad H_{ij} U_{jk} = U_{ij} \varepsilon_{jj} \delta_{jk}. \tag{7.172}$$

In the last equation, thanks to the Kronecker delta one can perform one of the sums. It then reduces to

$$H_{ij} U_{jk} = U_{ik} \varepsilon_k. \tag{7.173}$$

If we write this equation in the form

$$\begin{pmatrix} H_{1j}U_{j1} & H_{1j}U_{j2} & \cdots & H_{1j}U_{jN} \\ H_{2j}U_{j1} & H_{2j}U_{j2} & \cdots & H_{2j}U_{jN} \\ \vdots & \vdots & \cdots & \vdots \\ H_{Nj}U_{j1} & H_{Nj}U_{j2} & \cdots & H_{Nj}U_{jN} \end{pmatrix} = \begin{pmatrix} U_{11}\varepsilon_1 & U_{12}\varepsilon_2 & \cdots & U_{1N}\varepsilon_N \\ U_{21}\varepsilon_1 & U_{22}\varepsilon_2 & \cdots & U_{2N}\varepsilon_N \\ \vdots & \vdots & \cdots & \vdots \\ U_{N1}\varepsilon_1 & U_{N2}\varepsilon_2 & \cdots & U_{NN}\varepsilon_N \end{pmatrix} \tag{7.174}$$

and, if we denote the kth column of U as $U_{\cdot k}$, it is easy to see that the last equation leads precisely to

$$HU_{\cdot k} = \varepsilon_k U_{\cdot k}. \tag{7.175}$$

This is nothing else than the eigenvalue equation, where $U_{\cdot k}$ is the eigenvector ϕ_k corresponding to the eigenvalue ε_k of H. We can then conclude that the eigenvectors ϕ_k are the columns of the matrix U that diagonalizes H, and the elements of the diagonalized matrix are the eigenvalues ε_k. This procedure requires numerical methods, and is widely used to solve Schrödinger's equations that, as said before, are not treatable analytically. In fact, this is what happens more frequently in physics and research activities.

Exercise 28 What is the expectation value of the energy in a system like the one considered in the previous exercise?

In the previous exercise, we have seen that in the absence of analytical solutions one can solve the eigenvalue equation with the help of a complete basis of functions φ_n, that is used to construct the matrix representation of the Hamiltonian, which is then diagonalized. The expected value of the energy was expressed in the form

$$\langle E \rangle = \sum_{i,j} c_i^* c_j H_{ij} = V_c^{\dagger} H V_c, \tag{7.176}$$

If U diagonalizes the matrix H, one can write the expected value as

$$\langle E\rangle = V_c^\dagger U\, U^\dagger\ H\ UU^\dagger\ V_c = V_c^\dagger U \varepsilon U^\dagger V_c. \tag{7.177}$$

This equation makes it clear that the matrix U, simultaneously, transforms the vector V_c into

$$W_a = U^\dagger V_c = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix}. \tag{7.178}$$

In terms of these coefficients and the eigenvalues ε_n the expected value in (7.176) becomes

$$\langle E\rangle = \sum_k |a_k|^2 \varepsilon_k. \tag{7.179}$$

This expression is similar to that in (7.45), for the expected value of the energy. For time independent Hamiltonians, we can write the time dependent wave function $\psi(x,t)$ as

$$\psi(x,t) = \sum_k a_k e^{-i\varepsilon_k t/\hbar} \phi_k(x). \tag{7.180}$$

If we evaluate the expected value of the energy in the time dependent state $\psi(x,t)$ we have

$$\langle E\rangle = \sum_{jk} a_j^* a_k e^{i(\varepsilon_j - \varepsilon_k)t/\hbar} \langle \phi_j|\widehat{H}|\phi_k\rangle. \tag{7.181}$$

Since $\langle \phi_j|\widehat{H}|\phi_k\rangle = \varepsilon_j \delta_{jk}$, we also have

$$\langle E\rangle = \sum_k |a_k|^2 \varepsilon_k. \tag{7.182}$$

7.9 Problems

1. Use the transmission amplitude of the rectangular barrier

$$t = \frac{1}{\cosh qb + i\dfrac{k^2 - q^2}{2qk}\sinh qb}, \tag{7.183}$$

to obtain the tunneling time as a function of the electron energy, when the barrier height and width are $V_o = 0.3\,\mathrm{eV}$ and $b = 20\,\mathrm{nm}$, respectively.

2. Deduce the wave function in (7.18).
3. Use the wave packet of Eq. (7.18) and show that, as time moves on, the wave packet gets dispersed, i.e. it becomes broaden. Show graphically this phenomenon.
4. Show that if $\widehat{f}$ and $\widehat{g}$ are two non-commuting Hermitian operators, the anti-commutator $\{\widehat{f}, \widehat{g}\} = \widehat{f}\widehat{g} + \widehat{g}\widehat{f}$ is Hermitian.
5. Show that if $[\widehat{f}, \widehat{g}] = i\hbar$, also $[\widehat{\delta f}, \widehat{\delta g}] = i\hbar$.
6. Show the identity

$$[\widehat{f}\widehat{g}, \widehat{h}] = \widehat{f}\,[\widehat{g}, \widehat{h}] + [\widehat{f}, \widehat{g}]\,\widehat{h}. \tag{7.184}$$

Using this identity and the commutator $[\widehat{x_i}, \widehat{p_i}]$, show that

$$[\widehat{x_i}^N, \widehat{h}] = i\hbar x_i^{N-1}. \tag{7.185}$$

7. **The virial theorem**. Use Eq. (7.121) and show that

$$\frac{d}{dt}\langle\widehat{x}\widehat{p}\rangle = 2\langle\widehat{T}\rangle - \left\langle\widehat{x}\frac{d\widehat{V}}{dx}\right\rangle \tag{7.186}$$

where $\widehat{T}$ is the operator of the kinetic energy, such that $\widehat{H} = \widehat{T} + \widehat{V}$. Show that in the stationary case

$$2\langle\widehat{T}\rangle = \left\langle\widehat{x}\frac{d\widehat{V}}{dx}\right\rangle. \tag{7.187}$$

This relation defines the virial theorem.

8. Using the virial theorem show that for stationary states

$$\langle\widehat{T}\rangle = \langle\widehat{V}\rangle. \tag{7.188}$$

9. Suppose that at $t = 0$ we have the Gaussian package

$$\varphi(x) = \langle x|\varphi\rangle = Ae^{-x^2/2a^2}. \tag{7.189}$$

Obtain:

   a. the normalization constant A, and the expectation values $\langle|x|\rangle$, $\langle|x^2|\rangle$;
   b. show that $\langle|p_x|\rangle = 0$, $\langle|p_x^2|\rangle = \hbar^2/2a^2$;
   c. obtain the standard deviations Δx and Δp_x. Obtain the product $\Delta x \Delta p_x$.

10. Suppose we have a free particle $\widehat{H} = \frac{\widehat{p}^2}{2m}$, that at time $t = 0$ is represented by the Gaussian packet $\langle x|\varphi\rangle = Ae^{-x^2/2a^2}$. Show that at time t:

a. the wave packet is given by

$$\begin{aligned}\langle x|\psi(t)\rangle &= \langle x|\int dp e^{-ip^2t/2m\hbar}|p\rangle\langle p|\varphi\rangle,\\ &= \langle x|\int\int dp dx' e^{-ip^2t/2m\hbar}|p\rangle\langle p|x'\rangle\langle x'|\varphi\rangle;\end{aligned} \tag{7.190}$$

b. replacing the functions $\langle x|p\rangle$, $\langle p|x'\rangle$ and the function $\langle x'|\varphi\rangle$, determine the wave packet at time t.

11. Show that

$$\langle\varphi_m|\frac{d^2x(t)}{dt^2}|\varphi_n\rangle = -(\omega_m-\omega_n)^2\langle\varphi_m|x(t)|\varphi_n\rangle. \tag{7.191}$$

12. Show that $\widehat{\delta f}\widehat{\delta g} = \frac{1}{2}\left[\widehat{\delta f}, \widehat{\delta g}\right] + \frac{1}{2}\left\{\widehat{\delta f}, \widehat{\delta g}\right\}$.

Chapter 8
Harmonic Oscillator

8.1 Introduction

The harmonic oscillator is one of the most invoked physical systems in the formulation of simple, but illustrative models of real systems, which otherwise require a more involved formalism. The harmonic oscillator has been used, for example, to model atoms in the blackbody walls and atomic vibrations in crystalline systems.

If one has a potential function $V(r)$ as, for example, the potential shown in Fig. 8.1, and one is interested only in the lowest energy levels, it is possible to approximate the potential, around the minima, by parabolic functions, as shown in the figure. In general, when the potential $V(\mathbf{r})$ has a complicated structure, the simplest and widely used approximation is the so-called harmonic approximation.

It is known from functional analysis that an analytic function $V(x)$ about a point x_0, can be developed as the Taylor power series

$$V(x) = V(x_0) + \frac{1}{1!}V'(x_0)(x - x_0) + \frac{1}{2!}V''(x_0)(x - x_0)^2 + \ldots \tag{8.1}$$

If $V(x)$ is the potential, with a minimum at x_0, the harmonic approximation in the vicinity of the minimum is:

$$V(x) = V(x_0) + \frac{1}{2}V''(x_0)(x - x_0)^2. \tag{8.2}$$

Since the physical description remains the same if one adds or subtracts a constant to $V(x)$, for example the constant $V(x_0)$, the potential function in the harmonic approximation is written just as

$$V(x) \simeq \frac{1}{2}V''(x_0)x^2. \tag{8.3}$$

P. Pereyra, *Fundamentals of Quantum Physics*, Undergraduate Lecture Notes in Physics,
DOI: 10.1007/978-3-642-29378-8_8,

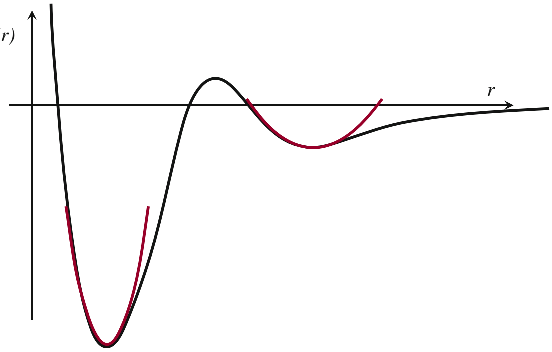


Fig. 8.1 Parabolic approximations in the potential minima

The coefficient $V''(x_0)$ contains very important physical information. The oscillation frequency ω is related with the curvature of the parabola. From now on we will write the harmonic potential in the form

$$V(x) = \frac{1}{2}mw^2x^2. \tag{8.4}$$

8.2 The Linear Harmonic Oscillator

The stationary Schrödinger equation of the harmonic oscillator, that will be solved in this section, is

$$-\frac{\hbar^2}{2m}\frac{d^2\varphi}{dx^2} + \frac{1}{2}mw^2x^2\varphi(x) = E\varphi(x). \tag{8.5}$$

The harmonic potential is a confining potential and our goal is to obtain the energy eigenvalues and the corresponding eigenfunctions. Before solving this problem we will simplify the notation. A convenient and widely used representation is the dimensionless one. This representation is easily achieved with the change of variable

$$x = \alpha_0\xi, \tag{8.6}$$

where ξ is a dimensionless variable. The coefficient α_0 has to be of unit length. With this change of variable the Schrödinger equation takes the form

$$-\frac{\hbar^2}{2m\alpha_0^2}\frac{d^2\varphi}{d\xi^2} + \frac{1}{2}mw^2\alpha_0^2\xi^2\varphi = E\varphi. \tag{8.7}$$

If we divide by the coefficient of the first term, we obtain

$$-\frac{d^2\varphi}{d\xi^2}+\frac{m^2\alpha_0^4 w^2}{\hbar^2}\xi^2\varphi=\frac{E2m\alpha_0^2}{\hbar^2}\varphi. \tag{8.8}$$

The first term is, obviously, dimensionless, all of the terms are also dimensionless. It is customary to choose α_0 such that

$$\frac{m^2\alpha_0^4 w^2}{\hbar^2}=1, \quad \Rightarrow \quad \alpha_0^2=\frac{\hbar}{mw}. \tag{8.9}$$

With this choice, the Schrödinger equation takes the form

$$-\frac{d^2\varphi}{d\xi^2}+\xi^2\varphi=\frac{2E}{\hbar w}\varphi. \tag{8.10}$$

To obtain the solution of this equation we will proceed as follows. We will first explore the equation and its solutions in the limit $|\xi|$ tending to ∞. Knowing the asymptotic solutions we will, in a second step, propose a solution for (8.10) that will, finally, lead us to a well known and soluble equation.

When $\xi \to \pm\infty$, it is clear that $2E/\hbar w \ll \xi^2$, thus, we can approximate (8.10) by

$$-\frac{d\varphi_\infty}{d\xi}+\xi^2\varphi_\infty=0. \tag{8.11}$$

To obtain φ_∞ we propose the function

$$\varphi_\infty(\xi)=e^{a\xi^2}, \tag{8.12}$$

and replace it in (8.11). So we have

$$(2a\xi)^2\, e^{a\xi^2}+2a\, e^{a\xi^2}-\xi^2\, e^{a\xi^2}=0. \tag{8.13}$$

In the asymptotic regions ($\xi \to \pm\infty$), the second term of this equation becomes negligible. We are then left with

$$(2a\xi)^2\, e^{a\xi^2}-\xi^2\, e^{a\xi^2}=0, \tag{8.14}$$

that is fulfilled when

$$a=\pm\frac{1}{2}. \tag{8.15}$$

Taking into account the finiteness requirement of $\varphi(\xi)$ that excludes $a=1/2$, we end up with the asymptotic solution

$$\varphi_\infty(\xi)=Be^{-\frac{1}{2}\xi^2}. \tag{8.16}$$

Since the harmonic oscillator solution φ must coincide, in the asymptotic region, with the exponential behavior of φ_∞, one can start with the trial function

$$\varphi(\xi) = e^{-\frac{1}{2}\xi^2} u(\xi), \tag{8.17}$$

with $u(\xi)$ a polynomial, to ensure the finiteness of the harmonic oscillator solution as ξ tends to ∞. If we define the dimensionless energy

$$\varepsilon = \frac{2E}{\hbar\omega}, \tag{8.18}$$

and replace the trial function $\varphi(\xi) = \varphi_\infty(\xi)u(\xi)$ in the differential Eq. (8.10), we have

$$\frac{d^2u}{d\xi^2} - 2\xi\frac{du}{d\xi} + (\varepsilon - 1)\, u = 0. \tag{8.19}$$

To solve this equation, it is common to propose the power series

$$u(\xi) = \sum_{j=0}^{\infty} a_j \xi^j, \tag{8.20}$$

that is then substituted in (8.19). This leaves us with

$$\sum_{j=2}^{\infty} a_j (j-1)\, j\, \xi^{j-2} - \sum_{j=0}^{\infty} a_j (2j + 1 - \varepsilon)\xi^j = 0. \tag{8.21}$$

We use this equation to determine the coefficients a_j. The normal procedure is to redefine the summation index j in order to factor ξ^j. In fact, if we change j by $j+2$, in the first sum, and factor, we have the combination

$$\sum_{j=0}^{\infty} \left[a_{j+2}(j+1)\,(j+2) - a_j\,(2j+1-\varepsilon)\right] \xi^j = 0, \tag{8.22}$$

that is fulfilled only if the coefficients of ξ^j vanish. This implies the recurrence relation

$$a_{j+2} = a_j \frac{2j + 1 - \varepsilon}{(j+1)(j+2)} \quad j = 0, 1, 2, \ldots, \tag{8.23}$$

that leads to two sets of coefficients a_j; those with odd indices, expressed in terms of the coefficient a_0, and those with even indices, expressed in terms of a_1. Therefore, the general solution has the form

$$\begin{aligned} u(\xi) = a_0 \Bigg(1 &+ \frac{1-\varepsilon}{2!}\,\xi^2 + \frac{(1-\varepsilon)(5-\varepsilon)}{4!}\,\xi^4 \\ &+ \ldots + \frac{(1-\varepsilon)\ldots(2n+1-\varepsilon)}{(n+2)!}\,\xi^{n+2} + \ldots \Bigg) \\ +a_1 \Bigg(\xi &+ \frac{3-\varepsilon}{3!}\,\xi^3 + \frac{(3-\varepsilon)(7-\varepsilon)}{5!}\,\xi^5 \\ &+ \ldots + \frac{(3-\varepsilon)\ldots(2n'+1-\varepsilon)}{(n'+2)!}\,\xi^{n'+2} + \ldots \Bigg) \end{aligned} \tag{8.24}$$

In each set of terms of the right hand side of this equation we write the general expression of the $n+1$ and the $n'+1$ terms. In the first series n is even ($n = 0, 2, 4, \ldots$) and in the second one n' is odd ($n' = 1, 3, 5, \ldots$). If the factors $(2n + 1 - \varepsilon)$ and $(2n' + 1 - \varepsilon)$ do not vanish, the series will have an infinite number of terms and will diverge as $\xi \to \infty$. Some lines above, we mentioned the finiteness requirement for φ, and also that it should be a polynomial function in order to keep the exponential behavior of φ_∞ in the asymptotic region. For $u(\xi)$ to become a polynomial, we need, first, that either $(2n + 1 - \varepsilon)$ or $(2n' + 1 - \varepsilon)$ vanish. If $(2n + 1 - \varepsilon)$ vanishes, the first series will terminate and become a polynomial, but not the second, in which case we have to choose $a_1 = 0$. If, on the other hand, $(2n' + 1 - \varepsilon)$ is the factor that vanishes, the second series terminates and we have to choose $a_0 = 0$. Therefore, the power series in (8.24) become a polynomial when

$$\varepsilon = 2n + 1, \qquad \text{which is equivalent to} \qquad \frac{2E}{\hbar\omega} = 2n + 1, \tag{8.25}$$

for n odd or even. Thus, for each n odd or even there is an energy

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right), \quad \text{with} \quad n = 0, 1, 2, 3, \ldots, \tag{8.26}$$

and the polynomials

$$u_n(\xi) = \sum_{j=0,2,4}^{n} a_{0j}\xi^j; \quad \text{when} \quad n = 0, 2, 4, \ldots, \tag{8.27}$$

and

$$u_n(\xi) = \sum_{j=1,3,5}^{n} a_{1j}\xi^j; \quad \text{when} \quad n = 1, 3, 5, \ldots, \tag{8.28}$$

with coefficients $a_{01} = a_0$,

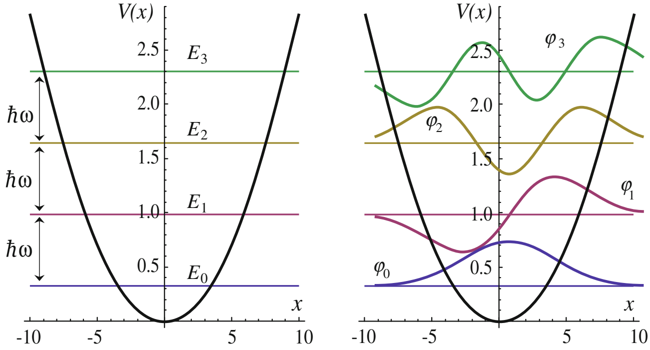


Fig. 8.2 Eigenvalues and eigenfunctions of the harmonic oscillator

$$a_{0j} = a_0(-2)^{j/2} \quad \frac{n(n-2)(n-4)\,\ldots\,(n+4-j)(n+2-j)}{j!} \qquad \mathrm{j} = 2, 4, 6, \ldots, \tag{8.29}$$

and $a_{11} = a_1$,

$$a_{1j} = a_1(-2)^{(j-1)/2}\, \frac{(n-1)(n-3)(n-5)\,\ldots\,(n+4-j)(n+2-j)}{j!} \\ \mathrm{j} = 3, 5, 7, \ldots, \tag{8.30}$$

respectively. The polynomials $u_n(\xi)$ in (8.27) and (8.28), normalized with a weight function $e^{-\xi^2}$, are essentially the Hermite polynomials $H_n(\xi)$. Consequently, the harmonic oscillator eigenfunctions are (remember that $\xi = x/\alpha_0 = x\sqrt{mw/\hbar}$)

$$\varphi_n(x) = C_n e^{-mwx^2/2\hbar} H_n\left(\sqrt{mw/\hbar}\; x\right). \tag{8.31}$$

The normalization constant can be found and is given by

$$C_n = \sqrt{\frac{1}{2^n n!}} \left(\frac{mw}{\pi\hbar}\right)^{1/4}. \tag{8.32}$$

In Fig. 8.2 we show the harmonic potential $m\omega^2 x^2/2$, with several eigenvalues E_n and their eigenfunctions φ_n. As for the infinite potential well, we have an infinite sequence of energy levels. The ground state corresponds to $n = 0$, the first excited state to $n = 1$ and so on. The eigenfunctions decrease exponentially for large $|x|$, and their parity corresponds with that of the quantum index n. One of the interesting features of the harmonic oscillator, perhaps the most interesting one, is the constant spacing of its energy levels. The difference between any two neighboring levels is $\hbar w$.

The Hermite polynomials, that are part of the eigenfunctions of the harmonic oscillator, are solutions of the differential equation

$$H_n'' - 2\xi H_n' + 2nH_n = 0. \tag{8.33}$$

This equation is easily obtained by combining Eqs. (8.19) and (8.25). Two useful relations that the Hermite polynomials fulfill are:

$$H_n'(x) = 2nH_{n-1}(x); \tag{8.34}$$

and the three term recurrence relation[1]

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x). \tag{8.35}$$

The student can deduce this recurrence relation by combining the differential Eqs. (8.33) and (8.34). Let us now introduce a small digression that will be useful for other applications.

8.3 Rising and Lowering Operators

The harmonic oscillator is a system with interesting properties like the already-mentioned constant spacing of its energy levels. This property makes the harmonic oscillator especially suitable for various theoretical developments. In particular, it allows one to define the rising and lowering operators (also known as creation and annihilation operators). With these operators, it is possible to represent the oscillator problem in a compact and very elegant mathematical notation and also to introduce an abstract notation, which becomes useful in applications or extensions of the quantum theory to other topics, especially to field theory. Though these applications are beyond the scope of this book, this is a good opportunity to introduce the rising and lowering operators.

If we consider the differential operators

$$\widehat{b} = \frac{1}{\sqrt{2}}\Big(\xi + \frac{\partial}{\partial\xi}\Big), \tag{8.36}$$

$$\widehat{b}^\dagger = \frac{1}{\sqrt{2}}\Big(\xi - \frac{\partial}{\partial\xi}\Big), \tag{8.37}$$

one can easily verify that

$$\widehat{b}^\dagger\widehat{b} = \frac{1}{2}\Big(\xi^2 - 1 - \frac{\partial^2}{\partial\xi^2}\Big), \tag{8.38}$$

$$\widehat{b}\,\widehat{b}^\dagger = \frac{1}{2}\Big(\xi^2 + 1 - \frac{\partial^2}{\partial\xi^2}\Big). \tag{8.39}$$

---

[1] Frequently orthogonal polynomials fulfill a three-term recurrence relation.

Combining properly these operators we can have others. For example, the commutator

$$\left[\widehat{b}^{\dagger},\widehat{b}\right]=\widehat{b}^{\dagger}\widehat{b}-\widehat{b}\,\widehat{b}^{\dagger}=-1, \tag{8.40}$$

and the anticommutator

$$\left\{\widehat{b}^{\dagger},\widehat{b}\right\}=\widehat{b}^{\dagger}\widehat{b}+\widehat{b}\,\widehat{b}^{\dagger}=\xi^2-\frac{\partial^2}{\partial\xi^2}. \tag{8.41}$$

The latter is exactly the dimensionless Hamiltonian in (8.10). This means that the Schrödinger equation of the harmonic oscillator can be written as

$$\frac{\hbar\omega}{2}\left[\widehat{b}^{\dagger}\widehat{b}+\widehat{b}\,\widehat{b}^{\dagger}\right]\varphi(\xi)=E\varphi(\xi), \tag{8.42}$$

or, better still, in the form

$$\hbar\omega\left(\widehat{b}^{\dagger}\widehat{b}+\frac{1}{2}\right)\varphi(\xi)=E\varphi(\xi). \tag{8.43}$$

Moreover, with the help of relations (8.33–8.35), for Hermite polynomials, it is possible to show that

$$\widehat{b}\,e^{-\frac{1}{2}\xi^2}H_n=\sqrt{2}\,n\,e^{-\frac{1}{2}\xi^2}H_{n-1} \tag{8.44}$$

$$\widehat{b}^{\dagger}e^{-\frac{1}{2}\xi^2}H_n=\frac{1}{\sqrt{2}}\,e^{-\frac{1}{2}\xi^2}H_{n+1}. \tag{8.45}$$

Therefore

$$\widehat{b}\,\varphi_n=\sqrt{n}\,\varphi_{n-1} \tag{8.46}$$

$$\widehat{b}^{\dagger}\,\varphi_n=\sqrt{n+1}\,\varphi_{n+1}. \tag{8.47}$$

In these equations, the operators $\widehat{b}^{\dagger}$ and $\widehat{b}$ behave, certainly, as *rising and lowering operators* in the space of the oscillator eigenfunctions. With the previous equations, it is easy to see that

$$\widehat{b}^{\dagger}\,\widehat{b}\,\varphi_n=n\,\varphi_n, \tag{8.48}$$

$$\widehat{b}\,\widehat{b}^{\dagger}\,\varphi_n=(n+1)\,\varphi_n. \tag{8.49}$$

The operator $\widehat{b}^{\dagger}\,\widehat{b}$ with eigenvalue n is also known as *the number operator*.

In (8.41), the Hamiltonian of the harmonic oscillator

$$\widehat{H}=\hbar\omega\left(\widehat{b}^{\dagger}\,\widehat{b}+\frac{1}{2}\right), \tag{8.50}$$

contains the number operator. Knowing that $\widehat{b}^\dagger \widehat{b}\, \varphi_n = n\, \varphi_n$, it is easy to verify that the eigenvalues of the harmonic oscillator are in fact

$$E_n = \hbar w \left(n + \frac{1}{2} \right). \tag{8.51}$$

8.4 Dipole Transitions and the Spontaneous Emission

The harmonic oscillator eigenfunctions have been used as an auxiliary basis in the solution of different types of quantum problems. In Chap. 7 we saw that using a basis of eigenfunctions, one can calculate matrix elements $H_{m,n} = \langle m|\widehat{H}|n\rangle$ and build up the matrix representation of $\widehat{H}$. We have seen that once the matrix H is built, one has to search for a matrix U, such that $U^\dagger H U = \varepsilon$ is diagonal. In that case, the matrix columns $U_{\cdot k}$ are the eigenfunctions ϕ_k of $\widehat{H}$ and the matrix elements ε_k the corresponding eigenvalues. If we have a time independent Hamiltonian, and we use these eigenfunctions, the most general solution can be written either as

$$\psi(x,t) = \sum_k a_k e^{-i\varepsilon_k t/\hbar} \phi_k(x) = W_a^T v \phi, \tag{8.52}$$

with $W_a^T = (a_1, a_2, \ldots)$, v a diagonal matrix with elements $e^{-i\varepsilon_k t/\hbar}$ and ϕ a vector which components are the eigenfunctions ϕ_k, or using the harmonic oscillator basis as

$$\psi(x) = \sum_n c_n e^{-i\omega_n t} \varphi_n(x) = V_c^T u \varphi, \tag{8.53}$$

with $V_c^T = W_a^T U^T$, u a diagonal matrix with elements $e^{-i\omega_n t}$ and $\varphi = U^* \phi$. Thus, if we are interested in the expected value of a time independent operator $\mathcal{O}$, in the quantum state $\psi(x,t)$, we have to evaluate

$$\langle\psi|\widehat{\mathcal{O}}|\psi\rangle = \sum_{j,k} a_j^* a_k e^{i(\varepsilon_j - \varepsilon_k)t/\hbar} \langle\phi_j|\widehat{\mathcal{O}}|\phi_k\rangle, \tag{8.54}$$

or, alternatively, the expression

$$\langle\psi|\widehat{\mathcal{O}}|\psi\rangle = \sum_{m,n} c_m^* c_n e^{i(\omega_m - \omega_n)t} \langle\varphi_m|\widehat{\mathcal{O}}|\varphi_n\rangle. \tag{8.55}$$

A simple and useful physical quantity is the electric dipole moment $e\mathbf{d}$. If we align the vector $\mathbf{d}$ along the x axis, we have just $e\mathbf{x}$. We will evaluate the matrix elements of this operator using the harmonic oscillator basis $\{\varphi_n(x)\}$. We will then calculate the spontaneous transition probabilities and the average lifetime of the excited oscillator states.

8.4.1 Selection Rules for Electric Dipole Transitions

Let us suppose that we are dealing with some quantum system with Hamiltonian $\widehat{H}$ and also that the corresponding Schrödinger equation is solved using the harmonic oscillator basis. We mentioned already that in this case we start writing the unknown solution as the superposition

$$\psi(x,t) = \sum_n c_n \psi_n(x,t), \tag{8.56}$$

with $\psi_n(x,t) = e^{-iE_nt/\hbar}\varphi_n(x)$, the eigenfunctions of a harmonic oscillator. As explained in chapter 7 and lines above, the coefficients c_n are fixed once we diagonalize the Hamiltonian matrix. We can then evaluate the expected value of ex using the eigenstates of $\widehat{H}$ or using the basis of the harmonic oscillator eigenfunctions. In the last case we have

$$\langle\psi|ex|\psi\rangle = \sum_{m,n} c_m^* c_n e^{i(\omega_m-\omega_n)t} e\langle\varphi_m|x|\varphi_n\rangle. \tag{8.57}$$

The matrix elements on the right hand side of this equation, will be denoted, indistinctly, as

$$x_{m,n} = \langle m|x|n\rangle = \langle\varphi_m|x|\varphi_n\rangle. \tag{8.58}$$

To evaluate these matrix elements we use the properties of the Hermite polynomials. If we write the integrand in the form

$$x_{m,n} = \int \varphi_m^* x \varphi_n dx = \alpha_0^2 C_m \int e^{-\frac{1}{2}\xi^2} H_m\, \xi\, \varphi_n d\xi. \tag{8.59}$$

and use the recurrence equation

$$\xi H_n = nH_{n-1} + \frac{1}{2}H_{n+1}, \tag{8.60}$$

we have

$$x_{m,n} = \alpha_0 C_m \left[\frac{m}{C_{m-1}}\int \varphi_{m-1}^*\varphi_n dx + \frac{1}{2C_{m+1}}\int \varphi_{m+1}\varphi_n dx\right]. \tag{8.61}$$

This equation shows that the matrix elements are zero except for $n = m - 1$ and $n = m + 1$. This means that:

$$x_{m,n} = \langle m|x|n\rangle = \begin{cases} \dfrac{m\,C_m}{C_{m-1}}\,\delta_{m-1,n} = \sqrt{\dfrac{n+1}{2}}\,\delta_{m,n+1} \\ \\ \dfrac{C_m}{2\,C_{m+1}}\,\delta_{m+1,n} = \sqrt{\dfrac{n}{2}}\,\delta_{m,n-1}. \end{cases} \tag{8.62}$$

The only matrix elements $x_{m,n}$ different from zero are those for which

$$m = n \pm 1. \tag{8.63}$$

These are the selection rules of the electric dipole moment in the harmonic oscillator basis. Since the matrix elements $\langle\varphi_m|x|\varphi_n\rangle$ involve two eigenstates. It is clear that the electric dipole induces transitions between those states of the basis, between $|\varphi_n\rangle$ and $\langle\varphi_m|$. The actual system may eventually pass from one eigenstate of the system to another. These are not the transitions we are talking about. The matrix elements $x_{m,n} = \langle\varphi_m|x|\varphi_n\rangle$ imply transitions in the space of the harmonic oscillator basis. The expected dipole moment $\langle\psi|e\,x|\psi\rangle$ is a superposition of these transitions. Each transition $\langle\varphi_m|x|\varphi_n\rangle$ is accompanied by a photon emission or a photon absorption with frequency $\omega_{mn} = \omega_n - \omega_m = (m-n)\omega$. Notice that due to the selection rule $\omega_{mn} = \pm\omega$. The expected value $\langle\varphi_m|x|\varphi_n\rangle$ will then be written as

$$\langle\varphi_m|x|\varphi_n\rangle = \begin{cases} \sqrt{\dfrac{n+1}{2}}\,e^{i\omega t}\,\delta_{m,n+1}, \\ \\ \sqrt{\dfrac{n}{2}}\,e^{-i\omega t}\,\delta_{m,n-1}. \end{cases} \tag{8.64}$$

One can easily verify that the matrix that represents the operator $\widehat{x}$ in the harmonic oscillator basis is self-adjoint with zeros everywhere except in the diagonals that are first neighbors of the main diagonal.

Taking into account the expected values $\langle\varphi_m|x|\varphi_n\rangle$ and the selection rules, one can finally write the expected value in (8.57) as

$$\langle\psi|e\,x|\psi\rangle = \frac{1}{\sqrt{2}}\sum_n e\left(c^*_{n+1}c_n\,\sqrt{n+1}\,e^{i\omega t} + c^*_{n-1}c_n\,\sqrt{n}\,e^{-i\omega,t}\right), \tag{8.65}$$

which can also be written as

$$\langle\psi|e\,x|\psi\rangle = e\sum_n\sqrt{\frac{n}{2}}\,\Re e\left[c^*_{n-1}c_n e^{i\omega t}\right], \tag{8.66}$$

or in the form

$$\langle\psi|e\,x|\psi\rangle = e\sum_n\sqrt{\frac{n}{2}}\,\left[\Re e\,[c_{n-1}^*c_n]\cos\omega t - \Im m\,[c_{n-1}^*c_n]\sin\omega t\right]. \tag{8.67}$$

According to this expression, the expected value of ex, and of course of x, oscillates with the characteristic frequency of the harmonic oscillator, and with an amplitude that depends also on the coefficients c_n of the expansion of $|\psi\rangle$ in the harmonic oscillator basis. Recall that these coefficients depend on the Hamiltonian $\widehat{H}$.

8.4.2 *Lifetime of Excited States*

In Chap. 1 we used the Einstein absorption and emission model to derive the Planck distribution. For that purpose, the absorption B_n^m, emission B_m^n and spontaneous emission A_m^n probabilities were introduced, for transitions between the energy levels E_n and E_m. Rigorous calculations of these coefficients can be done, without much difficulty, in the quantum field theory. We present here some intermediate results that may give an idea of the way in which the absorption and emission probabilities can be calculated. In the relativistic quantum field theory, the interaction of light and matter is described by the exchange of photons. The photonic fields are described by using vectors like

$$|\phi\rangle = |n_{\lambda_1}, n_{\lambda_2}, \ldots, n_{\lambda_j}, \ldots\rangle, \tag{8.68}$$

where n_{λ_i} represents the number of photons in the propagating mode[2] $\lambda_i = \mathbf{k}_i\gamma$ and *the electromagnetic field energy* with the Hamiltonian

$$\mathcal{H} = \sum_\lambda \hbar\omega_\lambda\left(\widehat{a}_\lambda^\dagger\widehat{a}_\lambda + \frac{1}{2}\right), \tag{8.69}$$

where $\widehat{a}_\lambda^\dagger$ and $\widehat{a}_\lambda$ are the rising and lowering operators. In this case, we call them creation and annihilation operators. These operators acting on the states $|\phi\rangle$ have the same effects as the rising and lowering operators on the oscillator eigenfunctions, i.e.

$$\widehat{a}_{\lambda_j}|\phi\rangle = \sqrt{n_{\lambda_j}}\;|n_{\lambda_1}, n_{\lambda_2}, \ldots, n_{\lambda_j}-1, \ldots\rangle, \tag{8.70}$$

$$\widehat{a}_{\lambda_j}^\dagger|\phi\rangle = \sqrt{n_{\lambda_j}+1}\;|n_{\lambda_1}, n_{\lambda_2}, \ldots, n_{\lambda_j}+1, \ldots\rangle. \tag{8.71}$$

We recall that when the number operator $\widehat{a}_\lambda^\dagger\widehat{a}_\lambda$, multiplied by $\hbar\omega_\lambda$ acts on the state $|n_\lambda\rangle$ with occupation number n_λ, the result is again the state $|n_\lambda\rangle$ times the energy $n_\lambda\hbar\omega_\lambda$.

To describe the electron-field system it is common to consider the Hamiltonian

---

[2] $\mathbf{k}_i$ refers to the wave vector and γ to the polarization.

$$\widehat{H} = \sum_j \frac{1}{2m}\mathbf{p}_j^2 + V(\mathbf{r}_1, \mathbf{r}_2, \ldots \mathbf{r}_N) + \sum_\lambda \hbar\omega_\lambda \left(\widehat{a}_\lambda^\dagger \widehat{a}_\lambda + \frac{1}{2}\right) + \widehat{H}_I. \tag{8.72}$$

where $\widehat{H}_I$ is the interaction Hamiltonian, between the charge carriers and the electromagnetic field, usually expressed in the form

$$\widehat{H}_I = \sum_j \frac{e}{2m} 2\mathbf{p}_j \cdot \mathbf{A}(\mathbf{r}_j) + \frac{e^2}{2m}\sum_j \mathbf{A}(\mathbf{r}_j)^2, \tag{8.73}$$

with **A** the potential vector that can be written as

$$\mathbf{A}(\mathbf{r}, t) = \frac{1}{\sqrt{V}} \sum_{\mathbf{k}} \sum_{\gamma=1}^{2} \mathbf{e}_{\mathbf{k}\gamma} \sqrt{\frac{h}{\omega_{\mathbf{k}\gamma}}} \left[\widehat{a}_\lambda e^{i\mathbf{k}\cdot\mathbf{r}} + \widehat{a}_\lambda^\dagger e^{-i\mathbf{k}\cdot\mathbf{r}}\right]. \tag{8.74}$$

Here $\mathbf{e}_{\mathbf{k}\gamma}$ are unit vectors of the polarizations. The transition probability from the initial state $|i\rangle = |\psi_i\rangle|\phi\rangle$ to the final state $|f\rangle = |\psi_f\rangle|\phi\rangle$ corresponds to an emission process if $E_f < E_i$. In this case we have to evaluate the expression

$$W_{fi}^{(a)} = A_i^f = \frac{4\pi^2 e^2}{m^2 \hbar V} \sum_\lambda \frac{n_\lambda + 1}{\omega_\lambda} \, |\langle \psi_f | \mathbf{e}_\lambda e^{-i\mathbf{k}\cdot\mathbf{r}} \cdot \mathbf{p}_j | \psi_i \rangle|^2 \, \delta(\omega_{fi} - \omega_\lambda). \tag{8.75}$$

This can be evaluated within the dipole approximation, when the wave number $k = 2\pi/\lambda$ of the radiation field is very small[3] and the exponential $exp(i\mathbf{k}\cdot\mathbf{r}) \simeq 1 + i\mathbf{k}\cdot\mathbf{r}$ is almost one. After a more or less straightforward calculation that can be found in specialized texts,[4] the above expression reduces to

$$W_{fi}^{(a)} = A_i^f = \frac{4e^2\omega^3}{3\hbar c^3} \, |\langle \varphi_f | \, \mathbf{r} \, | \varphi_i \rangle|^2, \tag{8.76}$$

which, in the one-dimensional approximation, with $n_i = n_f + 1$, is

$$W_{fi}^{(a)} = \frac{4e^2\omega^3}{3\hbar c^3} \frac{n_i}{2}. \tag{8.77}$$

This result tells us that the probability for the spontaneous emission per unit time is greater at higher energy levels. The inverse of the emission probability is *the mean lifetime* τ_i in the state $|\psi_i\rangle$, i.e.:

---

[3] This corresponds to radiation-field wavelengths much larger than the system size.

[4] See for example the excellent book by Yehuda B. Band, *Light and Matter*, John Wiley & Sons, West Sussex, England, 2006.

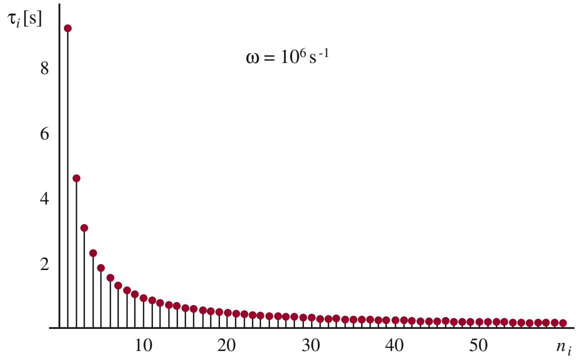


Fig. 8.3 The mean lifetime of an electron, in the excited states of the harmonic potential, as a function of the excitation level n

$$\tau_i = \frac{1}{W_{fi}^{(a)}} = \frac{1}{A_i^f} = \frac{3\hbar c^3}{2e^2\omega^3}\frac{1}{n_i}. \tag{8.78}$$

which, as can be seen in Fig. 8.3, drops very rapidly as the excitation level increases.

8.5 Solved Problems

Exercise 29 To evaluate the eigenvalues E_n and plot the eigenfunctions φ_n of the harmonic oscillator in Fig. 8.2, we consider a particle with mass equal to that of the electron moving with a frequency $\omega = 10^{15}\,\mathrm{s}^{-1}$. Determine the energies in eV and the classical turning points in Å, for the first four energy levels. What is the energy-levels spacing?

Solution The classical turning points are those where the kinetic energy becomes zero, therefore

$$E_n = \frac{1}{2}kx_r^2. \tag{8.79}$$

If we use the oscillator energy as a function of n, the coordinate x_r of the turning points will be given by

$$x_r = \pm\sqrt{\frac{2\hbar\omega}{k}\left(n+\frac{1}{2}\right)}. \tag{8.80}$$

Since $k = m\omega^2$, the last equation takes the form

$$x_r = \pm\sqrt{\frac{2\hbar}{m\omega}\left(n+\frac{1}{2}\right)}. \tag{8.81}$$

In the following table we have the energy and turning points for the first four values of n

n	E_n [eV]	x_r [Å]
0	0.329106	± 3.40249
1	0.987318	± 5.89329
2	1.64553	± 7.60821
3	2.30374	± 9.00215

The student can verify these numbers and prove that they correspond to the energies and turning points in Fig. 8.2. Beyond the turning points the eigenfunctions fall exponentially.

The spacing between energy levels is

$$\hbar\omega = 0.65821\,[\text{eV}] \tag{8.82}$$

Exercise 30 Show that: i) the expectation values $\langle x\rangle$ and $\langle\widehat{p}_x\rangle$, in the harmonic oscillator state φ_n, vanish; ii) the variances of x and p_x satisfy the Heisenberg inequality.
Solution We shall start evaluating the expectation values

$$\langle x\rangle = \langle n|x|n\rangle = \int \varphi_n^*(x)\, x\, \varphi_n(x) dx \tag{8.83}$$

$$\langle\widehat{p}_x\rangle = \langle n|\widehat{p}_x|n\rangle = -i\hbar \int \varphi_n^*(x)\, \frac{\partial}{\partial x}\, \varphi_n(x) dx. \tag{8.84}$$

To evaluate these integrals we will use some relations like

$$\begin{aligned} H_n'(\xi) &= 2n H_{n-1}(\xi), \\ \xi H_n(\xi) &= n H_{n-1}(\xi) + \frac{1}{2} H_{n+1}(\xi), \end{aligned}$$

as well as the orthonormality condition

$$\langle n'|n\rangle = \int \varphi_{n'}^*(\xi)\, \varphi_n(\xi) d\xi = \delta_{n'n}. \tag{8.85}$$

We have seen, in Sect. 8.4, that the only matrix elements $\langle n'|x|n\rangle$ different from zero are those with $n' = n \pm 1$, therefore

$$\langle n|x|n\rangle = 0. \tag{8.86}$$

To obtain the expected value $\langle n|\widehat{p}_x|n\rangle$ it is convenient to express the integrand as follows

$$\langle n|\widehat{p}_x|n\rangle = -i\hbar C_n \int \varphi_n^*(\xi)\, \frac{\partial}{\partial\xi}\, e^{-\xi^2/2} H_n(\xi) d\xi. \tag{8.87}$$

The operator $\partial/\partial\xi$ acting on the exponential function, multiplied by the Hermite polynomial, leaves us with

$$\langle n|\widehat{p}_x|n\rangle = -i\hbar C_n \int \varphi_n^*(\xi)\, e^{-\xi^2/2}\left(-\xi H_n(\xi) + H_n'(\xi)\right) d\xi. \tag{8.88}$$

With the help of the Hermite polynomial relations, and their relation with the eigenfunctions φ_n, we have

$$\langle n|\widehat{p}_x|n\rangle = -i\hbar C_n \int \varphi_n^*(\xi)\left(n\frac{\varphi_{n-1}(\xi)}{C_{n-1}} - \frac{\varphi_{n+1}(\xi)}{2C_{n+1}}\right) d\xi. \tag{8.89}$$

This certainly gives us the suggested result

$$\langle n|\widehat{p}_x|n\rangle = -i\hbar\left(\frac{nC_n}{C_{n-1}}\langle n|n-1\rangle - \frac{C_n}{2C_{n+1}}\langle n|n+1\rangle\right) = 0. \tag{8.90}$$

To obtain the variances we need to evaluate the expected values of x^2 and of $\widehat{p}_x^2$. We start with

$$\langle x^2\rangle = \langle n|x^2|n\rangle = \alpha_0^3 \int \varphi_n^*(\xi)\,\xi^2\,\varphi_n(\xi)d\xi, \tag{8.91}$$

where we have used the relation $x = \alpha_0\xi$. Taking into account the recurrence relations of $H_n(\xi)$, this equation transforms into

$$\begin{aligned}\langle x^2\rangle = \alpha_0^3 C_n &\int \varphi_n^*(\xi) e^{-\xi^2/2}\\ &\times\left(n(n-1)H_{n-2}(\xi) + \left(n+\frac{1}{2}\right)H_n(\xi) + \frac{1}{4}H_{n+2}(\xi)\right) d\xi.\end{aligned} \tag{8.92}$$

In this integral we can use the orthonormality of the eigenfunctions $\varphi_n(x)$. It is clear that only the third term remains, giving us

$$\langle x^2\rangle = \alpha_0^2\left(n+\frac{1}{2}\right). \tag{8.93}$$

Therefore, the variance of x is:

$$(\varDelta x)^2 = \langle x^2\rangle - \langle x\rangle^2 = \alpha_0^2\left(n+\frac{1}{2}\right). \tag{8.94}$$

We shall now calculate the expected value of $\widehat{p}_x^2$. In this case we have

$$\begin{aligned}\langle \widehat{p}_x^{\,2} \rangle = \langle n | \widehat{p}_x^{\,2} | n \rangle &= \int \varphi_n^*(x)\, \widehat{p}_x^{\,2}\, \varphi_n(x) dx \\ &= -\frac{\hbar^2}{\alpha_0} \int \varphi_n^*(\xi)\, \frac{\partial^2}{\partial \xi^2}\, e^{-\xi^2/2} H_n(\xi) d\xi. \end{aligned} \tag{8.95}$$

Using the Hermite polynomials relations and after a little algebra, it can be shown that

$$\frac{\partial^2}{\partial \xi^2}\, e^{-\xi^2/2} H_n(\xi) = e^{-\xi^2/2} \left[n(n-1) H_{n-2}(\xi) - \left(n + \frac{1}{2} \right) H_n(\xi) + \frac{1}{4} H_{n+2}(\xi) \right]. \tag{8.96}$$

If we substitute in the Eq. (8.95), it becomes clear that only the term proportional to H_n contributes to the integral. We then have

$$\langle \widehat{p}_x^{\,2} \rangle = \frac{\hbar^2}{\alpha_0^2} \left(n + \frac{1}{2} \right), \tag{8.97}$$

and the variance is

$$(\Delta \widehat{p}_x)^2 = \langle \widehat{p}_x^{\,2} \rangle - \langle \widehat{p}_x \rangle^2 = \frac{\hbar^2}{\alpha_0^2} \left(n + \frac{1}{2} \right). \tag{8.98}$$

With these results we have the equation

$$(\Delta x)^2 (\Delta \widehat{p}_x)^2 = \hbar^2 \left(n + \frac{1}{2} \right)^2, \tag{8.99}$$

that obviously satisfies the Heisenberg inequality

$$(\Delta x)^2 (\Delta \widehat{p}_x)^2 \geq \frac{\hbar^2}{4}. \tag{8.100}$$

Equation (8.99) shows that the state of minimum dispersion is the ground state.

Exercise 31 Show that the coefficients a_{0j} and a_{1j} in Eqs. (8.27) and (8.28) are given by the relations

$$a_{0j} = a_0\, (-2)^{j/2}\, \frac{n(n-2)(n-4)\, \ldots\, (n+4-j)(n+2-j)}{j!} \qquad \text{j=2, 4, 6, \ldots,} \tag{8.101}$$

$$a_{1j} = a_1 (-2)^{(j-1)/2}\, \frac{(n-1)(n-3)(n-5)\, \ldots\, (n+4-j)(n+2-j)}{j!} \qquad \text{j=3, 5, 7, \ldots} \tag{8.102}$$

Write explicitly the polynomials $u_n(\xi)$ for n even and odd, until the fourth term. Determine the polynomials $u_0(\xi)$, $u_1(\xi)$,...and $u_4(\xi)$. Compare them with the corresponding Hermite polynomials and check that $H_n(\xi) = f_n\, u_n(\xi)$, where f_n is only a numerical factor.

Solution If in the recurrence relation (8.23) we change $j+2$ by j and instead of ε we use ε_n, we have

$$a_j = a_{j-2}\frac{2j-3-\varepsilon_n}{(j-1)j} \qquad \mathrm{j} = 2, 3, 4, \ldots, \tag{8.103}$$

the even-index coefficients are of the form

$$a_{0j} = a_0\ \frac{(1-\varepsilon_n)(5-\varepsilon_n)\ \ldots\ (2j-7-\varepsilon_n)(2j-3-\varepsilon_n)}{j!} \qquad \mathrm{j} = 2, 4, 6, \ldots, \tag{8.104}$$

whereas those with odd-index are

$$a_{1j} = a_1\ \frac{(3-\varepsilon_n)(7-\varepsilon_n)\ \ldots\ (2j-7-\varepsilon_n)(2j-3-\varepsilon_n)}{j!} \qquad \mathrm{j} = 3, 5, 7, \ldots \tag{8.105}$$

Now recall that the condition to terminate the infinite series is $\varepsilon_n = 2n+1$, with n even or odd. If we replace ε_n by $2n+1$ for example, in the even-indices relation, we have

$$a_{0j} = a_0\ \frac{(-2n)(4-2n)\ \ldots\ (2j-8-2n)(2j-4-2n)}{j!} \qquad \mathrm{j} = 2, 4, 6, \ldots \tag{8.106}$$

Here it is possible to factor -2 in the $j/2$ factors of the numerator. In this way we obtain

$$a_{0j}{=}a_0 \quad (-2)^{j/2} \quad \frac{n(n{-}2)(n{-}4)\ \ldots\ (n{+}4{-}j)(n{+}2{-}j)}{j!} \qquad \mathrm{j}{=}2, 4, 6, \ldots \tag{8.107}$$

Similarly we get the expression for a_{1j}.

The polynomial $u_n(\xi)$ for n even, where the first four terms are explicitly written is then

$$\begin{aligned} u_n(\xi) &= a_{00} + a_{02}\xi^2 + a_{04}\xi^4 + a_{06}\xi^6 + \ldots \\ &= a_0\left(1 - \frac{2n}{2!}\xi^2 + \frac{2^2\, n(n-2)}{4!}\xi^4 - \frac{2^3\, n(n-2)(n-4)}{6!}\xi^6 + \ldots\right). \end{aligned} \tag{8.108}$$

The polynomial $u_n(\xi)$ for n odd, is

$$\begin{aligned} u_n(\xi) &= a_{11}\xi + a_{13}\xi^3 + a_{15}\xi^5 + a_{17}\xi^7 + \dots \\ &= a_1\left(\xi - \frac{2(n-1)}{3!}\xi^3 + \frac{2^2(n-1)(n-3)}{5!}\xi^5 \right. \\ &\quad \left. - \frac{2^3(n-1)(n-3)(n-5)}{7!}\xi^7 + \dots\right). \end{aligned} \tag{8.109}$$

If we give values to n and set the coefficients $a_0 = a_1 = 1$, we have the functions $u_0(\xi),\dots,u_4(\xi)$ shown below:

$$\begin{array}{ll} u_0(\xi) = 1; & H_0(\xi) = 1; \\ u_1(\xi) = \xi; & H_1(\xi) = 2\xi; \\ u_2(\xi) = -(2\xi^2 - 1); & H_2(\xi) = 4\xi^2 - 2; \\ u_3(\xi) = -(4\xi^3 - 6\xi)/3!; & H_3(\xi) = 8\xi^3 - 12\xi; \\ u_4(\xi) = (8\xi^4 - 24\xi^2 + 6)/3!; & H_4(\xi) = 16\xi^4 - 48\xi^2 + 12. \end{array}$$

It is clear that $u_n(\xi)$ and $H_n(\xi)$ differ only by a numerical factor, and in some cases also by a global sign.

8.6 Problems

1. Show that the solution to the Schrödinger equation in the asymptotic region is given by (8.16).
2. Check the expressions in Eqs. (8.38) and (8.39) and show that the harmonic oscillator Hamiltonian can be expressed as

$$\widehat{H} = \hbar\,\omega\,\left(\widehat{b}^{\dagger}\widehat{b} + \frac{1}{2}\right). \tag{8.110}$$

3. Verify Eqs. (8.44–8.47).
4. Determine the recurrence relation of the Hermite polynomials.
5. Prove that the normalization constant of the harmonic oscillator eigenfunctions is precisely the one given in (8.32).
6. Show that the Schrödinger equation

$$-\frac{d^2\varphi}{d\xi^2} + \xi^2\varphi = \frac{2E}{\hbar w}\varphi, \tag{8.111}$$

can be written as

$$(\widehat{b}^{\dagger}\widehat{b} + 1)\varphi = \frac{2E}{\hbar w}\varphi. \tag{8.112}$$

7. Prove that $[\widehat{b}, \widehat{b}^{\dagger}] = 1$.

8. The expected value of the energy of an oscillator is

$$\langle E \rangle = \int \varphi^*(x) \left\{ \frac{1}{2m}\widehat{p}^{\,2} + \frac{1}{2}mw^2\widehat{x}^{\,2} \right\} \varphi(x)dx = \frac{1}{2m}\langle \widehat{p}^{\,2} \rangle + \frac{1}{2}mw^2\langle \widehat{x}^{\,2} \rangle. \tag{8.113}$$

Using the definition of the variance

$$var F = \langle (\widehat{F} - \langle \widehat{F} \rangle)^2 \rangle = \langle (\delta \widehat{F})^2 \rangle = (\Delta F)^2, \tag{8.114}$$

for the operators $\widehat{p}$ y $\widehat{x}$ and the Heisenberg inequality $(\Delta p)^2(\Delta x)^2 \geq \frac{1}{4}\hbar^2$, show that whenever $\langle \widehat{x} \rangle = 0$ and $\langle \widehat{p} \rangle = 0$, the minimum energy $\langle E \rangle$ is $\hbar w/2$

9. Prove that between the time dependent matrices u and v, defined in (8.52) and (8.53), we have the following relation

$$v = U^{\dagger}uU, \tag{8.115}$$

with U the same matrix that diagonalizes the Hamiltonian H in the basis $\{\varphi_n\}$.

10. Consider a particle in a harmonic potential. Plot the probability density $|\varphi_n(x)|^2$ for n =5, 10 and 15 and compare these quantities with the classical probability density defined as

$$p(x)dx = \frac{2}{T}\frac{dx}{v(x)}, \tag{8.116}$$

where T is the classical period and $v(x)$ the classical velocity. Discuss the relation of these results with the correspondence principle.

11. Figure 8.3 shows an electron half-life time (lifetime) τ_n in the excited states of the harmonic oscillator, as a function of the excitation level n. When the excitation level is too large, the average lifetime approaches zero, i.e.:

$$\text{if } n \to n_\infty \cong \infty \;\Rightarrow\; \tau_{n_\infty} \to 0.$$

Show that if an electron is in an energy level E_{n_∞} (with $n_\infty \to \infty$), then it will take an infinite time to come, by spontaneous decay, to the ground state. Show that this time is proportional to the harmonic series, i.e., show that:

$$\tau = \frac{3\hbar c^3}{2e^2\omega^3}\left(1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots + \frac{1}{n_\infty}\right) \to \infty\,! \tag{8.117}$$

Chapter 9
Angular Momentum and Central Potentials

9.1 Introduction

Rotation is a fundamental motion in the real world. Thus an important and an appropriate number of physical variables have been introduced to study this type of motion, both in the classical and the quantum description. The relevant physical quantity is the angular momentum. In the quantum description, the angular momentum is defined exactly as in the classical physics, but with the momentum written as an operator, i.e. as the product

$$\widehat{\mathbf{L}} = \widehat{\mathbf{r}} \times \widehat{\mathbf{p}} = -i\hbar \mathbf{r} \times \nabla. \tag{9.1}$$

In this chapter we will study some properties of the operator $\widehat{\mathbf{L}}$. We will derive the angular-momentum conservation law in the absence of external torques. We will see also that, when particles move in the presence of central potentials, the product $\widehat{L}^2 = \widehat{\mathbf{L}} \cdot \widehat{\mathbf{L}}$ plays a fundamental role. Finally, having in mind the Hydrogen atom problem, that we will study in the next chapter, we will solve the angular-momentum eigenvalue problem.

9.2 Angular Momentum and Their Commutation Relations

The vectors $\mathbf{r}$ and ∇ can be expressed in the coordinates system that suits best or facilitates more the mathematical treatment, usually it is chosen in accord with the physical system symmetries. In the cartesian coordinates system we have

$$\widehat{\mathbf{L}} = \widehat{L}_x \mathbf{i} + \widehat{L}_y \mathbf{j} + \widehat{L}_z \mathbf{k}, \tag{9.2}$$

where

$$\widehat{L}_x = -i\hbar \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y} \right); \quad \widehat{L}_y = -i\hbar \left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z} \right) \tag{9.3}$$

P. Pereyra, *Fundamentals of Quantum Physics*, Undergraduate Lecture Notes in Physics, DOI: 10.1007/978-3-642-29378-8_9, © Springer-Verlag Berlin Heidelberg 2012

and

$$\widehat{L}_z = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right). \tag{9.4}$$

In compact notation, the previous expressions are written in the form (with summations over repeated indices)

$$\widehat{L}_i = -i\hbar\,\epsilon_{ijk}\,x_j\frac{\partial}{\partial x_k} \quad \text{with} \quad x_1 = x, \quad x_2 = y \quad \text{and} \quad x_3 = z \tag{9.5}$$

The symbol ϵ_{ijk} is called the Levi-Civita tensor which is equal to 1 if ijk is an even permutation of 123, -1 if it is an odd permutation and 0 if any index is repeated.

The most common realization of a rotating motion in the microscopic world is the motion of bounded electrons around the atomic nuclei. The symmetries of the potential suggests that the spherical coordinates are appropriate for studying this class of systems. Before discussing the commutation properties of $\widehat{\mathbf{L}}$ and its components, let us see how these operators look like in the spherical coordinates system. We know that

$$\mathbf{r} = r\mathbf{u}_r \quad \text{and} \quad \nabla = \mathbf{u}_r\frac{\partial}{\partial r} + \mathbf{u}_\theta\frac{1}{r}\frac{\partial}{\partial\theta} + \mathbf{u}_\phi\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}. \tag{9.6}$$

Therefore

$$\begin{aligned}\widehat{\mathbf{L}} &= -i\hbar\, r\,\mathbf{u}_r \times \left(\mathbf{u}_r\frac{\partial}{\partial r} + \mathbf{u}_\theta\frac{1}{r}\frac{\partial}{\partial\theta} + \mathbf{u}_\phi\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\right)\\ &= i\hbar\left(\mathbf{u}_\theta\frac{1}{\sin\theta}\frac{\partial}{\partial\phi} - \mathbf{u}_\phi\frac{\partial}{\partial\theta}\right).\end{aligned} \tag{9.7}$$

The angular momentum operator $\widehat{\mathbf{L}}$ depends only on the angular variables ϕ and θ. For this reason, $\widehat{\mathbf{L}}$ commutes with any function of r and $\partial/\partial r$. We will return to this point below. Given the relations

$$\begin{aligned}\mathbf{u}_r &= \mathbf{i}\sin\theta\cos\phi + \mathbf{j}\sin\theta\sin\phi + \mathbf{k}\cos\theta\\ \mathbf{u}_\theta &= \mathbf{i}\cos\theta\cos\phi + \mathbf{j}\cos\theta\sin\phi - \mathbf{k}\sin\theta\\ \mathbf{u}_\phi &= -\mathbf{i}\sin\phi + \mathbf{j}\cos\phi\end{aligned} \tag{9.8}$$

one can easily check that the $\widehat{L}_z$ component, defined as the scalar product of $\widehat{\mathbf{L}}$ with $\widehat{\mathbf{k}}$ (where $\widehat{\mathbf{k}}$ is a unit vector along the z axis), takes the simple form

$$\widehat{L}_z = -i\hbar\frac{\partial}{\partial\phi}. \tag{9.9}$$

We have seen earlier that the nature of the eigenvalues and the possibility of common eigenfunctions with other operators depend on the commutation relations. This is a particularly important issue in central potential problems. We will present now some commutation relations among the angular momentum components $\widehat{L}_i$ and also with the position operators $\widehat{x}_i$ and with the linear momentum operators $\widehat{p}_i$.

9.2.1 Commutation Relations Between $\widehat{\mathbf{L}}_i$ and $\widehat{\mathbf{x}}_j$

We will explicitly calculate a couple of commutators. It is important to recall that in order to determine a commutator correctly, one proceeds as if on the right side of the commutator one has a function, the function on which the commutator would eventually operate. In our first example, this precaution will not have effect because the operators that we have there commute, but in the second example it will be important to take into account this rule. Let us begin with the commutator $[\widehat{L}_x, \widehat{x}]$, which we develop as follows:

$$\begin{aligned}\left[\widehat{L}_x, \widehat{x}\right] &= \widehat{L}_x\widehat{x} - \widehat{x}\widehat{L}_x = -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) x + i\hbar x \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)\\ &= 0 - i\hbar x \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) + i\hbar x \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)\\ &= 0. \end{aligned} \tag{9.10}$$

From this relation we conclude that the operators $\widehat{L}_i$ commute with the operators $\widehat{x}_i$ of the same index. Let us see now a slightly different case: the commutator $[\widehat{L}_x, \widehat{y}]$, i.e:

$$\begin{aligned}\left[\widehat{L}_x, \widehat{y}\right] = \widehat{L}_x\widehat{y} - \widehat{y}\widehat{L}_x \;&=\; -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) y + i\hbar y \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)\\ &= i\hbar z - i\hbar y \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) + i\hbar y \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)\\ &= i\hbar z \end{aligned} \tag{9.11}$$

The operators $\widehat{L}_i$ do not commute with the operators $\widehat{x}_j$ when the indices i and j are different. Similarly we obtain:

$$\left[\widehat{L}_x, \widehat{z}\right] = -i\hbar y. \tag{9.12}$$

In the compact notation used before, these commutation relations are summarized in the form

$$\left[\widehat{L}_i, \widehat{x}_j\right] = i\hbar\epsilon_{ijk}\widehat{x}_k. \tag{9.13}$$

9.2.2 Commutation Relations Between $\widehat{\mathbf{L}}_i$ and $\widehat{\mathbf{p}}_j$

Let us now calculate the commutator of the angular momentum components $\widehat{\mathbf{L}}_i$ with the linear momentum components $\widehat{\mathbf{p}}_i$. These results will be useful to calculate more complex commutators. We start with operators having the same subindex; for example with the commutator

$$\begin{aligned}\left[\widehat{L}_x, \widehat{p}_x\right] = \widehat{L}_x \widehat{p}_x - \widehat{p}_x \widehat{L}_x \; &= \; -\hbar^2 \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y} \right) \frac{\partial}{\partial x} + \hbar^2 \frac{\partial}{\partial x} \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y} \right) \\ &= -\hbar^2\, x \left(y\frac{\partial^2}{\partial z \partial x} - z\frac{\partial^2}{\partial y \partial x} \right) + \hbar^2 \left(y\frac{\partial}{\partial x \partial z} - z\frac{\partial^2}{\partial x \partial y} \right) \\ &= 0 \end{aligned} \tag{9.14}$$

The operators $\widehat{L}_i$ and $\widehat{p}_j$ commute when $i = j$, but when $i \neq j$, the commutator is different from zero. Let us consider for example the commutator

$$\begin{aligned}\left[\widehat{L}_x, \widehat{p}_y\right] = \widehat{L}_x \widehat{p}_y - \widehat{p}_y \widehat{L}_x \; &= \; -\hbar^2 \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y} \right) \frac{\partial}{\partial y} + \hbar^2 \frac{\partial}{\partial y} \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y} \right) \\ &= -\hbar^2 \left(y\frac{\partial^2}{\partial z \partial y} - z\frac{\partial^2}{\partial y^2} \right) + \hbar^2 \frac{\partial}{\partial z} + \hbar^2 \left(y\frac{\partial^2}{\partial y \partial z} - z\frac{\partial^2}{\partial y^2} \right) \\ &= i\hbar\, \widehat{p}_z \end{aligned} \tag{9.15}$$

Similarly, we have

$$\left[\widehat{L}_x, \widehat{p}_z\right] = -i\hbar\, \widehat{p}_y. \tag{9.16}$$

In compact notation the commutation relations of angular momentum and linear momentum components are summarized in the general expression

$$\left[\widehat{L}_i, \widehat{p}_j\right] = i\hbar\, \epsilon_{ijk}\, \widehat{p}_k. \tag{9.17}$$

The similarity between these commutation relations and those obtained in (9.13) are remarkable. Similar relations will be derived below for the commutation relations of angular momentum components.

9.2.3 Commutation Between the $\widehat{\mathbf{L}}_j$ Components

We will explicitly calculate one of the three possible commutation relations between the $\widehat{L}_j$ components. Let the commutator

$$
\begin{aligned}
\left[\widehat{L}_x, \widehat{L}_y\right] &= \widehat{L}_x \widehat{L}_y - \widehat{L}_y \widehat{L}_x \\
&= -\hbar^2 \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \left(z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) - \widehat{L}_y \widehat{L}_x \\
&= -\hbar^2 \left(y \frac{\partial}{\partial x} + yz \frac{\partial^2}{\partial z \partial x} - yx \frac{\partial^2}{\partial z^2} - z^2 \frac{\partial^2}{\partial y \partial x} + zx \frac{\partial^2}{\partial y \partial z} \right) - \widehat{L}_y \widehat{L}_x
\end{aligned}
\tag{9.18}
$$

In the last two equations, we have explicitly developed the first term $\widehat{L}_x \widehat{L}_y$. The reader should notice that the goal is to move the differential operators to the right side. If one does the same with the second term $\widehat{L}_y \widehat{L}_x$ we end up with

$$
\begin{aligned}
\left[\widehat{L}_x, \widehat{L}_y\right] &= -\hbar^2 \left(y \frac{\partial}{\partial x} - x \frac{\partial}{\partial y} \right) \\
&= i\hbar \widehat{L}_z.
\end{aligned}
\tag{9.19}
$$

We have also commutators like

$$
\left[\widehat{L}_x, \widehat{L}_x\right] = \widehat{L}_x \widehat{L}_x - \widehat{L}_x \widehat{L}_x = 0. \tag{9.20}
$$

These commutation relations can be expressed in compact form as

$$
\left[\widehat{L}_i, \widehat{L}_j\right] = i\hbar\, \epsilon_{ijk}\, \widehat{L}_k. \tag{9.21}
$$

Given the commutation relations one can establish some important physical consequences. We have seen that whenever two Hermitian operators commute, they are dispersionless (and simultaneously well-defined) variables, and in addition they share the same eigenfunctions. We have seen also that when two variables, say $\widehat{f}$ and $\widehat{g}$, do not commute, the product of their average dispersions is greater than or equal to the expected value of their commutator $\widehat{h}$, i.e.,

$$
(\Delta f)^2 (\Delta g)^2 \geq \frac{1}{4} |\langle \widehat{h} \rangle|^2. \tag{9.22}
$$

If we apply this property to the angular momentum components $\widehat{L}_i$, we have

$$
(\Delta L_i)^2 (\Delta L_j)^2 \geq \frac{\hbar^2}{4}\, |\langle \widehat{L}_k \rangle|^2. \tag{9.23}
$$

This means that *it is not possible to determine simultaneously, without dispersion, all three components of* $\widehat{\mathbf{L}}$, unless $\langle \widehat{\mathbf{L}} \rangle = 0$ and $\Delta \mathbf{L} = 0$. These results will be taken into account some lines below.

9.2.4 The Operator $\widehat{L}^2$ and Its Commutation with $\widehat{L}_j$

It turns out that an important quantity, besides the angular momentum components $\widehat{L}_i$, is the operator

$$\widehat{L}^2 = \widehat{\mathbf{L}} \cdot \widehat{\mathbf{L}}. \tag{9.24}$$

We will see later that the angular part of the kinetic energy, in central potentials, is described precisely by this operator. Other interesting features of this operator come from its commutation relations with the angular momentum components $\widehat{L}_i$. At the end of this chapter we will show explicitly the commutation relations

$$\left[\widehat{L}^2, \widehat{L}_x\right] = \left[\widehat{L}^2, \widehat{L}_y\right] = \left[\widehat{L}^2, \widehat{L}_z\right] = 0. \tag{9.25}$$

If we only pay attention to these commutation relations, we can conclude that, for a well defined angular momentum, each of its components should also be a well defined quantity, i.e., a physical variable without dispersion. But this contradicts the results at the end of the last section, which show that the angular momentum components $\widehat{L}_i$ do not commute among themselves, hence they can not be, simultaneously, well defined quantities. We mentioned lines above that at least two components should have dispersion and satisfy the Heisenberg inequality. The convention is then to choose $\widehat{L}_x$, $\widehat{L}_y$ for the angular momentum components with dispersion and leave $\widehat{L}_z$ as the well-defined quantity. Therefore

$$\Delta L_x \Delta L_y \geq \frac{\hbar}{2}\, |\langle \widehat{L}_z \rangle|. \tag{9.26}$$

Since $\widehat{L}_z$ commutes with $\widehat{L}^2$, $\widehat{L}_z$ shares eigenfunctions with $\widehat{L}^2$. We come to this issue now.

9.3 Eigenvalues and Eigenfunctions of $\widehat{L}_z$ and $\widehat{L}^2$

If we use the operator $\widehat{\mathbf{L}}$ in the spherical coordinates representation and take into account that

$$\begin{aligned} \frac{\partial \mathbf{u}_\theta}{\partial \theta} &= -\mathbf{u}_r, & \frac{\partial \mathbf{u}_\phi}{\partial \theta} &= 0, \\ \frac{\partial \mathbf{u}_\theta}{\partial \phi} &= \cos\theta \mathbf{u}_\phi, & \frac{\partial \mathbf{u}_\phi}{\partial \phi} &= -\mathbf{i}\cos\phi - \mathbf{j}\sin\phi, \end{aligned} \tag{9.27}$$

we can verify, after some algebra, that $\widehat{L}^2$ reduces to

$$\widehat{L}^2 = -\hbar^2 \left(\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial}{\partial\theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2} \right), \tag{9.28}$$

which coincides with the angular part of the Laplacian ∇^2 (see Eq. (9.88)). We will see that dealing with particles moving in central force, the Schrödinger equation is separable and, even more, the angular part of the kinetic energy operator coincides with L^2.

At this point, once the commutation relations made clear the relevance of the angular momentum operators $\widehat{L}_z$ and $\widehat{L}^2$, we will move to the explicit calculation of the eigenvalues and eigenfunctions of $\widehat{L}_z$ and $\widehat{L}^2$.

If we write

$$\widehat{L}_z \; \varphi(\phi) = \hbar \; m_z \; \varphi(\phi), \tag{9.29}$$

and

$$\widehat{L}^2 \; \psi(\theta, \phi) = \; \hbar^2 \; \lambda \; \psi(\theta, \phi) \tag{9.30}$$

our goal is to obtain the eigenvalues $\hbar m_z$ and $\hbar^2\lambda$ and the eigenfunctions $\varphi(\phi)$ and $\psi(\theta, \phi)$.

Let us start with the Eq. (9.30). If we propose the factorization

$$\psi(\theta, \phi) = \Theta(\theta)\Phi(\phi), \tag{9.31}$$

we obtain the following two equations (the choice of separation constant m^2 will be understood two lines below)

$$-\frac{\partial^2 \Phi(\phi)}{\partial\phi^2} = m^2 \Phi(\phi) \tag{9.32}$$

$$\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial}{\partial\theta} \right) \Theta(\theta) - \frac{m^2}{\sin^2\theta} \Theta(\theta) + \lambda \; \Theta(\theta) = 0. \tag{9.33}$$

The last equation is known as the associated Legendre differential equation. We will begin solving the first of these equations. It is easy to verify that the exponential function $e^{im\phi}$ is a solution and, changing m by m_z, it is also the solution of (9.29). This means that

$$\Phi(\phi) = \varphi(\phi) = e^{im_z\phi}. \tag{9.34}$$

For this solution to be unique, it is necessary that

$$\Phi(\phi) = \Phi(\phi + 2\pi). \tag{9.35}$$

This condition is satisfied if

$$e^{im_z 2\pi} = 1, \quad \text{therefore} \quad m_z = 0, \pm 1, \pm 2, \ldots \tag{9.36}$$

This shows that the eigenvalues and eigenfunctions of $\widehat{L}_z$ are $m_z\hbar$ and $e^{im_z\phi}$ with $m_z = 0, \pm1, \pm2,...$ The quantum number m_z is known as the *magnetic quantum number*. This happens because the quantum number m_z arises naturally when charged particles move in a central field and in the presence of a magnetic field, whose direction defines the z axis and the magnetic moment orientation.

Let us now return to the Legendre differential equation. If we introduce the change of variable

$$\xi = \cos\theta, \tag{9.37}$$

and rename the function $\Theta(\theta)$ as

$$P(\xi) = \Theta(\theta), \tag{9.38}$$

Eq. (9.33) transforms into

$$\frac{d}{d\xi}\left[(1-\xi^2)\frac{dP(\xi)}{d\xi}\right] - \frac{m_z^2}{1-\xi^2}P(\xi) + \lambda P(\xi) = 0. \tag{9.39}$$

In order to solve this equation, we will start with the particular case $m_z = 0$. In this case the differential equation is simply

$$\frac{d}{d\xi}\left[(1-\xi^2)\frac{dP^0(\xi)}{d\xi}\right] + \lambda P^0(\xi) = 0. \tag{9.40}$$

This is the Legendre differential equation. This equation has non-divergent solutions at $\xi = \pm1$ and remains invariant under the transformation $\xi \to -\xi$. Thus, we seek for well defined parity solutions, i.e. solutions that are even or odd in the variable ξ. To obtain these solutions, let us use the Frobenius method based on the series expansion

$$P^0(\xi) = \sum_{k=0}^{\infty} a_k \xi^k. \tag{9.41}$$

After substitution of this function into the differential equation, and performing the standard changes in the summation indices, we obtain the recurrence relation

$$a_{k+2} = \frac{k(k+1)-\lambda}{(k+1)(k+2)}a_k. \tag{9.42}$$

It is clear from this recurrence relation that the series terminates at $k = l$ when

$$\lambda = l(l+1), \tag{9.43}$$

being l a nonnegative integer. The solutions are then the Legendre polynomials. It is not difficult to show that the first four Legendre polynomials are

$$\begin{aligned} &P_0^0(\xi) = 1, && P_1^0(\xi) = \xi, \\ &P_2^0(\xi) = \frac{1}{2}\left(3\xi^2 - 1\right), && P_3^0(\xi) = \frac{1}{2}\left(5\xi^3 - 3\xi\right). \end{aligned} \tag{9.44}$$

Another representation of the Legendre polynomials is

$$P_l^0(\xi) = \frac{1}{2^l l!}\frac{d^l}{d\xi^l}\left(\xi^2 - 1\right)^l, \qquad \text{with} \qquad l = 0, 1, 2, 3, .. \tag{9.45}$$

The integer l is known as the *orbital quantum number of the angular momentum* or simply the *orbital quantum number*. Given the solutions of the Legendre equation for $m_z = 0$, it is easy to verify that the functions

$$P_l^m(\xi) = \left(1 - \xi^2\right)^{m/2}\frac{d^m P_l^0(\xi)}{d\xi^m}, \qquad \text{with} \qquad 0 \le m \le l, \tag{9.46}$$

satisfy the differential equation

$$\frac{d}{d\xi}\left[\left(1 - \xi^2\right)\frac{dP_l^m(\xi)}{d\xi}\right] - \frac{m^2}{1 - \xi^2}P_l^m(\xi) + l(l+1)P_l^m(\xi) = 0, \tag{9.47}$$

which is exactly the associated Legendre equation (9.40) that we wanted to solve. The solutions $P_l^m(\xi)$ of this equation are the *associated Legendre polynomials* of order m and degree l.

It is important to notice that solving the eigenvalues for $\widehat{L}_z$ and $\widehat{L}^2$, two quantum numbers, m_z and l, come out naturally, making evident that the rotational motion and the associated physical variables are quantized. The quantum number l can take any natural number[1] $0, 1, 2, 3, \ldots$, while m_z any of the set of integers $0, \pm 1, \pm 2, \ldots, \pm l$. Therefore, the possible values for $\widehat{L}_z$, as shown in Fig. 9.1, are $0, \pm\hbar, \pm 2\hbar, \ldots, \pm l\hbar$ while those of $\widehat{L}^2$ are $l(l+1)\hbar^2$, with the values of l just mentioned. The corresponding eigenfunctions are:

$$\Psi(\theta, \phi) = e^{im_z\phi}P_l^{m_z}(\cos\theta), \quad \text{with} \quad |m_z| \le l = 0, 1, 2, 3, \ldots \tag{9.48}$$

If these functions are normalized we have the spherical harmonics

$$Y_l^{m_z}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi}\frac{l - |m_z|}{l + |m_z|}}(-1)^{m_z}e^{im_z\phi}P_l^{m_z}(\cos\theta), \tag{9.49}$$

with $|m_z| \le l = 0, 1, 2, 3, \ldots$. The spherical harmonics form a complete set of orthonormal functions, i.e.

---

[1] We will see in the next chapter that the orbital quantum number l will, actually, be limited by the quantum number n that defines the energy level.

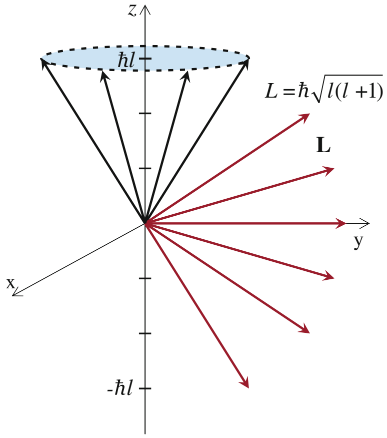


Fig. 9.1 The orientations of the angular momentum **L** are such that the length of its projection L_z along the z axis is a multiple of $\hbar$. The vector **L** precedes around the z axis with well defined projection L_z

$$\int_0^{2\pi}\int_0^{\pi} Y_{l'}^{m'*}(\theta,\phi)Y_l^m(\theta,\phi)\sin\theta d\theta d\phi = \delta_{l,l'}\delta_{m,m'}, \tag{9.50}$$

with

$$Y_l^{m^*}(\theta,\phi) = (-1)^m Y_l^{-m}(\theta,\phi). \tag{9.51}$$

It is common to call these states according to their orbital quantum number. Those with $l = 0$, are called s states (for sharp), those with $l = 1$, p states (for principal), with $l = 2$, d states (for diffusive), with $l = 3$, f states (for fundamental) and, for $l = 4, 5$,.. one has the $g, h, i, \ldots$ states, in alphabetical order. As can be seen from Fig. 9.2, where we draw the polar surfaces[2] of the spherical harmonics s, p and d, for all possible values of m_z, the angular distribution of the probability density function has specific symmetries and varies greatly from one quantum state to another. The first few spherical harmonics are:

$$\begin{aligned} Y_0^0(\theta,\phi) &= \frac{1}{\sqrt{4\pi}}, \\ Y_1^0(\theta,\phi) &= \sqrt{\frac{3}{8\pi}}\cos\theta, \\ Y_1^{\pm1}(\theta,\phi) &= \mp\sqrt{\frac{3}{4\pi}}e^{\pm i\phi}\sin\theta, \end{aligned} \tag{9.52}$$

$$Y_2^0(\theta,\phi) = \sqrt{\frac{5}{16\pi}}(3\cos^2\theta - 1),$$

---

[2] In the polar graphs the radius in the direction (θ,ϕ) is proportional to $|Y_l^{m_z}(\theta,\phi)|$.

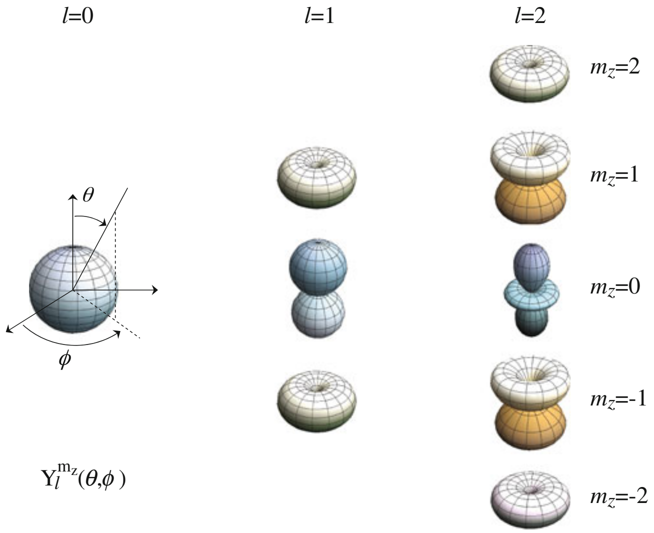


Fig. 9.2 Polar graphics of the spherical harmonics $Y_l^m(\theta, \phi)$ as functions of θ and ϕ, for different values of l and m

$$
\begin{aligned}
Y_2^{\pm1}(\theta, \phi) &= \mp\sqrt{\frac{15}{8\pi}}e^{\pm i\phi}\sin\theta\cos\theta, \\
Y_2^{\pm2}(\theta, \phi) &= \mp\sqrt{\frac{15}{32\pi}}e^{\pm i2\phi}\sin^2\theta.
\end{aligned} \tag{9.53}
$$

It is worth noticing that under a coordinate inversion through the origin, $\theta, \phi \to \pi - \theta, \phi + \pi$, $e^{im\phi}$ is multiplied by $(-1)^m$, and $P_l^{m_z}(\cos\theta)$ by $(-1)^{l+m}$. Thus $Y_l^{-m}(\theta, \phi)$ is multiplied by $(-1)^l$ and has the parity of l. To conclude this section, we recall that based on the commutation relations we were left with two relevant angular momentum operators, $\widehat{L}^2$ and $\widehat{L}_z$. Their eigenvalues and eigenfunctions satisfy the following equations

$$
\widehat{L}_z e^{im_z\phi} = \hbar m_z e^{im_z\phi}, \qquad \text{with} \qquad m_z = 0, \pm1, \pm2 ..., \pm l, \tag{9.54}
$$

and

$$
\widehat{L}^2 Y_l^{m_z}(\theta, \phi) = \hbar^2 l(l+1) Y_l^{m_z}(\theta, \phi), \qquad \text{with} \qquad l = 0, 1, 2, 3, \ldots \tag{9.55}
$$

In the compact Dirac notation, the spherical harmonics are represented by the quantum numbers l and m_z, hence one generally uses for these functions the ket

$$
|l, m_z\rangle = Y_l^{m_z}(\theta, \phi). \tag{9.56}
$$

9.4 Matrix Representations of the Angular Momentum

In this section, we will obtain matrix representations for the operators $\widehat{L}^2$, $\widehat{L}_z$, $\widehat{L}_x$ and $\widehat{L}_y$. In previous chapters we saw that the matrix representation of the Hamiltonian is diagonal when we use the Hamiltonian eigenfunctions basis, the same will occur with the representations of $\widehat{L}^2$ and $\widehat{L}_z$ if we use the spherical harmonics. The problem is slightly more complicated with the representations of the operators $\widehat{L}_x$ and $\widehat{L}_y$, which, as shown before, do not commute with $\widehat{L}_z$, and as a consequence their matrix representations are not diagonal in the spherical harmonics basis.

9.4.1 Matrix Representations of $\widehat{L}^2$ and $\widehat{L}_z$

The matrix representations of $\widehat{L}^2$ and $\widehat{L}_z$, in their eigenfunctions basis, can be written directly. If we denote the eigenfunctions of $\widehat{L}^2$ and $\widehat{L}_z$ as $|lm_z\rangle$, the matrix elements are

$$\langle lm'_z|\widehat{L}^2|lm_z\rangle = \hbar^2 l(l+1)\ \delta_{m',m_z} \tag{9.57}$$

and

$$\langle lm'_z|\widehat{L}_z|lm_z\rangle = \hbar m_z\ \delta_{m',m_z}. \tag{9.58}$$

These equations clearly show that the matrix representations of $\widehat{L}^2$ and $\widehat{L}_z$ are diagonal and of dimension $(2l+1) \times (2l+1)$. In fact, for a given l, the matrices that represent them[3] are

$$L^2 = \hbar^2 l(l+1) \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & 1 \end{pmatrix}, \quad \text{and} \quad L_z = \hbar \begin{pmatrix} l & 0 & \dots & 0 \\ 0 & l-1 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & -l \end{pmatrix}. \tag{9.59}$$

9.4.2 Matrix Representations of $\widehat{L}_x$ and $\widehat{L}_y$

If we define rising and lowering operators $\widehat{L}_+$ and $\widehat{L}_-$, like $\widehat{b}^\dagger$ and $\widehat{b}$ for the harmonic oscillator, we can use $\widehat{L}_+$ and $\widehat{L}_-$ to move throughout the spherical harmonics space. It is easy to show, and we will do that in this section, that the complex combinations

$$\widehat{L}_+ = \widehat{L}_x + i\widehat{L}_y \quad \text{and} \quad \widehat{L}_- = \widehat{L}_x - i\widehat{L}_y \tag{9.60}$$

---

[3] To arrange the matrix elements $(L_i)_{m',m}$, in the matrix representation of an operator $\widehat{L}_i$, we follow the convention to label the columns with the values of m, decreasing from left to right, and the arrows with those of m', decreasing from up to down.

act like rising and lowering operators. We will verify this feature and obtain some properties that will be used to determine the representations of $\widehat{L}_x$, $\widehat{L}_y$, that can be written as

$$\widehat{L}_x = \frac{1}{2}\left(\widehat{L}_+ + \widehat{L}_-\right) \quad \text{and} \quad \widehat{L}_y = \frac{1}{2i}\left(\widehat{L}_+ - \widehat{L}_-\right). \tag{9.61}$$

Before obtaining the matrix representations we will derive a couple of important relations.

9.4.2.1 Some properties of $\widehat{L}_+$ and $\widehat{L}_-$

Using the commutators $\left[\widehat{L}_i, \widehat{L}_j\right] = i\hbar\,\epsilon_{ijk}\,\widehat{L}_k$, one can easily show that $\widehat{L}_+$ and $\widehat{L}_-$ satisfy the following relations

$$\left[\widehat{L}_\mp, \widehat{L}_z\right] = \ \pm\hbar\widehat{L}_\mp, \tag{9.62}$$

$$\left[\widehat{L}_+, \widehat{L}_-\right] = 2\hbar\widehat{L}_z, \tag{9.63}$$

and, by simple substitution, they fulfill also the following equations

$$\widehat{L}_-\widehat{L}_+ = \widehat{L}^2 - \widehat{L}_z^2 + \hbar\widehat{L}_z, \tag{9.64}$$

and

$$\widehat{L}_+\widehat{L}_- = \widehat{L}^2 - \widehat{L}_z^2 - \hbar\widehat{L}_z. \tag{9.65}$$

We will see now that the operators $\widehat{L}_+$ and $\widehat{L}_-$, acting on the eigenfunctions $|lm_z\rangle$ of $\widehat{L}^2$ and $\widehat{L}_z$, behave as rising and lowering operators. From the commutation relation (9.62) we have

$$\widehat{L}_z\widehat{L}_+|lm_z\rangle - \widehat{L}_+\widehat{L}_z|lm_z\rangle = \hbar\widehat{L}_+|lm_z\rangle, \tag{9.66}$$

which becomes

$$\widehat{L}_z\widehat{L}_+|lm_z\rangle = \hbar(m_z+1)\widehat{L}_+|lm_z\rangle. \tag{9.67}$$

This equation shows that $\widehat{L}_+|lm_z\rangle$ is an eigenfunction of $\widehat{L}_z$ with eigenvalue $\hbar(m_z+1)$. This tells us that, indeed, the action of $\widehat{L}_+$ on $|lm_z\rangle$ leaves a function proportional to $|l, m_z+1\rangle$. i.e. that

$$\widehat{L}_+|lm_z\rangle = l^+_{m_z+1,m_z}|l, m_z+1\rangle. \tag{9.68}$$

Replacing this functión for $\widehat{L}_+|lm_z\rangle$ in (9.67) and dividing by $l^+_{m_z+1,m_z}$, we have the well known relation

$$\widehat{L}_z|l, m_z+1\rangle = \hbar(m_z+1)|l, m_z+1\rangle. \tag{9.69}$$

Similarly, one can show that

$$\widehat{L}_-|lm_z\rangle = l^-_{m_z-1,m_z}|l, m_z - 1\rangle. \tag{9.70}$$

The operators $\widehat{L}_+$ and $\widehat{L}_-$ are rising and lowering operators in the Hilbert space of the eigenfunctions of $\widehat{L}_z$. Although we know that the factors $l^+_{m_z+1,m_z}$ and $l^-_{m_z-1,m_z}$ should have units of angular momentum, they are still unknown. We will now determine these factors. It is worth mentioning that all properties that we are obtaining here for the angular momentum $\widehat{L}$ and its components $\widehat{L}_i$, are valid for other angular momentum operators, like the spin $\widehat{\mathbf{S}}$ and the total angular momentum $\widehat{\mathbf{J}} = \widehat{\mathbf{S}} + \widehat{\mathbf{L}}$, that we will find later. The commutation relations and properties we are pointing out for $\widehat{\mathbf{L}}$ can be written also for $\widehat{\mathbf{S}}$ and $\widehat{\mathbf{J}}$, with the corresponding change of L by S or by J.

To determine the factors $l^+_{m+1,m}, l^-_{m-1,m}$ let us consider the Eq. (9.65) and multiply it to the right by $|lm_z\rangle$ and to the left by $\langle lm_z|$; we thus have the matrix elements

$$\langle lm_z|\widehat{L}_+\widehat{L}_-|lm_z\rangle = \langle m_z|\widehat{L}^2|lm_z\rangle - \langle lm_z|\widehat{L}^2_z|lm_z\rangle + \langle lm_z|\hbar\widehat{L}_z|lm_z\rangle, \tag{9.71}$$

which transforms into

$$l^+_{m_z,m_z-1}l^-_{m_z-1,m_z} = \hbar^2\Big[l(l+1) - m_z(m_z-1)\Big]. \tag{9.72}$$

On the right side we have a real number, and since $\widehat{L}_+ = \widehat{L}^*_-$, one can easily conclude that

$$l^+_{m_z,m_z-1} = \big(l^-_{m_z-1,m_z}\big)^*. \tag{9.73}$$

Thus, we have, on one side that

$$|l^+_{m_z,m_z-1}|^2 = \hbar^2\big(l(l+1) - m_z(m_z-1)\big), \tag{9.74}$$

and, on the other

$$|l^-_{m_z-1,m_z}|^2 = \hbar^2\big(l(l+1) - m_z(m_z-1)\big). \tag{9.75}$$

These relations, with a change of m_z by $m_z + 1$ in the first case, can be written as follows:

$$|l^+_{m_z+1,m_z}| = \hbar\sqrt{l(l+1) - (m_z+1)m_z}, \tag{9.76}$$

$$|l^-_{m_z-1,m_z}| = \hbar\sqrt{l(l+1) - (m_z-1)m_z}. \tag{9.77}$$

With these results, and choosing the phases $e^{i\theta^\pm}$ of $l^+_{m_z,m_z-1}$ and $l^-_{m_z-1,m_z}$ in the Condon–Shortley convention, where $e^{i\theta^\pm} = 1$, it is easy to verify that

$$\widehat{L}_+ \, |lm_z\rangle = \hbar\sqrt{(l+m_z+1)(l-m_z)} \, |l, m_z+1\rangle, \tag{9.78}$$

and

$$\widehat{L}_-|lm_z\rangle = \hbar\sqrt{(l+m_z)(l-m_z+1)} \, |l, m_z-1\rangle. \tag{9.79}$$

9.4.2.2 The Matrix Representations of $\widehat{L}_x$ and $\widehat{L}_y$

If we multiply (9.78) on the left with the bra $\langle lm'_z|$, we have the matrix elements

$$\begin{aligned}(L_+)_{m'_z,m_z} &= \langle lm'_z|\widehat{L}_+|lm_z\rangle = l^+_{m+1,m}\langle lm'_z|l, m_z+1\rangle \\ &= l^+_{m+1,m}\delta_{m'_z,m_z+1},\end{aligned} \tag{9.80}$$

that, using the convention to arrange the matrix elements mentioned before, leads us to write, for a given value of l, the matrix representation of $\widehat{L}_+$ as

$$L_+ = \begin{pmatrix} 0 & l^+_{l,l-1} & 0 & \dots & 0 & 0 \\ 0 & 0 & l^+_{l-1,l-2} & \dots & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & \dots & 0 & l^+_{-l+1,-l} \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}. \tag{9.81}$$

In the same way, we have the matrix elements

$$\begin{aligned}(L_-)_{m'_z,m_z} &= \langle lm'_z|\widehat{L}_-|lm_z\rangle = l^-_{m-1,m}\langle lm'_z|l, m_z-1\rangle \\ &= l^-_{m-1,m}\delta_{m'_z,m_z-1}.\end{aligned} \tag{9.82}$$

In (9.83) we expressed the operators $\widehat{L}_x$ and $\widehat{L}_y$ in terms of the rising and lowering operators. The matrix representations L_x and L_y are obtained from similar combinations of the matrix representations L_+ and L_-. Therefore, given

$$\widehat{L}_x = \frac{1}{2}\left(\widehat{L}_+ + \widehat{L}_-\right) \quad \text{and} \quad \widehat{L}_y = \frac{1}{2i}\left(L_+ - L_-\right), \tag{9.83}$$

we find the matrices

$$\begin{aligned}(L_x)_{m'_z,m_z} &= \frac{1}{2}\left(L_+ + L_-\right)_{m'_z,m_z} \\ &= \frac{1}{2}\left(l^+_{m+1,m}\delta_{m'_z,m_z+1} + l^-_{m-1,m}\delta_{m'_z,m_z-1}\right),\end{aligned} \tag{9.84}$$

and

$$(L_y)_{m_z',m_z} = \frac{1}{2i}\,(L_+ - L_-)_{m_z',m_z}$$
$$= \frac{1}{2i}\left(l^+_{m+1,m}\delta_{m_z',m_z+1} - l^-_{m-1,m}\delta_{m_z',m_z-1}\right). \qquad (9.85)$$

Both matrices have nonzero elements only in the diagonals that are the first neighbors of the main diagonal. Using these expressions, the reader can verify that the matrix representations L_x and L_y, for $l = 1$, are

$$L_x = \frac{\hbar}{2}\begin{pmatrix} 0 & \sqrt{2} & 0 \\ \sqrt{2} & 0 & \sqrt{2} \\ 0 & \sqrt{2} & 0 \end{pmatrix} \quad \text{and} \quad L_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i\sqrt{2} & 0 \\ i\sqrt{2} & 0 & -i\sqrt{2} \\ 0 & i\sqrt{2} & 0 \end{pmatrix}. \qquad (9.86)$$

9.5 Central Potentials

Among the interactions that particles experience, the central potentials represent an important and very special class of interaction. If two particles, with position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ interact through a central force, the potential energy depends only on the distance $r = |\mathbf{r}_2 - \mathbf{r}_1|$. The gravitational and electric interactions are examples of central potentials. In some systems as the Hydrogen atom, the electron-proton interaction is central. A non-central interaction is, for example, the interaction between a charge and a uniform external field. In Hydrogenic metals the interaction between the peripheral electron[4] (the valence electron) with the nucleus and the remaining (core) electrons, is described, approximately, by an *effective potential* with spherical symmetry. In these cases, as in the Hydrogen atom, the Schrödinger equation has the form

$$-\frac{\hbar^2}{2m}\nabla^2\varphi(\mathbf{r}) + V(r)\varphi(\mathbf{r}) = E\,\varphi(\mathbf{r}). \qquad (9.87)$$

To deal with this family of systems in this approximation, it is desirable to use the spherical coordinates system where the differential operator ∇^2 is

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}. \qquad (9.88)$$

The first term of the right hand side describes the kinematics in the radial direction and the last two terms describe the angular motion. In the previous section, we have seen that the operator $\widehat{L}^2$, expressed in the spherical coordinates, is

$$\widehat{L}^2 = \widehat{\mathbf{L}}\cdot\widehat{\mathbf{L}} = -\hbar^2\left(\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right). \qquad (9.89)$$

---

[4] In general by peripheral electrons, we have in mind electrons in the upper occupied energy level.

Therefore, the Schrödinger equation for a central potential can be written as

$$\left(-\frac{\hbar^2}{2m}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)+\frac{\widehat{L}^2}{2mr^2}+V(r)\right)\varphi(r,\theta,\phi)=E\,\varphi(r,\theta,\phi). \tag{9.90}$$

Our purpose is to obtain the complete solution $\varphi(r,\theta,\phi)$ of this problem for specific potentials $V(r)$. We will make a progress in this chapter discussing the solution of the angular part. In Sect. 9.3, we found already the eigenvalues and eigenfunctions of $\widehat{L}^2$. We will use them, to define the angular part of the central potential solution. In the next chapter, devoted to study the hydrogen atom, we will discuss the radial part of $\varphi(r,\theta,\phi)$.

The role that the angular momentum $\widehat{\mathbf{L}}$ plays in a central potential problem is now clear. The operator $\widehat{L}^2$ in the Hamiltonian, contains all the information of the angular motion in the central potential problem. Although this term $\widehat{L}^2/2mr^2$ originates from the kinetic energy, it is known as the *centrifugal potential*.

From a mathematical point of view, the differential equation in (9.90) is separable. If we propose a solution like

$$\varphi(r,\theta,\phi)=R(r)\psi(\theta,\phi) \tag{9.91}$$

and replace it in (9.90), we obtain the following equations

$$\widehat{L}^2\psi(\theta,\phi)=\hbar^2\lambda\psi(\theta,\phi), \tag{9.92}$$

$$\left(-\frac{\hbar^2}{2m}\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right)+\frac{\hbar^2\lambda}{2mr^2}+V(r)\right)R(r)=E\,R(r), \tag{9.93}$$

with $\hbar^2\lambda$ the separation constant. We know already the first equation, it is exactly the Eq. (9.30) whose eigenvalues are

$$\hbar^2\lambda=\hbar^2 l(l+1) \tag{9.94}$$

and whose eigenfunctions are the spherical harmonics

$$\psi(\theta,\phi)=Y_l^{m_z}(\theta,\phi)=\sqrt{\frac{2l+1}{4\pi}\frac{l-|m_z|}{l+|m_z|}}(-1)^{m_z}e^{im_z\phi}P_l^{m_z}(\cos\theta). \tag{9.95}$$

This result confirms that $\widehat{L}^2$ and the Hamiltonian with a central potential share eigenfunctions. Since the angular part is already solved, what remains open is the differential equation

$$\left[-\frac{\hbar^2}{2m}\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right)+\frac{\hbar^2 l(l+1)}{2mr^2}+V(r)\right]R(r)=E\,R(r), \tag{9.96}$$

that we will refer to as the radial part of the central-potential Schrödinger equation. In the next chapter we will discuss this problem with $V(r)$ the Coulomb potential.

9.6 Solved Problems

Exercise 32 Show that the angular momentum $\widehat{\mathbf{L}}$ commutes with the central potential Hamiltonian.
Solution To show that the angular momentum operator $\widehat{\mathbf{L}}$ commutes with a central potential Hamiltonian, we have two alternatives. In the first and easier, we need just to observe the Hamiltonian in (9.90) and since $\widehat{\mathbf{L}}$ commutes with itself and commutes also with functions of r and $\partial/\partial r$, it commutes with the Hamiltonian. Another alternative is through the explicit evaluation of the commutator

$$\left[\widehat{\mathbf{L}}, H\right] = \frac{1}{2m}\left[\widehat{\mathbf{L}}, \widehat{p}^2\right] + \left[\widehat{\mathbf{L}}, V(r)\right]. \tag{9.97}$$

Since

$$\widehat{p}^2 = \widehat{\mathbf{p}} \cdot \widehat{\mathbf{p}} = \widehat{p}_x^2 + \widehat{p}_y^2 + \widehat{p}_z^2 \tag{9.98}$$

and

$$\begin{aligned}\left[\widehat{\mathbf{L}}, \widehat{p}_i^2\right] &= \widehat{\mathbf{L}}\widehat{p}_i\widehat{p}_i - \widehat{p}_i\widehat{p}_i\widehat{\mathbf{L}} = \widehat{\mathbf{L}}\widehat{p}_i\widehat{p}_i - \widehat{p}_i\widehat{\mathbf{L}}\widehat{p}_i + \widehat{p}_i\widehat{\mathbf{L}}\widehat{p}_i - \widehat{p}_i\widehat{p}_i\widehat{\mathbf{L}}\\ &= \left[\widehat{\mathbf{L}}, \widehat{p}_i\right]\widehat{p}_i + \widehat{p}_i\left[\widehat{\mathbf{L}}, \widehat{p}_i\right],\end{aligned} \tag{9.99}$$

it is easy to conclude, using the commutation relations obtained before, that

$$\left[\widehat{\mathbf{L}}, \widehat{p}^2\right] = 0. \tag{9.100}$$

We have still to determine the commutator $\left[\widehat{\mathbf{L}}, V(r)\right]$. For this purpose we develop the commutator as follows

$$\left[\widehat{\mathbf{L}}, V(r)\right] = -i\hbar\mathbf{r} \times \left[\nabla V(r)\right] - i\hbar\mathbf{r} \times V(r)\nabla + i\hbar V(r)\mathbf{r} \times \nabla. \tag{9.101}$$

If we use the well known relation $\mathbf{F} = -\nabla V(r)$, where $\mathbf{F}$ is the external force, we are left with

$$\left[\widehat{\mathbf{L}}, V(r)\right] = i\hbar\mathbf{r} \times \mathbf{F} = i\hbar\mathbf{M}. \tag{9.102}$$

Here $\mathbf{M}$ is the torque. Thus

$$\left[\widehat{\mathbf{L}}, H\right] = i\hbar\mathbf{M}. \tag{9.103}$$

In central potentials $\mathbf{F}$ is parallel to $\mathbf{r}$, hence the torque is zero and the angular momentum $\widehat{\mathbf{L}}$ commutes with the central potential Hamiltonian.

Exercise 33 Show that

$$\left[\widehat{L}^2, \widehat{L}_x\right] = 0. \tag{9.104}$$

Solution We write this commutator explicitly as

$$\left[\widehat{L}_x^2, \widehat{L}_x\right] + \left[\widehat{L}_y^2, \widehat{L}_x\right] + \left[\widehat{L}_z^2, \widehat{L}_x\right] = 0. \tag{9.105}$$

Since $\widehat{L}_x^2$, obviously, commutes with $\widehat{L}_x$, the first commutator is zero. We still need to show that

$$\left[\widehat{L}_y^2, \widehat{L}_x\right] + \left[\widehat{L}_z^2, \widehat{L}_x\right] = 0. \tag{9.106}$$

We will evaluate these commutators separately. We start evaluating the commutator $\left[\widehat{L}_y^2, \widehat{L}_x\right]$. If we develop this quantity as

$$\begin{aligned}\left[\widehat{L}_y^2, \widehat{L}_x\right] &= \widehat{L}_y^2\widehat{L}_x - \widehat{L}_x\widehat{L}_y^2 \\ &= \widehat{L}_y^2\widehat{L}_x - \widehat{L}_y\widehat{L}_x\widehat{L}_y + \widehat{L}_y\widehat{L}_x\widehat{L}_y - \widehat{L}_x\widehat{L}_y^2 \\ &= \widehat{L}_y\left[\widehat{L}_y, \widehat{L}_x\right] + \left[\widehat{L}_y, \widehat{L}_x\right]\widehat{L}_y,\end{aligned} \tag{9.107}$$

and replace the commutator $\left[\widehat{L}_y, \widehat{L}_x\right]$ by $-i\hbar\widehat{L}_z$ we obtain

$$\left[\widehat{L}_y^2, \widehat{L}_x\right] = -i\hbar\big(\widehat{L}_y\widehat{L}_z + \widehat{L}_z\widehat{L}_y\big). \tag{9.108}$$

In the same way, we have

$$\begin{aligned}\left[\widehat{L}_z^2, \widehat{L}_x\right] &= \widehat{L}_z^2\widehat{L}_x - \widehat{L}_x\widehat{L}_z^2 \\ &= \widehat{L}_z^2\widehat{L}_x - \widehat{L}_z\widehat{L}_x\widehat{L}_z + \widehat{L}_z\widehat{L}_x\widehat{L}_z - \widehat{L}_x\widehat{L}_z^2 \\ &= \widehat{L}_z\left[\widehat{L}_z, \widehat{L}_x\right] + \left[\widehat{L}_z, \widehat{L}_x\right]\widehat{L}_z.\end{aligned} \tag{9.109}$$

Replacing the commutator $\left[\widehat{L}_z, \widehat{L}_x\right]$ by $i\hbar\widehat{L}_y$, we obtain

$$\left[\widehat{L}_z^2, \widehat{L}_x\right] = i\hbar\big(\widehat{L}_z\widehat{L}_y + \widehat{L}_y\widehat{L}_z\big), \tag{9.110}$$

which is exactly the negative of the commutator $\left[\widehat{L}_y^2, \widehat{L}_x\right]$. Therefore, if we substitute these results in (9.106), we have what we wanted to prove, that

$$\left[\widehat{L}^2, \widehat{L}_x\right] = 0 \tag{9.111}$$

Exercise 34 Suppose we have a central potential system with angular momentum $\widehat{\mathbf{L}}$ that does not depend explicitly on time. Show that in this case the angular momentum is a constant of motion.

Solution Taking into account the general relation (7.119), applied to the angular momentum, we have

$$\frac{d\widehat{\mathbf{L}}}{dt} = \frac{\partial \widehat{\mathbf{L}}}{\partial t} + \left[\widehat{\mathbf{L}}, H\right]. \tag{9.112}$$

If the angular momentum does not depend explicitly on time, the first term on the right hand side is zero. If in addition we use that $\left[\widehat{\mathbf{L}}, H\right] = 0$, we end up with

$$\frac{d\widehat{\mathbf{L}}}{dt} = 0. \tag{9.113}$$

Therefore, the angular momentum is a conserved quantity. In central potential systems, the external torque is zero and the angular momentum $\widehat{\mathbf{L}}$ is a constant of motion.

9.7 Problems

1. Suppose a free particle in the state $\psi = e^{i\mathbf{k}\cdot\mathbf{r}}$ with $\mathbf{r} = x_1\widehat{\mathbf{x}}_1 + x_2\widehat{\mathbf{x}}_2 + x_3\widehat{\mathbf{x}}_3$, and $\mathbf{k} = k_1\widehat{\mathbf{x}}_1 + k_2\widehat{\mathbf{x}}_2 + k_3\widehat{\mathbf{x}}_3$ constant:
   a) show that $L_i = \hbar(\mathbf{r} \times \mathbf{k}) \cdot \widehat{\mathbf{x}}_i$;
   b) determine the angular momentum $\widehat{\mathbf{L}}$
2. If the particle of the previous problem moves in the plane $x_2 - x_3$ with $\mathbf{k} = k_2\widehat{\mathbf{x}}_2$, what is the expected value of $\widehat{L}_z$?
3. Explain whether the angular momentum $\widehat{\mathbf{L}}$ of a free particle commutes with its Hamiltonian or not.
4. The orbital quantum number of a particle in a central potential is $l = 3$. What are the possible values of m_z? What is the probability that $\langle L_z \rangle = \hbar$?
5. Using the commutation relations for $\widehat{L}_i$, show that the operator

$$\widehat{\mathbf{L}} \times \widehat{\mathbf{L}} = \left(\widehat{L}_y\widehat{L}_z - \widehat{L}_z\widehat{L}_y\right)\mathbf{i} + \left(\widehat{L}_z\widehat{L}_x - \widehat{L}_x\widehat{L}_z\right)\mathbf{j} + \left(\widehat{L}_x\widehat{L}_y - \widehat{L}_y\widehat{L}_z\right)\mathbf{k}, \tag{9.114}$$

   can be written as

$$\widehat{\mathbf{L}} \times \widehat{\mathbf{L}} = i\hbar\widehat{\mathbf{L}}. \tag{9.115}$$

6. Prove that

$$\left[\widehat{L}_{\mp}, \widehat{L}_z\right] = \pm\hbar\widehat{L}_{\mp}, \tag{9.116}$$

   and

$$\left[\widehat{L}_{+}, \widehat{L}_{-}\right] = 2\hbar\widehat{L}_z, \tag{9.117}$$

7. Prove that

$$\widehat{L}_{-}|lm_z\rangle = l^{-}_{m_z-1,m_z}|l, m_z - 1\rangle. \tag{9.118}$$

   Show also that

$$l^{+}_{m_z,m_z-1} = \left(l^{-}_{m_z-1,m_z}\right)^{*}. \tag{9.119}$$

and verify that

$$|l^{+}_{m_z+1,m_z}| = \hbar\sqrt{l(l+1)-(m_z+1)m_z}, \tag{9.120}$$

and

$$|l^{-}_{m_z-1,m_z}| = \hbar\sqrt{l(l+1)-(m_z-1)m_z}. \tag{9.121}$$

8. Prove explicitly the following commutation relations

$$\left[\widehat{L}^2, \widehat{L}_x\right] = \left[\widehat{L}^2, \widehat{L}_y\right] = \left[\widehat{L}^2, \widehat{L_z}\right] = 0. \tag{9.122}$$

9. and show also the following relations

$$\begin{aligned} \frac{\partial \widehat{\mathbf{u}}_\theta}{\partial \theta} &= -\widehat{\mathbf{u}}_r & \frac{\partial \widehat{\mathbf{u}}_\phi}{\partial \theta} &= 0 \\ \frac{\partial \widehat{\mathbf{u}}_\theta}{\partial \phi} &= \cos\theta \widehat{\mathbf{u}}_\phi & \frac{\partial \widehat{\mathbf{u}}_\phi}{\partial \phi} &= -\widehat{\mathbf{i}}\cos\phi - \widehat{\mathbf{j}}\sin\phi. \end{aligned} \tag{9.123}$$

10. Use the Frobenius series in (9.41) and shows the relations (9.42) and (9.45).
11. Plot the first spherical harmonics as functions of θ, for $\phi = \pi/4$.
12. Show that for $l = 1$, one has certainly the matrix representations

$$L_x = \frac{\hbar}{2}\begin{pmatrix} 0 & \sqrt{2} & 0 \\ \sqrt{2} & 0 & \sqrt{2} \\ 0 & \sqrt{2} & 0 \end{pmatrix} \quad \text{and} \quad L_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i\sqrt{2} & 0 \\ i\sqrt{2} & 0 & -i\sqrt{2} \\ 0 & i\sqrt{2} & 0 \end{pmatrix}. \tag{9.124}$$

13. Obtain the matrix representations of L_x and L_y for $l = 1/2$.

Chapter 10
The Hydrogen Atom

10.1 Introduction

Some of the problems that were the center of discussions during the long genesis of the quantum theory were the atomic structure and the radiated spectral lines, in particular the emission lines of Hydrogen. A convincing explanation of this problem had to wait to the advent of the quantum theory. The Bohr postulates were, as will be seen below, unable to explain convincingly the electronic configuration and the spectral lines of atoms in the presence of a magnetic field. The application of the quantum theory to the atomic problem was undoubtedly the litmus test of the new theory. We will now present an introduction to this problem. It is usual to postpone the discussion of the atomic Schrödinger equation, even for the simplest atom, the Hydrogen atom, because it requires solving a three-dimensional problem. In Chap. 9 we advanced a significant part for this task. We saw that, when the potential is central, the angular part is basically described by the angular momentum $\widehat{L}^2$, whose eigenfunctions are the spherical harmonics $Y_l^m(\theta, \phi)$. We know then the angular part of the whole solution and we still need to solve the radial equation. In this chapter we will have a well defined function for the potential $V(r)$ and we will obtain solutions for this specific potential.

The Hydrogen atom consists of an electron orbiting around its nucleus, which contains a proton. Some Hydrogen isotopes contain, besides the proton, one neutron (the deuterium) or two neutrons (the tritium). The proton is 1840 times heavier than the electron. For this reason, a good approximation is to assume that the nucleus is at rest, with the electron orbiting around it.[1] The most important interaction between the nucleus and the electron is the Coulomb interaction

---

[1] If we do not make this assumption, the Hydrogen kinetic energy will be the sum of the proton and the electron kinetic energies. At the end of this chapter, in the solved problems section, we show that the two-particle problem can be reduced to that of two quasi-particles: the center of mass, with mass $m + m_p$ (m is the mass of electron and m_p is the mass of proton), and the relative particle with mass $\mu = mm_p/(m + m_p)$. As this mass differs by less than 1% from m, the problem is essentially equivalent to assuming that the proton is at rest, and the electron orbits around it.

P. Pereyra, *Fundamentals of Quantum Physics*, Undergraduate Lecture Notes in Physics,
DOI: 10.1007/978-3-642-29378-8_10,

$$V(r) = -\frac{e^2}{r} \quad \text{CGS} \quad \text{and} \qquad V(r) = -\frac{e^2}{4\pi\epsilon_o r} \quad \text{MKS}. \tag{10.1}$$

For simplicity we will use the CGS system of units. We can change to the MKS units replacing e^2 by $e^2/4\pi\epsilon_o$. To extend our results for Hydrogen-like atoms, with one electron in the valence band, but Z protons in the nucleus, we have to consider the potential energy

$$V(r) = -\frac{Ze^2}{r} \tag{10.2}$$

which will be approximately the potential energy of the valence band electron.[2] When Z is one, the results correspond to Hydrogen atom. The Schrödinger equation of the valence band electron in a Hydrogen-like atom is then

$$-\frac{\hbar^2}{2m}\nabla^2\varphi(\mathbf{r}) - \frac{Ze^2}{r}\varphi(\mathbf{r}) = E\,\varphi(\mathbf{r}), \tag{10.3}$$

that in the system of spherical coordinates takes the form

$$\left[-\frac{\hbar^2}{2m}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\widehat{L}^2}{2mr^2} - \frac{Ze^2}{r}\right]\varphi(r,\theta,\phi) = E\,\varphi(r,\theta,\phi). \tag{10.4}$$

As mentioned before, the whole angular dependence is comprised in the angular momentum operator $\widehat{L}^2$, whose eigenfunctions are the spherical harmonics $Y_l^{m_z}(\theta,\phi)$, see Eq. (9.55). If we propose the factorization

$$\varphi(r,\theta,\phi) = R(r)Y_l^{m_z}(\theta,\phi), \tag{10.5}$$

and substitute it into (10.4), we have

$$\left[-\frac{\hbar^2}{2m}\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right) + \frac{\hbar^2 l(l+1)}{2mr^2} - \frac{Ze^2}{r}\right]R(r) = E\,R(r), \tag{10.6}$$

which is similar to the radial Eq. (9.96) that we had at the end of Chap. 9. Here we have a specific function for the central potential and our goal is to solve the eigenvalue problem. The solutions of this problem will tell us what are the allowed energy values for the electron in the atom of Hydrogen, and which the peripheral and less bounded electron of the Hydrogen-like atoms.

---

[2] The $Z-1$ electrons of the inner levels have roughly a spherical distribution, their effect on the valence band electron is then negligible.

10.2 The Energy Levels of the Hydrogen Atom

The centrifugal term of the differential Eq. (10.6), is proportional to r^{-2}, and decreases with r more rapidly than the Coulomb potential, but when $r \rightarrow 0$, it gets larger than the Coulomb potential. To solve the radial differential equation, it is convenient to explore the radial solutions in both limits: $r \rightarrow \infty$ and $r \rightarrow 0$. Before doing this, we will introduce some useful simplifications. Taking into account that

$$-\frac{\hbar^2}{2m}\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right) = -\frac{\hbar^2}{2m}\frac{1}{r}\frac{d^2}{dr^2}r, \tag{10.7}$$

and using this identity in (10.6), we have the equation

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dr^2} + \frac{\hbar^2 l(l+1)}{2mr^2} - \frac{Ze^2}{r}\right] rR(r) = E\,rR(r), \tag{10.8}$$

that makes clear the convenience of introducing, instead of $R(r)$, the function

$$u(r) = rR(r). \tag{10.9}$$

The differential Eq. (10.8) and its relations with well known special differential equations will be clear if we write it in terms of dimensionless variables. To transform it into a dimensionless equation we introduce the change of variable

$$\rho = 2\alpha r, \tag{10.10}$$

and define the parameters (we have seen in Chap. 1, and we will see lines below that the electron energies in the Hydrogen atom are negative)

$$\alpha^2 = \frac{2m|E|}{\hbar^2} = -\frac{2mE}{\hbar^2} \quad \text{and} \quad a_o = \frac{\hbar^2}{me^2}. \tag{10.11}$$

It is convenient to check that α and a_o have units of inverse length and length, respectively. The differential equation becomes then of the form

$$\frac{d^2u}{d\rho^2} + \left(-\frac{1}{4} + \frac{Z}{\alpha a_o}\frac{1}{\rho} - \frac{l(l+1)}{\rho^2}\right)u(\rho) = 0. \tag{10.12}$$

Our purpose now is to solve this equation. We will use a procedure similar to that used in the harmonic oscillator case, we will first explore how the Hydrogen atom solutions look like, or behave, in the limits $\rho \rightarrow 0$ and $\rho \rightarrow \infty$. Knowing the asymptotic solutions we will be able to guess part of $u(r)$, and to determine the correct solution. Once we obtain $u(r)$, we will have the radial solution $R(r)$. At the same time, we will determine the energy eigenvalues for the Hydrogen atom, and for the Hydrogen-like atoms as well.

The solution near to the origin

When $\rho \rightarrow 0$, the centrifugal potential is much greater than the absolute value of the Coulomb potential, and this is greater than $|E|$, i.e.:

$$\frac{\hbar^2 l(l+1)}{2mr^2} \gg \frac{Ze^2}{r} \gg |E|. \tag{10.13}$$

In this regime one can approximate the radial equation by

$$-\frac{d^2 u_o(\rho)}{d\rho^2} + \frac{l(l+1)}{\rho^2} u_o(\rho) = 0. \tag{10.14}$$

It is easy to see that if we propose a polynomial solution $u_o(\rho)$ of the form ρ^{s+1}, and replace it in the approximate differential equation, we will have:

$$(s+1)s - l(l+1) = 0, \tag{10.15}$$

with solutions

$$s = -\frac{1}{2} \pm (l + \frac{1}{2}) = l, -l-1. \tag{10.16}$$

When $s = -l-1$, the solution diverges at the origin, hence we throw it away and we keep only the solution

$$u_o(\rho) = \rho^{l+1}. \tag{10.17}$$

This means that near the origin the solution $u(\rho)$, for the radial equation, must be proportional to ρ^{l+1}, i.e.:

$$u(\rho) \propto \rho^{l+1}. \tag{10.18}$$

The solution far from the origin

When $\rho \rightarrow \infty$, the centrifugal potential is much smaller than the Coulomb potential, and this much smaller than the energy $|E|$, i.e.:

$$\frac{\hbar^2 l(l+1)}{2mr^2} \ll \frac{Ze^2}{r} \ll |E|. \tag{10.19}$$

In this region, the approximate radial equation will be

$$\frac{d^2 u_\infty(\rho)}{d\rho^2} = \frac{1}{4} u_\infty(\rho). \tag{10.20}$$

If we propose for the radial solution $u_\infty(\rho)$ an exponential function of the form $e^{q\rho}$, we will have

$$q^2 = \frac{1}{4}. \tag{10.21}$$

Thus,

$$q = \pm\frac{1}{2}. \tag{10.22}$$

Because of the finiteness condition, which the radial solution must satisfy, we keep only the function

$$u_\infty(\rho) = e^{-\rho/2}. \tag{10.23}$$

This means that in the far region

$$u(\rho) \propto e^{-\rho/2}. \tag{10.24}$$

With this information we are almost able to take the last step to obtain the solution of the radial equation of the Hydrogen atom.

Knowing the behavior of the radial solutions in the near and far regions, we can propose the function

$$u = \rho^{l+1} e^{-\rho/2} Q(\rho). \tag{10.25}$$

The unknown function $Q(\rho)$ should be controlled by the asymptotic solutions, therefore as $\rho \to \infty$ the function $Q(\rho)$ should increase less rapidly than the exponential factor $e^{-\rho/2}$. In other words, we require a polynomial behavior for $Q(\rho)$. Replacing the proposed function into the differential equation we have

$$\frac{d^2 Q}{d\rho^2} + \left(\frac{2(l+1)}{\rho} - 1\right)\frac{dQ}{d\rho} + \frac{(Z/\alpha a_o) - l - 1}{\rho} Q = 0. \tag{10.26}$$

This equation is similar to the Laguerre differential equation

$$\frac{d^2 y}{dx^2} + \left(\frac{\mu+1}{x} - 1\right)\frac{dy}{dx} + \frac{k}{x} y = 0. \tag{10.27}$$

which for $\mu = 1, 3, 5, \ldots$ and $k = 0, 1, 2, 3, \ldots$, is fulfilled by the generalized (or associated) Laguerre polynomial[3] $L_k^\mu(x)$. Taking

$$\mu = 2l + 1, \tag{10.28}$$

and

$$\frac{Z}{\alpha a_o} - l - 1 = k, \qquad k = 0, 1, 2, 3, \ldots \tag{10.29}$$

---

[3] These polynomials are also called associated Laguerre's polynomials. In the literature of these polynomials there are variations in the notation and definition that should be taken into account.

the solution of (10.26) will be the generalized Laguerre polynomial $L_k^{\mu}(\rho)$. The last relation is of great importance in this problem. From this relation comes out the energy quantization of the Hydrogen atom. It is wonted to define, instead of k, the number $n = k + l + 1$, and call it *the principal quantum number*, that satisfies the important order relation

$$l \leq n - 1. \tag{10.30}$$

In terms of the principal quantum number, the quantization condition (10.29) takes the form

$$Z/\alpha a_o = n \quad \text{with} \quad n = 1, 2, 3, \ldots \tag{10.31}$$

Replacing α in this equation we have

$$E_n = -\frac{Z^2 e^2}{2a_o n^2} = -\frac{Z^2 m e^4}{2\hbar^2 n^2}. \tag{10.32}$$

This defines the energy eigenvalues of the Hydrogen atom, in the CGS units. In the MKS units, we have

$$E_n = -\frac{Z^2 m e^4}{8\epsilon_o^2 h^2 n^2}, \tag{10.33}$$

and the radial functions

$$R_{nl}(\rho) = A_n \rho^l e^{-\rho/2} L_{n-l-1}^{2l+1}(\rho). \tag{10.34}$$

with

$$L_k^{\mu} = e^{\rho} \frac{\rho^{-\mu}}{k!} \frac{d^k}{d\rho^k} e^{-\rho} \rho^{k+\mu}. \tag{10.35}$$

It is important to notice that the electron energies E_n in the Hydrogen and Hydrogen-like atoms depend only on the principal quantum number n. This means that all states with different quantum numbers l and m_z, but with the same n, are degenerate states. A little later we will see how many electrons can be arranged in each of the energy levels.

We got in this way one of the physical results in the explanation of the spectral lines, the electron energy in the Hydrogen atom:

$$E_n = -\frac{m e^4}{2\hbar^2 n^2}. \tag{10.36}$$

Before mentioning the fundamental problems with the electronic configuration, we will discuss some basic properties of the Laguerre polynomials and of the Hydrogen atom eigenfunctions.

10.3 Properties of the Hydrogen-Atom Eigenfunctions

With $R_{nl}(r)$ we complete the solution to the Schrödinger Eq. (10.4) of the Hydrogen atom. The constant A_n in the radial function of (10.34) is determined from the normalization condition for the eigenfunctions $\varphi_{nl}^{m_z}(r,\theta,\phi) = R_{nl}(r)Y_l^{m_z}(\theta,\phi)$, i.e. from

$$\begin{aligned}\int |\varphi_{nl}^{m_z}(r,\theta,\phi)|^2 r^2 \sin\theta dr d\theta d\phi &= \int |R_{nl}(r)|^2 r^2 dr \int |Y_l^{m_z}(\theta,\phi)|^2 \sin\theta d\theta d\phi.\\ &= 1\end{aligned} \tag{10.37}$$

Since the spherical harmonics are normalized functions, the normalization condition reduces to

$$\int |R_{nl}(r)|^2 r^2 dr = 1. \tag{10.38}$$

If we use the representation of the radial function in (10.34) the normalization condition becomes

$$|A_n|^2 \frac{1}{8\alpha^3} \int \rho^2 \rho^{2l} e^{-\rho} |L_{n-l-1}^{2l+1}(\rho)|^2 d\rho = 1. \tag{10.39}$$

Here, we have used the relation $\rho = 2\alpha r$ to write the integrand in terms of the variable ρ only. To determine the constant A_n we need to evaluate the integral

$$\frac{1}{8\alpha^3} \int \rho^{2l+2} e^{-\rho} |L_{n-l-1}^{2l+1}(\rho)|^2 d\rho. \tag{10.40}$$

In the appendix B, it is shown that two important relations of the Laguerre polynomials are:

$$\int_0^\infty x^k e^{-x} L_\nu^k(x) L_{\nu'}^k(x) dx = \delta_{\nu,\nu'} \frac{(\nu+k)!}{\nu!} \tag{10.41}$$

and

$$\int_0^\infty x^{k+1} e^{-x} L_\nu^k(x) L_\nu^k(x) dx = (k+2\nu+1)\frac{(\nu+k)!}{\nu!}. \tag{10.42}$$

If we make $k = 2l+1$ and $\nu = nl - 1$, the last equation takes the form

$$\int_0^\infty x^{2l+2} e^{-x} L_{n-l-1}^{2l+1}(x) L_{n-l-1}^{2l+1}(x) dx = 2n \frac{(n+l)!}{(n-l-1)!}. \tag{10.43}$$

This integral is precisely what we have in the normalization condition (10.40). Consequently

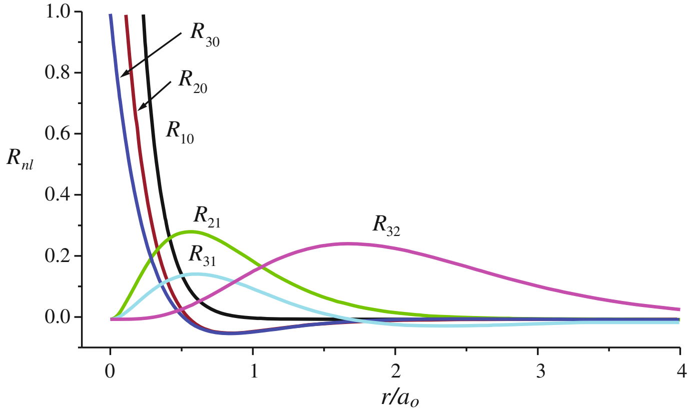


Fig. 10.1 The radial eigenfunctions R_{nl} as functions of r (measured in units of a_o), for n =1, 2 and 3 and all possible values of $l < n$

$$|A_n|^2 \frac{1}{8\alpha^3} 2n \frac{(n+l)!}{(n-l-1)!} = 1. \tag{10.44}$$

If we recall that $Z = \alpha a_o n$, the normalization constant of the radial function becomes

$$A_n = \left(\frac{Z}{a_o}\right)^{3/2} \frac{2}{n^2} \sqrt{\frac{(n-l-1)!}{(n+l)!}}, \tag{10.45}$$

and the radial part of the Hydrogen atom eigenfunction takes the form

$$R_{nl}(r) = \left(\frac{Z}{a_o}\right)^{l+3/2} \left(\frac{2}{n}\right)^{l+1} \frac{1}{n} \sqrt{\frac{(n-l-1)!}{(n+l)!}} r^l e^{-Zr/a_o n} L_{n-l-1}^{2l+1} \left(\frac{2Zr}{a_o n}\right). \tag{10.46}$$

In Fig. 10.1 we plot $R_{nl}(r)$ as a function of r/a_o. With these results we have well defined the eigenvalues E_n and the eigenfunctions $\varphi_{nl}^{m_z}(r, \theta, \phi)$ of the Hydrogen-like atoms. In a compact Dirac notation the eigenfunctions $\varphi_{nl}^{m_z}(r, \theta, \phi)$, defined by the quantum numbers n, l and m_z, are represented by the kets $|nlm_z\rangle$. We will see in the next section that, in order to describe the electronic configuration, the central potential model for the Hydrogen atom needs to consider the existence of an important electron's characteristic: the spin. Moreover, to explain the experimental observations in the presence of magnetic fields, we will extend the analysis to include the spin-field interactions.

A physical variable of interest, that can easily be calculated, is the expected value of the electron radius in a Hydrogen-like atom. At the end of this chapter we will show that the average radius of the electron orbit, in these atoms, is given by

$$\langle r\rangle = \frac{a_o}{2Z}\left(3n^2 - l(l+1)\right). \tag{10.47}$$

This radius increases with n. With this formula, we can evaluate expected radii for different electronic states. The valence electrons of the Hydrogen-like atoms with large atomic number Z have also large n' s. The orbit radii for these electrons depend on n and l but not on the magnetic quantum number m_z. The states $|nlm_z\rangle$ with the same n are degenerate. In the Hydrogen atom the electron orbit radius in the ground state (with the lowest energy) is

$$\langle r\rangle = \frac{3a_o}{2}. \tag{10.48}$$

Of the order of the Bohr radius a_o.

10.3.1 The Electronic Configuration of Hydrogen-Like Atoms

To conclude this topic, it is important to notice that the application of the order relations

$$|m_z| \le l \le n-1, \tag{10.49}$$

to determine the electronic configuration in Hydrogen-like atoms, is essential but not sufficient to explain the sequence 2, 8, 18, 32, ... that the experiments indicate for the number of electrons in the energy levels $E_1, E_2, E_3, \ldots$ According to (10.49), given an energy level E_n, we can have for the same energy, states with $l = 0, 1, 2, \ldots, n-1$ and for each value of l, $2l+1$ possible values for m_z. This, as shown in Table 10.1, predicts (see the fifth column) that the number of degenerate states is $\sum_{l=0}^{n-1}(2l+1) = n^2$. This number differs by a factor of 2 of the correct numbers suggested in the periodic chart of the elements, shown in the last column of Table 10.1. To explain this difference it was necessary to recognize, and later to discover, the existence of a new electron's physical property: the spin. Electrons, as protons and other quantum particles, besides their charge and mass possess an intrinsic angular momentum called spin, which manifest in the presence of other particles with spin and in the presence of magnetic fields. As will be seen with more detail in the next chapter, the z component S_z of the electron's spin, $\widehat{\mathbf{S}}$, has only two values: $\hbar/2$ and $-\hbar/2$. Taking into account the spin, the number of degenerate states gets then multiplied by a factor of 2, as indicated in the sixth column of Table 10.1. Indeed, W. E. Pauli,[4] in 1925, shortly before the final formulation of the quantum theory, showed that, in order to explain the periodic chart of elements, the orbiting electrons must be described by four quantum numbers plus an *exclusion principle* that forbids the possibility for two electrons, of the same atom, to have the same quantum numbers.

It is common to represent the orbital electron states $\varphi_{nl}^{m_z}\chi_{ms} = |nlm_z m_s\rangle$, where χ_{ms} corresponds the spin state, simply as $ns, np, nd, nf, \ldots$ when $l = 0, 1, 2, 3, \ldots,$

---

[4] W. Pauli , *Z. Physik* **31**, 765 (1925).

Table 10.1 States and number of electrons *without* and *with* spin for $n = 1, 2$ and 3

n	$l\ (\leq n-1)$	$m_z\ (= -l, \ldots l)$	$\lvert n, l, m_z\rangle$	electrons without spin	electrons with spin
1	0	0	$\lvert 1, 0, 0\rangle$	1	2
2	0	0	$\lvert 2, 0, 0\rangle$	4	8
	1	−1	$\lvert 2, 1, -1\rangle$		
		0	$\lvert 2, 1, 0\rangle$		
		1	$\lvert 2, 1, 1\rangle$		
3	0	0	$\lvert 3, 0, 0\rangle$	9	18
	1	−1	$\lvert 3, 1, -1\rangle$		
		0	$\lvert 3, 1, 0\rangle$		
		1	$\lvert 3, 1, 1\rangle$		
	2	−2	$\lvert 3, 2, -2\rangle$		
		−1	$\lvert 3, 2, -1\rangle$		
		0	$\lvert 3, 2, 0\rangle$		
		1	$\lvert 3, 2, 1\rangle$		
		2	$\lvert 3, 2, 2\rangle$		

respectively. This notation with a superindex k to indicate the number of electrons in the state nl, is used to represent the electronic configuration of atoms. For example, for Oxygen the electronic configuration is $1s^2 2s^2 2d^4$, for sodium with 11 electrons the electronic configuration is $1s^2 2s^2 2d^6 3s$.

10.4 The Hydrogen Atom in a Magnetic Field

At the end of the nineteenth century, continuing the analysis of the emitted light by atoms, Pieter Zeeman performed a series of studies on the effect of the magnetic field and observed the splitting of spectral lines as predicted by Lorentz, based on classical arguments.[5] Lorentz showed that in an atom (conceived as an oscillator) the relation between the frequencies ν and ν_o (with and without field, respectively), is

$$\nu^2 = \pm \frac{eB}{2\pi m}\nu + \nu_o^2 \quad \text{and} \quad \nu \cong \nu_o \pm \frac{eB}{4\pi m}. \tag{10.50}$$

The signs ± in the first equation refer to clockwise or anti-clockwise motion. This result implies that the emission lines split into three, with frequencies w_{21}, $w_{21} - 2\delta\omega$ and $w_{21} + 2\delta\omega$, as shown in Fig. 10.2. Pieter Zeeman found splittings into three and more lines. The attempts to explain those observations, 20 years later, fed the controversy that preceded the correct explanation of that problem that came through the final formulation of the quantum theory. Some of these phenomena known as

---

[5] Hendrik Antoon Lorentz and Pieter Zeeman shared the Nobel Prize in physics in 1902.

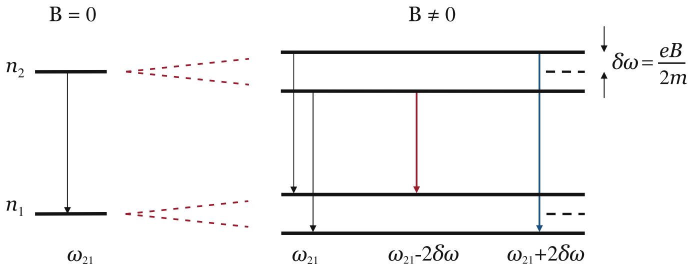


Fig. 10.2 The classical Lorentz model predicts that, in the presence of a magnetic field, the emission line splits into three emission lines

the *normal Zeeman effect*, could be explained with semi-classical arguments; others known as the *anomalous Zeeman effect* required the spin-field interaction.

When the Hydrogen atom or any other charged system is in the presence of a magnetic field, the charge dynamics is modified by the magnetic component of the Lorentz force $q\big(\mathbf{E} + (1/c)\mathbf{v} \times \mathbf{B}\big)$. It is known from the classical description of particles in a magnetic field $\mathbf{B}$ that the linear momentum $\mathbf{p}$ transforms into $\mathbf{p}+q\mathbf{A}/c$, i.e.

$$\mathbf{p} \quad \rightarrow \quad \mathbf{p} + \frac{q}{c}\mathbf{A}, \tag{10.51}$$

where $\mathbf{A}$ is the vector potential, such that $\nabla \times \mathbf{A} = \mathbf{B}$. The Hamiltonian for a charged particle changes in the presence of an electromagnetic field and takes the form (we assume that the scalar electric potential is zero)

$$\widehat{H} = \frac{1}{2m}\big(\widehat{\mathbf{p}} + \frac{q}{c}\mathbf{A}\big)^2 + V(\mathbf{r}). \tag{10.52}$$

Before specializing this discussion to the Hydrogen atom problem, we shall briefly refer to the general consequences that the presence of a magnetic field has on the kinetic term of the Schrödinger equation, regardless of the potential function $V(\mathbf{r})$. To visualize the effect and the physical sense we develop the kinetic energy term as follows

$$\frac{1}{2m}\big(\widehat{\mathbf{p}} + \frac{q}{c}\mathbf{A}\big)^2 = \frac{1}{2m}\widehat{\mathbf{p}}^2 + \frac{q}{mc}\mathbf{A}\cdot\mathbf{p} - \frac{iq\hbar}{2mc}\nabla\cdot\mathbf{A} + \frac{q^2}{2mc^2}\mathbf{A}^2. \tag{10.53}$$

If we choose the z axis along the field $\mathbf{B}$, we write the magnetic field as $\mathbf{B} = B\widehat{\mathbf{z}}$. This equation can be simplified, without changing the physical description, using equivalent representations or gauge transformations for the potential vector $\mathbf{A}$. One of the possible gauges for the vector potential is[6]

---

[6] This is known as the symmetric Coulomb gauge.

$$\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r} = \frac{B}{2}(-y\widehat{\mathbf{x}} + x\widehat{\mathbf{y}}) = A_x\widehat{\mathbf{x}} + A_y\widehat{\mathbf{y}}, \tag{10.54}$$

which satisfies the equations

$$\nabla \times \mathbf{A} = \mathbf{B} \quad \text{and} \quad \nabla \cdot \mathbf{A} = 0. \tag{10.55}$$

Therefore

$$\mathbf{A} \cdot \widehat{\mathbf{p}} = \frac{1}{2}\mathbf{B} \cdot (\mathbf{r} \times \widehat{\mathbf{p}}) = \frac{1}{2}\mathbf{B} \cdot \widehat{\mathbf{L}} \quad \text{and} \quad \mathbf{A}^2 = \frac{B^2}{4}\left(x^2 + y^2\right). \tag{10.56}$$

Making use of these relations, the Hamiltonian takes the form

$$\widehat{H} = \frac{1}{2m}\widehat{\mathbf{p}}^2 + \frac{q}{2mc}\mathbf{B} \cdot \widehat{\mathbf{L}} + \frac{q^2}{2mc^2}\frac{B^2}{4}\left(x^2 + y^2\right) + V(\mathbf{r}). \tag{10.57}$$

This Hamiltonian has two new terms: the *momentum-field interaction*, responsible of the normal Zeeman effect and the magnetic confining harmonic potential, which can be neglected because in the Hydrogen atom it is relatively small. If we study the dynamics of charge carriers in fields, one should take into account the harmonic potential. It is essential to explain important phenomena like the quantum Hall effect. The Schrödinger equation for the Hydrogen atom in a magnetic field is then

$$-\frac{\hbar^2}{2m}\nabla^2\varphi(\mathbf{r}) - \frac{e}{2mc}\mathbf{B} \cdot \widehat{\mathbf{L}}\,\varphi(\mathbf{r}) + V(r)\,\varphi(\mathbf{r}) = E\,\varphi(\mathbf{r}), \tag{10.58}$$

or just

$$-\frac{\hbar^2}{2m}\nabla^2\varphi(\mathbf{r}) - \frac{e}{2mc}B\widehat{L}_z\,\varphi(\mathbf{r}) + V(r)\,\varphi(\mathbf{r}) = E\,\varphi(\mathbf{r}). \tag{10.59}$$

As will be seen in the next section, the new terms in the Hamiltonian will be helpful to explain the normal Zeeman effect but not the anomalous Zeeman effect. We leave the discussion on the anomalous Zeeman effect for the next chapter, where we will study the spin. It is worth noticing that since the angular momentum components $\widehat{L}_i$ do not commute with $\widehat{\mathbf{L}}$, the angular momentum $\widehat{\mathbf{L}}$ will not commute with the Hamiltonian. Therefore, the angular momentum $\widehat{\mathbf{L}}$ in the presence of an external magnetic field, is not a conserved quantity. Now the z component $\widehat{L}_z$ is the conserved quantity. It commutes with $\widehat{L^2}$ and also with functions of r.

10.5 The Normal Zeeman Effect

If the orbital magnetic moment $\widehat{\mu}_{\mathrm{L}}$, is defined as

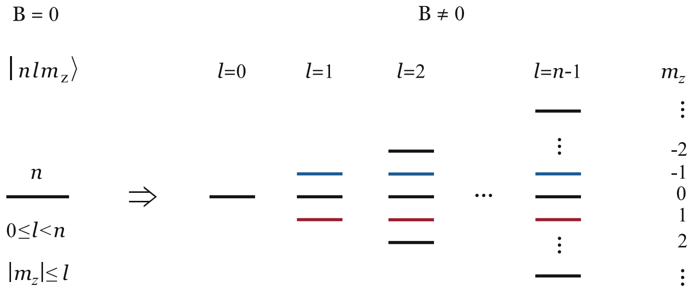


Fig. 10.3 The magnetic field removes the degeneracy in m_z, which is equal to $2l + 1$ for each allowed value of l. Since the degeneracy in the angular momentum remains, states with different l but the same m_z have the same energy. For example, $|n11\rangle$ and $|n21\rangle$ have the same energy

$$\widehat{\mu}_{\mathrm{L}} = g_{\mathrm{L}} \frac{e\hbar}{2mc} \widehat{\mathbf{L}} = g_{\mathrm{L}} \mu_B \widehat{\mathbf{L}}, \tag{10.60}$$

where μ_B is the Bohr's magneton $\mu_B = 5.7884\ 10^{-5}\,\mathrm{eV/T}$ and g_{L} (=1) is the orbital gyromagnetic ratio, the Schrödinger Eq. (10.59) can be written as

$$-\frac{\hbar^2}{2m}\nabla^2\varphi(\mathbf{r}) - \frac{\mathbf{B}\cdot\widehat{\mu}_{\mathrm{L}}}{\hbar}\varphi(\mathbf{r}) + V(r)\,\varphi(\mathbf{r}) = E\,\varphi(\mathbf{r}), \tag{10.61}$$

where one is tempted to interpret the second term as part of the potential energy.

Since the eigenvalue of the operator $\widehat{L}_z$ is $m_z\hbar$, the Schrödinger equation of the Hydrogen atom in a magnetic field can be written in the form

$$-\frac{\hbar^2}{2m}\nabla^2\varphi(\mathbf{r}) + V(r)\varphi(\mathbf{r}) = \left(E + \mu_B B m_z\right)\varphi(\mathbf{r}) \tag{10.62}$$

which is similar to the Schrödinger equation of the Hydrogen atom in the absence of a magnetic field with the sum

$$E + \mu_B B m_z \tag{10.63}$$

playing the role of the energy E of that equation. Consequently

$$E_n + \mu_B B m_z = -\frac{Z^2 m e^4}{2\hbar^2 n^2}. \tag{10.64}$$

For fixed values of n and l, we have now $2l + 1$ different values for the energy.

Therefore, the effect of the magnetic field in the linear approximation, is to remove the degeneracy associated with the magnetic quantum number m_z. This is the normal Zeeman effect.

In summary, the electron energies in the Hydrogen atom (or in the Hydrogen-like atoms) in the presence of a magnetic field are

$$E_n = -\frac{Z^2 m e^4}{2\hbar^2 n^2} - \mu_B B m_z, \quad \text{with} \quad m_z = 0, \pm 1, \pm 2, \ldots, \pm l. \tag{10.65}$$

The interaction $\widehat{\mathbf{L}} \cdot \mathbf{B}$ splits the energies as shown in Fig. 10.3 and explains part of the experimental observations and the splitting of the emission lines in the presence of magnetic fields.

10.6 Solved Problems

Exercise 35 Show that the expected value of the election radius in the Hydrogen atoms is

$$\langle r \rangle = \frac{a_o}{2Z}\big(3n^2 - l(l+1)\big). \tag{10.66}$$

Solution The expected value of r, in an eigenstate $\varphi_{nl}^{m_z} = R_{nl}(r) Y_l^{m_z}(\theta, \phi)$ of a Hydrogen atom, is given by the relation

$$\langle r \rangle = \int\int\int \big(\varphi_{nl}^{m_z}(r, \theta, \phi)\big)^* r \varphi_{nl}^{m_z}(r, \theta, \phi) r^2 \sin\theta dr d\theta d\phi. \tag{10.67}$$

The integral of the angular part is simple and gives a factor 1, we have then to evaluate the integral

$$\langle r \rangle = \int_0^\infty r^3 [R_{nl}(r)]^2 dr = |A_n|^2 \int r^3 \rho^{2l} e^{-\rho} \left[L_{n-l-1}^{2l+1}(\rho)\right]^2 dr, \tag{10.68}$$

which is similar to the integral in Eq. (10.39). In this integral we have a different power of r. If we substitute r by $\rho/2\alpha$, we have

$$\langle r \rangle = |A_n|^2 \frac{1}{(2\alpha)^4} \int \rho^{2l+3} e^{-\rho} \left[L_{n-l-1}^{2l+1}(\rho)\right]^2 d\rho. \tag{10.69}$$

To evaluate this integral we use the recurrence relation (B.11) written in the form

$$x L_\nu^k(x) = (k + 2\nu + 1) L_\nu^k(x) - (\nu + 1) L_{\nu+1}^k(x) - (\nu + k) L_{\nu-1}^k(x). \tag{10.70}$$

The integral of $(x L_\nu^k(x))^2$ multiplied by $x^k e^{-x}$, with $k = 2l + 1$ and $\nu = n - l - 1$, is precisely the integral that we need to evaluate $\langle r \rangle$. If we square the right hand side of (10.70), multiply by $x^k e^{-x}$, integrate and use the orthogonality and normalization properties of the Laguerre polynomials that were derived in the Appendix A, we find, after some algebra, the following expression

$$\int x^{k+2}e^{-x}[L_{\nu}^{k}(x)]^2dx = \frac{(\nu+k)!}{\nu!}(6\nu^2+6k\nu+6\nu+k^2+3k+2). \quad (10.71)$$

This equation with $k=2l+1$ and $\nu=n-l-1$ reduces to the expected result.

Exercise 36 Show that if m_N is the mass of the nucleus and m_e the mass of the electron in the Hydrogen atom, the approximation of the nucleus being at rest and the assumption of $m_N \gg m_e$ are equivalent.

Solution If the nucleus is at rest, the Schrödinger equation of the Hydrogen atom

$$\left(\frac{\widehat{\mathbf{p}}_N^2}{2m_N}+\frac{\widehat{\mathbf{p}}_e^2}{2m_e}+V(|\mathbf{r}_N-\mathbf{r}_e|)\right)\psi(\mathbf{r}_N,\mathbf{r}_e) = E\psi(\mathbf{r}_N,\mathbf{r}_e), \quad (10.72)$$

is just

$$\left(\frac{\widehat{\mathbf{p}}_e^2}{2m_e}+V(|\mathbf{r}_e|)\right)\psi(\mathbf{r}_e) = E\psi(\mathbf{r}_e), \quad (10.73)$$

with the origin of the coordinate system at the position of the atomic nucleus. If the nucleus is not at rest, we can write the two-particle Hamiltonian in terms of the center of mass R and coordinate r defined by

$$\mathbf{R} = \frac{m_N\mathbf{r}_N+m_e\mathbf{r}_e}{m_N+m_e} \quad \text{and} \quad \mathbf{r} = \mathbf{r}_e-\mathbf{r}_N, \quad (10.74)$$

to obtain the Schrödinger equation

$$\left(\frac{\widehat{\mathbf{P}}^2}{2M}+\frac{\widehat{\mathbf{p}}^2}{2\mu}+V(r)\right)\psi(\mathbf{R},\mathbf{r}) = E\psi(\mathbf{R},\mathbf{r}), \quad (10.75)$$

with $M = m_N+m_e$ and $\mu = m_Nm_e/M$. The first term describes the kinetic energy of a quasi-particle with mass M and momentum $\mathbf{P} = \dot{\mathbf{R}}$ while the second the kinetic energy of a quasi-particle with mass μ and momentum $\mathbf{p} = \dot{\mathbf{r}}$. If we factor the two-particle wave function as

$$\psi(\mathbf{R},\mathbf{r}) = \Psi(\mathbf{R})\phi(\mathbf{r}) \quad (10.76)$$

we end up with the following equations

$$\frac{\widehat{\mathbf{P}}^2}{2M}\Psi(\mathbf{R}) = E_M\Psi(\mathbf{R}), \quad (10.77)$$

and

$$\left(\frac{\widehat{\mathbf{p}}^2}{2\mu}+V(r)\right)\phi(\mathbf{r})=\epsilon\phi(\mathbf{r}), \tag{10.78}$$

with $\epsilon = E - E_M$. The first of these equations describes the motion of the center of mass as a free particle with kinetic energy E_M. One can always choose the zero of the energy at any value without consequence on the system dynamics. We can choose the zero at E_M. In that case, the center of mass is at rest and we are left with the Schrödinger equation

$$\left(\frac{\widehat{\mathbf{p}}^2}{2\mu}+V(r)\right)\phi(\mathbf{r})=E\phi(\mathbf{r}). \tag{10.79}$$

It is easy now to see that, in the particular case of $m_N \gg m_e$, the relative mass $\mu \to m_e$, hence

$$\left(\frac{\widehat{\mathbf{p}}^2}{2m_e}+V(r)\right)\phi(\mathbf{r})=E\phi(\mathbf{r}), \tag{10.80}$$

which coincides with the Schrödinger Eq. (10.73) for the nucleus at rest.

10.7 Problems

1. Show that the two-particle Hamiltonian

$$\widehat{\mathbf{H}}=\frac{\widehat{\mathbf{p}}_1^2}{2m_1}+\frac{\widehat{\mathbf{p}}_2^2}{2m_2}+V(|\mathbf{r}_2-\mathbf{r}_1|), \tag{10.81}$$

written in terms of the center of mass R an relative coordinate r:

$$\mathbf{R}=\frac{m_1\mathbf{r}_1+m_2\mathbf{r}_2}{m_1+m_2} \qquad \text{and} \qquad \mathbf{r}=\mathbf{r}_2-\mathbf{r}_1, \tag{10.82}$$

transforms into

$$\widehat{\mathbf{H}}=\frac{\widehat{\mathbf{P}}^2}{2M}+\frac{\widehat{\mathbf{p}}^2}{2\mu}+V(r), \tag{10.83}$$

with $M = m_1 + m_2$ and $\mu = m_1m_2/M$.

2. What is the difference between the reduced mass μ and the mass of the electron in the Hydrogen atom?
3. Show that

$$\int x^{k+2}e^{-x}[L_\nu^k(x)]^2dx=\frac{(\nu+k)!}{\nu!}(6\nu^2+6k\nu+6\nu+k^2+3k+2). \tag{10.84}$$

4. Show that

$$-\frac{\hbar^2}{2m}\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right)=-\frac{\hbar^2}{2m}\frac{1}{r}\frac{d^2}{dr^2}r, \tag{10.85}$$

and obtain the Eq. (10.8).

5. Consider the change of variable $\rho = 2\alpha r$, the definitions

$$\alpha^2=\frac{2m|E|}{\hbar^2} \qquad \text{and} \qquad a_o=\frac{\hbar^2}{me^2}, \tag{10.86}$$

   and deduce the Eq. (10.12).
6. Show that the products αr and αa_o are dimensionless.
7. Show that if $u(\rho) = \rho^{l+1}e^{-\rho/2}Q(\rho)$, the radial equation becomes the Eq. (10.26).
8. Show that the number of degenerate states in the energy level E_n is n^2, and $2n^2$ when the two spin states are taken into account.
9. Shows in detail that the normalization constant of the radial function R_{nl} is

$$A_n=\frac{Z}{a_o}^{3/2}\frac{2}{n^2}\sqrt{\frac{(n-l-1)!}{(n+l)!}}, \tag{10.87}$$

   and show also that for the ground state of the Hydrogen atom is

$$A_1=\frac{2}{a_o^{3/2}}. \tag{10.88}$$

10. Show that

$$\left\langle\frac{1}{r}\right\rangle=\frac{Z}{a_o n^2} \quad \text{and} \quad \left\langle\frac{1}{r^2}\right\rangle=\frac{2Z^2}{a_o^2 n^3}\frac{1}{2l+1}. \tag{10.89}$$

11. When you have an operator $\widehat{F}(\lambda)$ with eigenvalue $f_n(\lambda)$ and eigenfunction $\psi_n(\lambda)$, the Feynman-Hellman theorem states that

$$\frac{\partial f_n(\lambda)}{\partial\lambda}=\langle\psi_n(\lambda)|\frac{\partial\widehat{F}(\lambda)}{\partial\lambda}|\psi_n(\lambda)\rangle. \tag{10.90}$$

   If the Hamiltonian of the Hydrogen atom is written as

$$\widehat{H}(\lambda)=\frac{\widehat{p}_r}{2m}+\frac{\hbar^2\lambda(\lambda+1)}{2mr^2}+V(r), \tag{10.91}$$

   with eigenvalue

$$E_n(\lambda)=-\frac{mZ^2e^4}{2\hbar^2(k+\lambda+1)^2}, \tag{10.92}$$

   use the Feynman-Hellman theorem to show that

$$\left\langle \frac{1}{r^2} \right\rangle_{\lambda=l} = \frac{2Z^2}{a_o^2 n^3} \frac{1}{2l+1}. \tag{10.93}$$

12. To explain the observed lines by P. Zeeman for the Hydrogen atom in the presence of a magnetic field **B**, Hendrik Lorentz proposed a theory where the electron of the Hydrogen atom oscillates with a frequency ν_o under a harmonic force $F_k = kr$ with $k = 4\pi^2 \nu_o^2 m$. Derive the relation

$$\nu^2 = \pm \frac{eB}{2\pi m} \nu + \nu_o^2, \tag{10.94}$$

   when a magnetic field **B** is established. Assume that the Lorentz force and the harmonic force are responsible for the circular motion with electron acceleration v^2/r. Discuss the conditions to obtain

$$\nu = \nu_o \pm \frac{eB}{4\pi m}. \tag{10.95}$$

13. If a Hydrogen atom is in a field $B = 1$ T, mention all the possible transitions from level $n = 2$ and the frequencies of the emission lines.

Chapter 11
Spin and the Pauli Equation

11.1 Introduction

The electronic configuration and the explanation of the emission lines of atoms in a magnetic field, observed by Zeeman, were the major topics of the emerging quantum physics and of controversy around 1920. Despite the success to explain the Balmer and Rydberg series and the efforts by physicists of the stature of Arnold Sommerfeld to formalize the Bohr model, the old semi-classical quantum theory was unable to explain the emission lines of atoms in magnetic fields as well as other atomic problems. The resistance of many physicists to accept a theory of *ad-hoc* postulates and models kept growing. Among the dissenters was Otto Stern, who had been an assistant to Einstein in Zurich. Stern had experience in producing beams of atoms. In 1921 he decided to study the dynamics of atoms in the presence of an external field. His objective was not to find evidences of the spin, but to visualize the space quantization, as was named in those days, the magnetic moment quantization, which is essentially the same as the quantization of angular momentum.

It is known from the electromagnetic theory that a particle with magnetic moment μ, in the presence of an inhomogeneous magnetic field, feels a force equal to the gradient of its potential energy $-\mu \cdot \mathbf{B}$, i.e.

$$\mathbf{F} = \nabla(\mu \cdot \mathbf{B}). \tag{11.1}$$

If the field is along the $\hat{\mathbf{z}}$-axis, and $\mu = \mu_B \widehat{\mathbf{L}}$, the force on the atoms is:

$$F_z = \mu_B \widehat{L}_z \frac{\partial B}{\partial z}. \tag{11.2}$$

Consequently, the paths which the atoms could follow were as many as the possible values of the quantum number m_z. In the previous chapter it was shown that $m_z = -l, -l+1, \ldots, l-1, l$. For this reason one can expect to observe $2l+1$ traces. The arrangement in the Stern–Gerlach experiment, as shown in Fig. 11.1, was such that

P. Pereyra, *Fundamentals of Quantum Physics*, Undergraduate Lecture Notes in Physics,
DOI: 10.1007/978-3-642-29378-8_11,

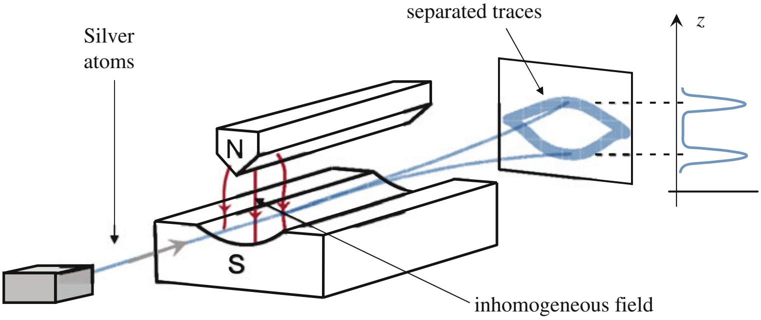


Fig. 11.1 In the Stern–Gerlach experiment the magnetic force of an inhomogeneous magnetic field deflects the silver atoms moving across. The two traces observed, instead of the predicted three by the theoretical arguments of those years, led to the discovery of the spin

the direction of **B** did not change much but its magnitude changed from one point to another.

After many attempts, they found that only two traces appeared.[1] This result introduced more confusion but also new arguments to justify them. It was not until 1926, when Goudsmit and Uhlenbeck[2] reinterpreted the Stern–Gerlach experiment and introduced the electron spin as an *intrinsic angular momentum*, that the experimental results became clear. The electron's spin $\widehat{\mathbf{S}}$ is a physical quantity with two components along the magnetic field. Like the orbital magnetic momentum, $\widehat{\mu}_{\mathrm{L}} = g_{\mathrm{L}}\frac{e\hbar}{2mc}\widehat{\mathbf{L}}$, the intrinsic magnetic momentum μ_e is proportional to the intrinsic angular momentum $\widehat{\mathbf{S}}$. In 1921, Compton was the first to suggest the possibility of a half-integer intrinsic angular momentum for the electron spin. It seems also that in 1925 Kronig suggested to Pauli the idea of an electron spin. It is known that Pauli postulated, later in 1925, the existence of a two-value quantum property without classical analog[3] and an exclusion principle to forbid the possibility that two electrons, in the same atom, could have the same quantum numbers.

If the electron possesses, besides its orbital momentum $\widehat{\mu}_{\mathrm{L}}$, an intrinsic magnetic momentum defined as

$$\mu_e = g_{\mathrm{S}}\mu_B\widehat{\mathbf{S}}, \tag{11.3}$$

one can think of a potential energy

$$-\frac{\mathbf{B}\cdot\mu_e}{\hbar}, \tag{11.4}$$

---

[1] An interesting story of the circumstances with regard to this experiment can be found in B. Friedrich and D. Herschbach, *Physics Today* December 2003 pg. 53.

[2] S. Goudsmit and G. Uhlenbeck, *Nature* **117** 264 (1926).

[3] W.E. Pauli, *Z. Physic* **31**, 765 (1925) and W.E. Pauli Nobel Lecture (1946).

that should be added to the Hamiltonian of the Schrödinger equation. Therefore, we have to consider, for atoms in a magnetic field, the equation

$$\left(\frac{-\hbar^2}{2m}\nabla^2+g_{\mathrm{L}}\frac{\mu_B}{\hbar}\mathbf{B}\cdot\widehat{\mathbf{L}}+g_{\mathrm{S}}\frac{\mu_B}{\hbar}\mathbf{B}\cdot\widehat{\mathbf{S}}+\frac{q^2}{c^2}\frac{B^2}{4}\left(x^2+y^2\right)+V(\mathbf{r})\right)\phi(\mathbf{r},s)=E\phi(\mathbf{r},s). \tag{11.5}$$

The inclusion of a term in the Schrödinger equation that depends on the intrinsic angular momentum $\widehat{\mathbf{S}}$ adds to the spatial degrees of freedom, the spin degree of freedom. This fact, explicitly indicated in the wave function $\phi(\mathbf{r},\mathbf{s})$, corresponds to the product

$$\phi(\mathbf{r},s)=\varphi(\mathbf{r})\chi(s), \tag{11.6}$$

in which $\varphi(\mathbf{r})$ is the usual wave function and $\chi(s)$ the spin-dependent vector, called *spinor*, which depends on the spin. The spinor is in general a linear combination of two states: $\chi_\uparrow$ and $\chi_\downarrow$ that correspond to the spin projections, parallel and antiparallel to the field. These eigenfunctions of $\widehat{S}_z$ are usually represented with vectors (spinors)

$$\chi_\uparrow=\begin{pmatrix}1\\0\end{pmatrix}, \qquad \chi_\downarrow=\begin{pmatrix}0\\1\end{pmatrix}, \tag{11.7}$$

and satisfy the eigenvalue equations

$$\widehat{S}_z\begin{pmatrix}1\\0\end{pmatrix}=S_z\begin{pmatrix}1\\0\end{pmatrix}, \quad \widehat{S}_z\begin{pmatrix}0\\1\end{pmatrix}=-S_z\begin{pmatrix}0\\1\end{pmatrix}. \tag{11.8}$$

If the spin is an angular momentum, it must satisfy commutation relations like those of the angular momentum $\widehat{\mathbf{L}}$ and its components. This means that

$$[\widehat{S}_i,\widehat{S}_j]=i\hbar\epsilon_{ijk}\widehat{S}_k, \tag{11.9}$$

and

$$[\widehat{\mathbf{S}},\widehat{S}_i]=i\hbar\left(\widehat{S}_j\hat{\mathbf{x}}_k-\widehat{S}_k\hat{\mathbf{x}}_j\right). \tag{11.10}$$

We will come back into the Schrödinger equation for particles with spin after a brief discussion of the spin properties. In the next section, we will use the commutation relations (11.9) to show that the spin quantum number s, compatible with these relations, can be an integer or a half-integer number. We will obtain the eigenvalues of $\widehat{S}_z$ and $\widehat{S}^2$, and we will also derive their matrix representations.

11.2 Spin Eigenvalues and Matrix Representations

11.2.1 Eigenvalues of $\widehat{S}_z$ and $\widehat{S}^2$

In Chap. 9 we obtained the eigenvalues and eigenfunctions of $\widehat{L}_z$ and $\widehat{L}^2$ using their explicit differential operators. Here we will derive the eigenvalues and the eigenfunctions of $\widehat{S}_z$ and $\widehat{S}^2$ based on their commutation relations. To this purpose we need to introduce the rising and lowering operators $\widehat{S}_+$ and $\widehat{S}_-$, analogous to $\widehat{L}_+$ and $\widehat{L}_-$, defined in Chap. 9. We will obtain some identities that will lead to several interesting results. Let us then define the operators:

$$\widehat{S}_+ = \widehat{S}_x + i\widehat{S}_y \quad \text{and} \quad \widehat{S}_- = \widehat{S}_x - i\widehat{S}_y. \tag{11.11}$$

It is easy show that

$$\left[\widehat{S}_\mp, \widehat{S}_z\right] = \ \pm\hbar\widehat{S}_\mp, \tag{11.12}$$

and

$$\left[\widehat{S}_+, \widehat{S}_-\right] = 2\hbar\widehat{S}_z. \tag{11.13}$$

The operators $\widehat{S}_-$ and $\widehat{S}_+$, like $\widehat{L}_+$ and $\widehat{L}_-$, are rising and lowering operators, i.e. operators that, acting on the eigenfunctions of $\widehat{S}_z$ result in eigenfunctions of the neighbor states, corresponding to quantum numbers increased or decreased by 1. Indeed, suppose we have the eigenvalue equations

$$\widehat{S}^2\,\chi = a\,\chi \quad \text{and} \quad \widehat{S}_z\,\chi = b\,\chi. \tag{11.14}$$

If we act with the operator $\widehat{S}_+$ on the second of these equations we have

$$\widehat{S}_+\widehat{S}_z\chi = b\,\widehat{S}_+\chi, \tag{11.15}$$

using the relation (11.12), we can write the last equation as

$$\widehat{S}_z\,\widehat{S}_+\chi = (b+\hbar)\,\widehat{S}_+\chi. \tag{11.16}$$

This shows not only that $\widehat{S}_+\chi$ is an eigenfunction of $\widehat{S}_z$, but also that $\widehat{S}_+\chi$ is a state with an eigenvalue of $\widehat{S}_z$ increased by $\hbar$, i.e. $\widehat{S}_+$ is a rising operator. If we repeat n times the action of $\widehat{S}_+$, we will have the relation

$$\widehat{S}_z\,\widehat{S}_+^n\,\chi = (b+n\hbar)\,\widehat{S}_+^n\,\chi. \tag{11.17}$$

It is clear also that, if χ_M is the state with the maximum eigenvalue b_M, then

$$\widehat{S}_+\,\chi_M = 0. \tag{11.18}$$

Similarly, one can show that

$$\widehat{S}_z\,\widehat{S}_-\chi = (b-\hbar)\,\widehat{S}_-\chi, \tag{11.19}$$

with

$$\widehat{S}_z\,\widehat{S}_-^n\,\chi = (b-n\hbar)\,\widehat{S}_-^n\,\chi. \tag{11.20}$$

Again, if χ_m is the state of minimum eigenvalue b_m, we will have

$$S_-\,\chi_m = 0. \tag{11.21}$$

With these results and using the following relations

$$\widehat{S}_-\widehat{S}_+ = \widehat{S}^2 - \widehat{S}_z^2 - \hbar\widehat{S}_z. \tag{11.22}$$

$$\widehat{S}_+\widehat{S}_- = \widehat{S}^2 - \widehat{S}_z^2 + \hbar\widehat{S}_z. \tag{11.23}$$

we will show that the spin can be half-integer and, even more, we will obtain the eigenvalues of $\widehat{S}_z$ and $\widehat{S}^2$. Let us consider the operator $\widehat{S}_-\widehat{S}_+$ acting on the state χ_M. In that case, we have

$$\begin{aligned}\widehat{S}_-\widehat{S}_+\chi_M &= \big(\widehat{S}^2 - \widehat{S}_z^2 - \hbar\,\widehat{S}_z\big)\chi_M\\ &= \big(a - b_M^2 - \hbar\,b_M\big)\chi_M = 0.\end{aligned} \tag{11.24}$$

Similarly, if we consider the operator $\widehat{S}_+\widehat{S}_-$ acting on the state χ_m, we have

$$\begin{aligned}\widehat{S}_+\widehat{S}_-\chi_m &= \big(\widehat{S}^2 - \widehat{S}_z^2 + \hbar\,\widehat{S}_z\big)\chi_m\\ &= \big(a - b_m^2 + \hbar\,b_m\big)\chi_m = 0.\end{aligned} \tag{11.25}$$

These equations are satisfied if and only if

$$a - b_m^2 + \hbar\,b_m = 0 \quad\text{and}\quad a - b_M^2 - \hbar\,b_M = 0. \tag{11.26}$$

Now let us suppose that $b_M = b_m + n\hbar$. Using this and combining the last equations, we obtain, after some algebra, the following equation

$$2b_M\hbar(n+1) - (n+1)n\hbar^2 = 0 \quad\Rightarrow\quad b_M = \frac{n\hbar}{2}. \tag{11.27}$$

Substituting b_M in the second equation of (11.26) we end up with the eigenvalue

$$a = \hbar^2\frac{n}{2}\left(1+\frac{n}{2}\right), \tag{11.28}$$

of $\widehat{S}^2$. It is clear that the number n can be even or odd. If it is an even number, say $2l$, the eigenvalue of $\widehat{S}^2$ will be $\hbar^2 l\,(l+1)$, with $2l+1$ projections along the z-axis. if n is odd, for example $n=1$, the eigenvalue of $\widehat{S}^2$ is $\hbar^2 \frac{1}{2}\left(\frac{1}{2}+1\right)$, and the number of projections along the z-axis is $2(1/2)+1=2$. This is precisely the number found in the Stern–Gerlach experiment. This suggests that the quantum number that corresponds to the intrinsic electron angular momentum **S** is $s=1/2$. In this case

$$\widehat{S}^2\,\chi = \hbar^2 s(s+1)\,\chi = \frac{3}{4}\hbar^2\,\chi. \tag{11.29}$$

We can ask what are the eigenvalues of $\widehat{S}_z$? If $n=1$, the maximum and minimum eigenvalues are

$$b_M = \frac{\hbar}{2} \quad \text{and} \quad b_m = b_M - n\hbar = -\frac{\hbar}{2}, \tag{11.30}$$

respectively. Therefore:

$$\widehat{S}_z\,\chi_M = \frac{\hbar}{2}\chi_M \quad \Rightarrow \quad \chi_M = \chi_\uparrow, \tag{11.31}$$

and

$$\widehat{S}_z\,\chi_m = -\frac{\hbar}{2}\chi_m \quad \Rightarrow \quad \chi_m = \chi_\downarrow. \tag{11.32}$$

From here on, the eigenvalues of $\widehat{S}_z$ will be denoted as $m_s\hbar$. It is clear that $m_s=\pm 1/2$. m_s is the fourth quantum number that was missing to solve properly the electronic configuration problem. If we represent the electronic states with the kets $|nlm_zm_s\rangle$, it is clear also that instead of the states $|nlm_z\rangle$ we actually have two states: $|nlm_z 1/2\rangle$ and $|nlm_z -1/2\rangle$. This explains that the number of degenerate states for each energy level is twice the predicted number n^2, in the absence of spin. It is important to observe that if $\mathbf{B}=0$, the spin-field interaction disappears from the Hamiltonian. This implies that the Hamiltonian in Chap. 10 was correct. Not knowing the existence of the spin up $\chi_\uparrow$ and the spin down $\chi_\downarrow$ states, the counting of the degenerate states per energy level was wrong. On the other hand, we learn here that, when the hydrogen and hydrogen-like atoms are in the presence of a magnetic field, one has to include the spin-field interaction in order to explain the anomalous Zeeman effect.

11.2.2 Spin Representations and the Pauli Matrices

To obtain the matrix representations of the angular momentum in Chap. 9, we have first derived the rising and lowering formulas (9.78) and (9.79). These formulas with $\widehat{S}$ instead of $\widehat{L}$ are

$$\widehat{S}_+\,|sm_s\rangle = \hbar\sqrt{(s+m_s+1)(s-m_s)}\,|s, m_s+1\rangle, \tag{11.33}$$

$$\widehat{S}_-|sm_s\rangle = \hbar\sqrt{(s+m_s)(s-m_s+1)}\;|s, m_s-1\rangle, \tag{11.34}$$

and we can use them to obtain all representations of the spin operators. We leave as an exercise for the student the explicit evaluation of these representations. Here we will discuss this issue by using a different procedure, based on the commutation relations and the Hermitian character of the spin operators. We will begin with the matrix representation of $\widehat{S}_z$ in the basis of eigenfunctions $\chi_\uparrow$ and $\chi_\downarrow$. In this basis, we build the matrix representation S_z of $\widehat{S}_z$, with the matrix elements $\langle\chi_i|\widehat{S}_z|\chi_j\rangle$, i.e.

$$S_z = \begin{pmatrix} \langle\chi_\uparrow|\widehat{S}_z|\chi_\uparrow\rangle & \langle\chi_\uparrow|\widehat{S}_z|\chi_\downarrow\rangle \\ \langle\chi_\downarrow|\widehat{S}_z|\chi_\uparrow\rangle & \langle\chi_\downarrow|\widehat{S}_z|\chi_\downarrow\rangle \end{pmatrix}. \tag{11.35}$$

If we use the eigenvalues equations and the orthogonality of $\chi_\uparrow$ and $\chi_\downarrow$, we have:

$$S_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{11.36}$$

with

$$S_z^2 = \frac{\hbar^2}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{\hbar^2}{4} I_2. \tag{11.37}$$

From the commutation relations, which are the same as for the matrix representations, one can deduce other quite useful relations. For example, if we take the relation

$$S_z S_x - S_x S_z = i\hbar S_y, \tag{11.38}$$

and multiply from the left and right by S_z we have, after replacing S_z^2 by $\hbar^2 I_2/4$, the identity

$$\frac{\hbar^2}{4}\big(S_x S_z - S_z S_x\big) = i\hbar S_z S_y S_z, \tag{11.39}$$

that transforms into

$$-\frac{\hbar^2}{4} S_y = S_z S_y S_z. \tag{11.40}$$

Here we can multiply again from the right, by S_z. Simplifying factors, one has the anticommutator

$$S_y S_z + S_z S_y = 0. \tag{11.41}$$

In a similar way we obtain

$$S_z S_x + S_x S_z = 0, \tag{11.42}$$

$$S_x S_y + S_y S_x = 0. \tag{11.43}$$

These anticommutators are important relations. Combined with the commutators, we deduce the following

$$2S_i S_j = i\hbar\epsilon_{ijk} S_k. \tag{11.44}$$

With these relations we are ready to obtain the matrix representations of $\widehat{S}_x$ and $\widehat{S}_y$. We propose that

$$S_x = \begin{pmatrix} a_x & b_x \\ c_x & d_x \end{pmatrix} \quad \text{and} \quad S_y = \begin{pmatrix} a_y & b_y \\ c_y & d_y \end{pmatrix}, \tag{11.45}$$

with complex entries, in general. If we replace these matrices in (11.44), we find that

$$S_x = \begin{pmatrix} 0 & \frac{\hbar^2}{4}\frac{1}{c_x} \\ c_x & 0 \end{pmatrix} \quad \text{and} \quad S_y = \begin{pmatrix} 0 & -i\frac{\hbar^2}{4}\frac{1}{c_x} \\ ic_x & 0 \end{pmatrix}. \tag{11.46}$$

It is easy to verify that for these matrices to be Hermitian we require that $c_x = \hbar/2$. Therefore, the final forms of S_x and S_y are

$$S_x = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{with} \quad S_x^2 = \frac{\hbar^2}{4} I_2, \tag{11.47}$$

and

$$S_y = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{with} \quad S_y^2 = \frac{\hbar^2}{4} I_2. \tag{11.48}$$

It is common and convenient to define the vector

$$\sigma = \frac{2}{\hbar}\mathbf{S}, \tag{11.49}$$

whose components are the *Pauli matrices*

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{and} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{11.50}$$

It is easy to verify that

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{11.51}$$

$$\sigma_i \sigma_j = -\sigma_j \sigma_i = \ i\epsilon_{ijk}\ \sigma_k \tag{11.52}$$

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k. \tag{11.53}$$

In the next section we will present the Pauli equation. In the derivation of this equation we will require the following identities

$$(\sigma \cdot \mathbf{a})(\sigma \cdot \mathbf{b}) = \mathbf{a} \cdot \mathbf{b} + i\sigma \cdot \mathbf{a} \times \mathbf{b}, \tag{11.54}$$

$$\big(\sigma \cdot (\nabla + \mathbf{f})\big)\big(\sigma \cdot (\nabla + \mathbf{f})\big) = (\nabla + \mathbf{f})(\nabla + \mathbf{f})I_2 + i\sigma \cdot (\nabla \times \mathbf{f}). \tag{11.55}$$

11.2.3 Pauli's Equation

In the Hamiltonian of equation (11.5) we introduced the spin-dependent potential energy, analogous to the interaction of magnetic moment with magnetic field, that came out with the transformation $\widehat{\mathbf{p}} \rightarrow \widehat{\mathbf{p}} + \mathbf{qA}/\mathbf{c}$, prescribed by the classical theory. As the spin is a purely quantum property, there is no classical analogue. The alternative, thus, is to change the Schrödinger equation so that the spin-field interaction will appear only in the presence of a magnetic field and disappear in its absence. This kind of extension was done by Pauli in 1927. He showed that replacing the vector $\widehat{\mathbf{p}}$ by the matrix[4] $\sigma \cdot \widehat{\mathbf{p}}$, the new equation is exactly the same as the old Schrödinger equation, but, when an external magnetic field is present, two energy terms appear: the spin-field and the orbital momentum-field interactions, as in (11.5). Let us see this with some detail.

If we consider $\mathbf{a} = \mathbf{b} = \widehat{\mathbf{p}}$ in (11.54), we have

$$(\sigma \cdot \widehat{\mathbf{p}})(\sigma \cdot \widehat{\mathbf{p}}) = \widehat{\mathbf{p}} \cdot \widehat{\mathbf{p}}\, I_2. \tag{11.56}$$

This relation tells us that if we change the Hamiltonian

$$\frac{\widehat{\mathbf{p}} \cdot \widehat{\mathbf{p}}}{2m} + V(\mathbf{r}) \quad \text{by} \quad \frac{(\sigma \cdot \widehat{\mathbf{p}})(\sigma \cdot \widehat{\mathbf{p}})}{2m} + V(\mathbf{r})I_2, \tag{11.57}$$

and, as a consequence, we change the wave function $\varphi(\mathbf{r})$ by $\phi(\mathbf{r}, s) = \varphi(\mathbf{r})\chi(s)$, the Schrödinger equation becomes

$$\frac{(\sigma \cdot \widehat{\mathbf{p}})(\sigma \cdot \widehat{\mathbf{p}})}{2m}\phi(\mathbf{r}, s) + V(\mathbf{r})\phi(\mathbf{r}, s) = E\phi(\mathbf{r}, s). \tag{11.58}$$

In the absence of a magnetic field, the kinetic term is independent of σ, according to (11.56), and this equation coincides with the Schrödinger equation independent of σ. But, when an external magnetic field is present the linear momentum $\widehat{\mathbf{p}}$ becomes $\widehat{\mathbf{p}} + \mathbf{qA}/\mathbf{c}$, and the previous equation transforms into

$$\frac{\big(\sigma \cdot (\widehat{\mathbf{p}} + \mathbf{qA}/\mathbf{c})\big)\big(\sigma \cdot (\widehat{\mathbf{p}} + \mathbf{qA}/\mathbf{c})\big)}{2m}\phi(\mathbf{r}, s) + V(\mathbf{r})\phi(\mathbf{r}, s) = E\phi(\mathbf{r}, s). \tag{11.59}$$

---

[4] Simultaneously we must change $V(\mathbf{r})$ by $V(\mathbf{r})I_2$. Usually when we are dealing with matrix interaction, and we do not write explicitly the unit matrix I_2, one must understand that it is there.

If we replace ∇ by $-i\hbar\nabla$and $\mathbf{f}$ for $q\mathbf{A}/c$ in the identity (11.55), the kinetic energy term

$$\frac{\left(\sigma\cdot(\widehat{\mathbf{p}}+\mathbf{qA/c})\right)\left(\sigma\cdot(\widehat{\mathbf{p}}+\mathbf{qA/c})\right)}{2m}, \tag{11.60}$$

becomes

$$\frac{(\widehat{\mathbf{p}}+\mathbf{qA/c})\cdot(\widehat{\mathbf{p}}+\mathbf{qA/c})}{2m}I_2+\frac{\hbar q}{2mc}\sigma\cdot\mathbf{B}. \tag{11.61}$$

At the end of Chap. 10, dealing with the Hydrogen atom in a magnetic field, we saw that the kinetic term, in the Coulomb gauge, could be written as

$$(\widehat{\mathbf{p}}+\mathbf{qA/c})\cdot(\widehat{\mathbf{p}}+\mathbf{qA/c})=\widehat{\mathbf{p}}^2+\frac{q}{c}\mathbf{B}\cdot\widehat{\mathbf{L}}+\frac{q^2}{c^2}\frac{B^2}{4}\left(x^2+y^2\right). \tag{11.62}$$

Therefore, the Schrödinger equation for particles with spin in the presence of a magnetic field is now

$$\left(\frac{-\hbar^2}{2m}\nabla^2+\frac{q}{2mc}\mathbf{B}\cdot\widehat{\mathbf{L}}+\frac{\hbar q}{2mc}\sigma\cdot\mathbf{B}+\frac{q^2}{c^2}\frac{B^2}{4}\left(x^2+y^2\right)+V(\mathbf{r})\right)\phi(\mathbf{r},s)=E\phi(\mathbf{r},s). \tag{11.63}$$

If we replace q by $-e$, σ for $2\mathbf{S}/\hbar$ and introduce the Bohr magneton μ_B we have

$$\left(\frac{-\hbar^2}{2m}\nabla^2-g_{\mathrm{L}}\frac{\mu_B}{\hbar}\mathbf{B}\cdot\widehat{\mathbf{L}}-g_{\mathrm{S}}\frac{\mu_B}{\hbar}\mathbf{B}\cdot\mathbf{S}+\frac{e^2}{c^2}\frac{B^2}{4}\left(x^2+y^2\right)+V(\mathbf{r})\right)\phi(\mathbf{r},s)=E\phi(\mathbf{r},s). \tag{11.64}$$

with $g_{\mathrm{L}}=1$ and $g_{\mathrm{S}}=2$ the Landé factors. This is, precisely, the Pauli equation for spin $1/2$ electrons in a magnetic field that was proposed in (11.5). For the Hydrogen atom $V(\mathbf{r})$ is a central potential and the parabolic potential can be neglected. We should note that all changes that were made here were restricted to the kinetic term, without any conditions on the potential term. Consequently, the equations that we obtained here are valid also to describe charge carriers dynamics through different types of potentials. In particular, we can use this equation if we study the transport of charged particles across semiconductor structures or magnetic superlattices. This kind of systems with additional spin-orbit interactions are part of the actual spintronic devices.

Since the function $\phi(\mathbf{r},s)$ is vectorial, the Hamiltonian in (11.5) should be a matrix equation. When the particle spin is 1/2, it is a 2×2 matrix equation. We will understand that, when the unit matrix does not appear explicitly in some Hamiltonian terms, they are actually multiplied by the corresponding unit matrix. Notice that

$$\widehat{H}I_2\phi(\mathbf{r},s)=\widehat{H}\begin{pmatrix}1&0\\0&1\end{pmatrix}\varphi(\mathbf{r})\begin{pmatrix}a\\b\end{pmatrix}=\widehat{H}\varphi(\mathbf{r})\begin{pmatrix}1&0\\0&1\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}=\widehat{H}\varphi(\mathbf{r})\begin{pmatrix}a\\b\end{pmatrix}. \tag{11.65}$$

This shows that

$$\widehat{H} I_2 \phi(\mathbf{r}, s) = \widehat{H} \phi(\mathbf{r}, s). \tag{11.66}$$

In what follows, we write the matrix I_2 only when we need it.

11.3 The Spin-Orbit Interaction

The spin-field interaction was used to explain the anomalous Zeeman effect and *the fine structure* of the emission lines. In fact, if an orbiting electron, with magnetic moment $\mu_e = g_S \mu_B \widehat{\mathbf{S}}/\hbar$, moves with a velocity $\mathbf{v}$ in an electric field $\mathbf{E}$, it feels in its own reference frame a Lorentz-transformed magnetic field $\mathbf{B}'$ given by (assuming $v \ll c$)

$$\mathbf{B}' = -\frac{\mathbf{v} \times \mathbf{E}}{c^2}. \tag{11.67}$$

This suggests that in the presence of electric fields, the external magnetic field $\mathbf{B}$ in the spin-field interaction of the Hamiltonian should be replaced by $\mathbf{B}+\mathbf{B}'$. Therefore, for electrons in the presence of an electric field, we must include the interaction term

$$-g_S \frac{\mu_B}{\hbar} \mathbf{B}' \cdot \mathbf{S} = -g_S \frac{\mu_B}{\hbar} \left(-\frac{\mathbf{v} \times \mathbf{E}}{c^2} \right) \cdot \mathbf{S} \tag{11.68}$$

In the radial atomic electric field

$$\mathbf{E} = \frac{\mathbf{r}}{r} E, \tag{11.69}$$

the spin-field interaction becomes

$$U = -\frac{E}{r} \mu_e \cdot \frac{\mathbf{v} \times \mathbf{r}}{c^2} = \frac{g_S \mu_B}{m \hbar c^2} \frac{1}{r} \frac{\partial \phi}{\partial r} \mathbf{S} \cdot \mathbf{L}. \tag{11.70}$$

This energy predicts a spacing of the emission lines, *the fine structure*, that is larger by a factor of 2 than the observed spacing. This difficulty could not be solved by changing the Landé factor g_S from 2 to 1. The solution came from the relativistic electron theory of Dirac. A theory that we will not study here but is available in many textbooks. We will only present briefly the Thomas precession picture. In 1927 Thomas[5] gave the following argument to correct the factor 2. When the orbiting electron moves in an external field $\mathbf{E}$, the time derivative of its intrinsic angular momentum $\widehat{\mathbf{S}}$, in the electron's reference frame, is

$$\left.\frac{d\widehat{\mathbf{S}}}{dt}\right|_{rest} = -\mu_e \times \left(\frac{\mathbf{v}}{c^2} \times \mathbf{E} \right). \tag{11.71}$$

---

[5] L. H. Thomas, Nature, vol. 117, 514 (1926).

If the electron's frame moves with an angular velocity ω_T (Thomas frequency), the time derivative of $\widehat{\mathbf{S}}$, in the non rotating reference system (the nucleus), is

$$\left.\frac{d\widehat{\mathbf{S}}}{dt}\right|_{nr} = \left.\frac{d\widehat{\mathbf{S}}}{dt}\right|_{rest} + \widehat{\mathbf{S}} \times \omega_T, \tag{11.72}$$

with

$$\omega_T = \frac{e}{2m_e c^2}\mathbf{v} \times \mathbf{E}. \tag{11.73}$$

If we recall that $\mu_B = e\hbar/2m_e$ and substitute, we have

$$\begin{aligned}\left.\frac{d\widehat{\mathbf{S}}}{dt}\right|_{nr} &= -\mu_e \times \frac{\mathbf{v} \times \mathbf{E}}{c^2} + \frac{1}{g_S}\mu_e \times \frac{\mathbf{v} \times \mathbf{E}}{c^2} \\ &= \frac{1 - g_S}{g_S}\mu_e \times \frac{\mathbf{v} \times \mathbf{E}}{c^2} = \mu_e \times \mathbf{B_T}.\end{aligned} \tag{11.74}$$

This means that the potential energy must be

$$U = -\mu_e \cdot \frac{1 - g_S}{g_S}\frac{\mathbf{v} \times \mathbf{E}}{c^2}. \tag{11.75}$$

This energy, for $g_S = 2$ and for the radial electric field considered before, becomes

$$U = -\frac{E}{2r}\mu_e \cdot \frac{\mathbf{r} \times \mathbf{v}}{c^2}, \tag{11.76}$$

which is smaller by a factor of 2 than the energy in (11.70). Using again the relation $\mu_e = g_S\mu_B\widehat{\mathbf{S}}/\hbar$ we have

$$U = -\frac{\mu_B}{\hbar m_e c^2}\frac{E}{r}\mathbf{S} \cdot \mathbf{L} = \frac{\mu_B}{\hbar m_e c^2}\frac{1}{r}\frac{\partial \phi(r)}{\partial r}\mathbf{S} \cdot \mathbf{L}, \tag{11.77}$$

that is also known as the *spin-orbit interaction*. The coefficient of this interaction depends on the external electric field that was assumed radial.

11.4 The Total Angular Momentum

So far we have introduced, related to the atomic electrons, the orbital angular momentum $\widehat{\mathbf{L}}$ and the intrinsic angular momentum $\widehat{\mathbf{S}}$. To describe these electrons we make use of quantum states defined as the product $\varphi(\mathbf{r})\chi_S$, which means a direct product of the angular momentum and spin state spaces. This product implies the sum of the operators

$$\widehat{\mathbf{J}} = \widehat{\mathbf{L}} + \widehat{\mathbf{S}}, \tag{11.78}$$

and defines the total angular momentum $\widehat{\mathbf{J}}$. This is a particular case. In general, a system may consist of two or more subsystems with angular momenta $\widehat{\mathbf{J}}_1, \widehat{\mathbf{J}}_2, \ldots$ In that case, one can define also the total angular momentum as

$$\widehat{\mathbf{J}} = \widehat{\mathbf{J}}_1 + \widehat{\mathbf{J}}_2 + \ldots \tag{11.79}$$

In the presence of an external magnetic field and of the spin-orbit interaction, the angular momentum $\widehat{\mathbf{L}}$ is not anymore a constant of motion. The Hamiltonian does not commute with $\widehat{\mathbf{L}}$, but $\widehat{\mathbf{L}}$ commutes with $\widehat{\mathbf{S}}$; and the total angular momentum $\widehat{\mathbf{J}}$ commutes with $\widehat{\mathbf{L}}$ and with $\widehat{\mathbf{S}}$. In the absence of external magnetic fields and presence of spin-orbit interaction, the Hamiltonian commutes with $\widehat{\mathbf{J}}$. Hence, the total angular momentum is a conserved quantity and

$$\widehat{J}_x = \widehat{L}_x + \widehat{S}_x, \quad \widehat{J}_y = \widehat{L}_y + \widehat{S}_y, \quad \widehat{J}_z = \widehat{L}_z + \widehat{S}_z, \tag{11.80}$$

The operators $\widehat{\mathbf{J}}$, $\widehat{J}^{\,2}$ and $\widehat{J}_i$ mimic the commutation relations of $\widehat{\mathbf{L}}$, $\widehat{L}^{\,2}$ and $\widehat{L}_i$. Besides this, the operators $\widehat{J}^{\,2}$, $\widehat{J}_z$, $\widehat{L}^{\,2}$ and $\widehat{S}^{\,2}$ commute among them.[6] This means that for definite values of l and s the eigenfunctions are represented by functions characterized by the quantum numbers j, l, s and m_j. If we represent the eigenfunctions with the bras $|jlsm_j\rangle$, the eigenvalue equations of $\widehat{J}^{\,2}$ and $\widehat{J}_z$ will be:

$$\widehat{J}^{\,2}|jlsm_j\rangle = \hbar^2 j(j+1)|jlsm_j\rangle \tag{11.81}$$

$$\widehat{J}_z|jlsm_j\rangle = \hbar m_j|jlsm_j\rangle \quad \text{with} \quad |m_j| \leq j. \tag{11.82}$$

But, what values will the quantum numbers j and m_j take? Suppose that l and s are given. There are a total of $(2l+1)(2s+1)$ possible orientations of $\widehat{\mathbf{L}}$ and $\widehat{\mathbf{S}}$, represented by the different values of m_z and m_s. The largest value that j can take is $l+s$ and the smallest $|l-s|$. The maximum of m_j occurs when m_z and m_s take also their maximum values, and the minimum when they take their minimum values, this means that

$$l+s \geq j \geq |l-s| \quad \text{and} \quad m_j = -l-s, -l-s+1, \ldots, l+s-1, l+s. \tag{11.83}$$

One can easily verify that the number of states characterized by the quantum numbers j, l, s and m_j is equal to the number of states characterized by the quantum numbers l, m_z , s and m_s. Is is always possible to pass from the representation $|jlsm_j\rangle$ to the representation $|lm_zsm_S\rangle$. This is a broad topic in the literature of angular momentum, that is beyond our purpose here. The interested student can consult the specialized literature.[7] From the definition of $\widehat{\mathbf{J}}$ it is obvious that

---

[6] There are at least two sets of mutually commuting operators: the operators $\widehat{J}^{\,2}$, $\widehat{J}_z$, $\widehat{L}^{\,2}$ and $\widehat{S}^{\,2}$ and the operators $\widehat{L}^{\,2}$, $\widehat{L}_z$ $\widehat{S}^{\,2}$ and $\widehat{S}_z$.

[7] See, for example, A.R. Edmonds, *Angular Momentum in Quantum Mechanics*, Princeton University Press, Princeton, N. J. 1974.

$$\widehat{J}^2 = \widehat{L}^2 + \widehat{S}^2 + 2\widehat{\mathbf{S}} \cdot \widehat{\mathbf{L}} \tag{11.84}$$

The eigenvalues of the operator $\widehat{\mathbf{S}} \cdot \widehat{\mathbf{L}}$ can be obtained easily because

$$\widehat{\mathbf{S}} \cdot \widehat{\mathbf{L}} = \frac{1}{2}\left(\widehat{J}^2 - \widehat{L}^2 - \widehat{S}^2\right). \tag{11.85}$$

Indeed,

$$(\widehat{\mathbf{S}} \cdot \widehat{\mathbf{L}})|jlsm_j\rangle = \hbar^2 \frac{j(j+1) - l(l+1) - s(s+1)}{2} |jlsm_j\rangle. \tag{11.86}$$

11.5 Problems

1. Show that the operators

$$\widehat{S}_+ = \widehat{S}_x + i\widehat{S}_y \quad \text{and} \quad \widehat{S}_- = \widehat{S}_x - i\widehat{S}_y, \tag{11.87}$$

   satisfy the commutation relations

$$\left[\widehat{S}_\mp, \widehat{S}_z\right] = \pm\hbar\widehat{S}_\mp. \tag{11.88}$$

2. Show that if $\widehat{S}_z\, \chi = b\, \chi$, then

$$\widehat{S}_z\, \widehat{S}_+ \chi = (b + \hbar)\, \widehat{S}_+ \chi, \tag{11.89}$$

$$\widehat{S}_z\, \widehat{S}_- \chi = (b - \hbar)\, \widehat{S}_- \chi. \tag{11.90}$$

3. Derive the relations

$$\widehat{S}_- \widehat{S}_+ = \widehat{S}^2 - \widehat{S}_z^2 - \hbar\widehat{S}_z, \tag{11.91}$$

$$\widehat{S}_+ \widehat{S}_- = \widehat{S}^2 - \widehat{S}_z^2 + \hbar\widehat{S}_z. \tag{11.92}$$

   Combining these equations with the assumption $b_M = b_m + n\hbar$ to show that $b_m = -n\hbar/2$.
4. Show the relations

$$\widehat{S}_+ \, |sm_s\rangle = \hbar\sqrt{(s+m_s+1)(s-m_s)}\ |s, m_s+1\rangle, \tag{11.93}$$

$$\widehat{S}_- |sm_s\rangle = \hbar\sqrt{(s+m_s)(s-m_s+1)}\ |s, m_s-1\rangle, \tag{11.94}$$

   and obtain, starting from these expressions, the matrix representations S_x and S_y.
5. Using the eigenvalue equation (11.8), show that the matrix representation of S_z is

$$S_z = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{11.95}$$

6. Show that the spin components anticommute, i.e.

$$\widehat{S}_i\widehat{S}_j + \widehat{S}_j\widehat{S}_i = 0 \quad \text{for} \quad i, j = x, y, z \tag{11.96}$$

and furthermore, that

$$2S_i S_j = i\hbar\, \epsilon_{ijk} S_k. \tag{11.97}$$

7. If $\sigma = \frac{2}{\hbar}\mathbf{S}$, determine the matrix representations of σ_x, σ_y and σ_z.
8. Prove the following identities:

$$(\sigma \cdot \mathbf{a})(\sigma \cdot \mathbf{b}) = \mathbf{a} \cdot \mathbf{b} + i\sigma \cdot \mathbf{a} \times \mathbf{b}, \tag{11.98}$$

$$\big(\sigma \cdot (\nabla + \mathbf{f})\big)\big(\sigma \cdot (\nabla + \mathbf{f})\big) = (\nabla + \mathbf{f})(\nabla + \mathbf{f}) I_2 + i\sigma \cdot (\nabla \times \mathbf{f}). \tag{11.99}$$

9. Show the identity

$$(\sigma \cdot \widehat{\mathbf{p}})(\sigma \cdot \widehat{\mathbf{p}}) = \widehat{\mathbf{p}} \cdot \widehat{\mathbf{p}}\; I_2, \tag{11.100}$$

and verify that in the presence of a magnetic field, one has

$$(\sigma \cdot \widehat{\mathbf{p}})(\sigma \cdot \widehat{\mathbf{p}}) \;\longrightarrow\; \widehat{\mathbf{p}}^2 + \frac{q}{c}\mathbf{B} \cdot \widehat{\mathbf{L}} + \frac{q^2}{c^2}\frac{B^2}{4}\big(x^2 + y^2\big). \tag{11.101}$$

Chapter 12
Perturbation Theory

In previous chapters we have seen an important set of systems whose Schrödinger's equations can be solved analytically. There is, however, an even larger number of quantum systems whose Schrödinger's equations can not be solved analytically. Generally one has to use approximate methods like the WKB approximation, the perturbation method, and numerous numerical methods. In this chapter we will study only the fundamentals of the perturbation theory, and the interaction representation. The perturbation theory can be used to approach the correct solution, when part of the Hamiltonian is characterized by a small perturbation parameter η, such that, neglecting the perturbation part $\widehat{V}_p = \eta\widehat{U}$, one is left with a soluble problem for the Hamiltonian $\widehat{H}_o = \widehat{H} - \widehat{V}_p$. When this is the case, the first step of the perturbation method is to solve the problem for $\widehat{H}_o$. In the following steps the theory tells us how, starting from the unperturbed eigenfunctions $\varphi_n^{(0)}$ and eigenvalues $E_n^{(0)}$, one obtains corrections of different orders in the perturbation parameter that will leave us, at the end, with good approximations to the correct solution. We will study, separately, the perturbation theory for time-independent and for time dependent perturbations. For the time-independent perturbations, we will distinguish the degenerate from the non-degenerate case. In the last section of this chapter, we will present some basic expressions of an alternative method to treat time-dependent and time-independent perturbation interactions, the interaction representation. The interaction representation is a formalism that combines the Schrödinger representation, where all the time dependence is in the wave function

$$\psi(\mathbf{r}, t) = e^{-i\widehat{H}t/\hbar}\psi(\mathbf{r}, 0), \tag{12.1}$$

with the Heisenberg representation, where all the time dependence is (see Sect. 7.6) on the operators expressed as:

$$\widehat{V}(t) = e^{i\widehat{H}t/\hbar}\widehat{V}e^{-i\widehat{H}t/\hbar}. \tag{12.2}$$

P. Pereyra, *Fundamentals of Quantum Physics*, Undergraduate Lecture Notes in Physics, DOI: 10.1007/978-3-642-29378-8_12, © Springer-Verlag Berlin Heidelberg 2012

In the interaction representation only the perturbation part $\widehat{V}_p$ of the whole potential $\widehat{V} = \widehat{V}_o + \widehat{V}_p$, becomes a function of time when it is written in the form

$$\widehat{V}_p(t) = e^{i\widehat{H}_o t/\hbar}\widehat{V}_p e^{-i\widehat{H}_o t/\hbar}, \tag{12.3}$$

with $\widehat{H}_o$ the unperturbed Hamiltonian.

12.1 Time-Independent Perturbation Theory

If we have a Hamiltonian like

$$\widehat{H} = \widehat{H}_o + \widehat{V}_p = \widehat{H}_o + \eta\widehat{U}, \tag{12.4}$$

where, as mentioned before, the unperturbed Schrödinger equation for $\widehat{H}_o$ is soluble, one can use the perturbation theory, with different levels of approximation defined by different orders of the smallness parameter η.

Suppose now that the perturbation potential is independent of time and the Schrödinger equation

$$\widehat{H}_o\varphi_n^{(0)} = E_n^{(0)}\varphi_n^{(0)}, \quad \text{or} \quad \widehat{H}_o|n\rangle_0 = E_n^{(0)}|n\rangle_0, \tag{12.5}$$

is soluble. The procedure to obtain the approximate solutions of

$$\widehat{H}\varphi_n = E_n\varphi_n, \tag{12.6}$$

depends on whether the states $\varphi_n^{(0)}$ are degenerate or not. Let us study the non-degenerate and degenerate cases separately.

12.1.1 Perturbation Theory for Non-Degenerate States

If the eigenfunctions $\varphi_n^{(0)}$ are non-degenerate, we can propose for the exact solutions, $|n\rangle$ and E_n, a development in powers of the perturbation parameter η, as follows:

$$|n\rangle = |n\rangle_0 + \eta|n\rangle_1 + \eta^2|n\rangle_2 + \ldots; \tag{12.7}$$

$$E_n = E_n^{(0)} + \eta E_n^{(1)} + \eta^2 E_n^{(2)} + \ldots \tag{12.8}$$

If we replace these expressions in (12.6), we have

$$\big(\widehat{H}_o + \eta\widehat{U}\big)\big(|n\rangle_0 + \eta|n\rangle_1 + \ldots\big) = \big(E_n^{(0)} + \eta E_n^{(1)} +\big)\big(|n\rangle_0 + \eta|n\rangle_1 + \ldots\big). \tag{12.9}$$

Equating the coefficients of the same powers of η, on the left and right sides of this equation, we have:

$$\widehat{H}_o|n\rangle_0 = E_n^{(0)}|n\rangle_0 \tag{12.10}$$

$$\widehat{H}_o|n\rangle_1 + \widehat{U}|n\rangle_0 = E_n^{(0)}|n\rangle_1 + E_n^{(1)}|n\rangle_0 \tag{12.11}$$

$$\widehat{H}_o|n\rangle_2 + \widehat{U}|n\rangle_1 = E_n^{(o)}|n\rangle_2 + E_n^{(1)}|n\rangle_1 + E_n^{(2)}|n\rangle_0. \tag{12.12}$$

The first of these equations is the Schrödinger equation of the unperturbed system, which we have assumed soluble. Thus we have $E_n^{(0)}$ and $|n\rangle_0$. Let us assume that $|n\rangle_0$ is normalized. The second equation, allows us to obtain the first order corrections to the energy and wave function. This equation depends also on $E_n^{(0)}$ and $|n\rangle_0$. Before using the second equation we write it as:

$$\big(\widehat{H}_o - E_n^{(0)}\big)|n\rangle_1 = \big(E_n^{(1)} - \widehat{U}\big)|n\rangle_0. \tag{12.13}$$

If we multiply this equation from the left by ${}_0\langle n|$ and take into account, here and in the following, that

$${}_0\langle n|\widehat{H}_o = {}_0\langle n|E_n^{(0)}, \quad {}_0\langle k|n\rangle_0 = \delta_{k,n}, \quad {}_0\langle k|\widehat{U}|n\rangle_0 = U_{k,n}^{(0)}, \tag{12.14}$$

the left side vanishes, and from the right side we have that

$$E_n^{(1)} = U_{n,n}^{(0)}. \tag{12.15}$$

This means that, to first order in the perturbation, the correction to the unperturbed energy is just the expected value of the perturbation potential evaluated in the unperturbed state $|n\rangle_0$. Thus, to first order in η, we have

$$E_n = E_n^{(0)} + \eta U_{n,n}^{(0)}. \tag{12.16}$$

For some problems this degree of approximation could be enough. Let us now obtain the first-order correction $\eta|n\rangle_1$ to the unperturbed solution $|n\rangle_0$. To obtain this correction we will use the complete set of the unperturbed functions $|n\rangle_0$. We can express $\eta|n\rangle_1$ as the linear combination

$$|n\rangle_1 = \sum_k c_{nk}^{(1)}|k\rangle_0 = \sum_k {}_0\langle k|n\rangle_1|k\rangle_0, \tag{12.17}$$

which requires the coefficients $c_{nk}^{(1)} = {}_0\langle k|n\rangle_1$. Our next task is to obtain these coefficients. Let us go back to (12.13), and write that equation in the form

$$\big(\widehat{H}_o - E_n^{(0)}\big)|n\rangle_1 = \big(U_{n,n}^{(0)} - \widehat{U}\big)|n\rangle_0, \tag{12.18}$$

where we have replaced $E_n^{(1)}$ by $U_{n,n}^{(0)}$. Multiplying this equation, from the left, with $_0\langle k|$, we have

$$\big(E_k^{(0)} - E_n^{(0)}\big)_0 \langle k|n\rangle_1 = \big(U_{n,n}^{(0)}\delta_{k,n} - U_{k,n}^{(0)}\big). \tag{12.19}$$

This gives us the coefficient

$$_0\langle k|n\rangle_1 = \frac{\big(U_{n,n}^{(0)}\delta_{k,n} - U_{k,n}^{(0)}\big)}{\big(E_k^{(0)} - E_n^{(0)}\big)}, \quad \text{for} \quad k \neq n. \tag{12.20}$$

Therefore, the expansion (12.17) can be written as

$$|n\rangle_1 = \sum_{k\neq n} \frac{U_{k,n}^{(0)}}{\big(E_n^{(0)} - E_k^{(0)}\big)} |k\rangle_0, \tag{12.21}$$

and the wave function to first order in η is

$$|n\rangle = |n\rangle_0 + \eta \sum_{k\neq n} \frac{U_{k,n}^{(0)}}{\big(E_n^{(0)} - E_k^{(0)}\big)} |k\rangle_0. \tag{12.22}$$

With quite similar procedures one can obtain the higher-order corrections. It can be shown that for the second order corrections we have

$$E_n^{(2)} = \sum_{k\neq n} \frac{|U_{k,n}^{(0)}|^2}{\big(E_n^{(0)} - E_k^{(0)}\big)} \equiv \sum_{k\neq n} \frac{|U_{k,n}^{(0)}|^2}{\hbar\omega_{nk}}; \tag{12.23}$$

$$|n\rangle^{(2)} = \sum_{l,k\neq n} \frac{U_{l,k}^{(0)} U_{k,n}^{(0)}}{\hbar^2\omega_{nk}\omega_{nl}} |l\rangle_0 - \sum_{k\neq n} \frac{U_{n,n}^{(0)} U_{k,n}^{(0)}}{\hbar^2\omega_{nk}^2} |k\rangle_0 - \frac{1}{2} \sum_{k\neq n} \frac{|U_{k,n}^{(0)}|^2}{\hbar^2\omega_{nk}^2} |n\rangle_0. \tag{12.24}$$

12.1.2 Perturbation Theory for Degenerate States

Let us suppose now that we are dealing with an unperturbed Hamiltonian H_o that has degenerate eigenstates $|n\rangle_0$. When the number of states corresponding to the same energy level is, say d, we have a degeneracy degree d. Let us suppose also that the perturbation potential V_p breaks or removes this degeneracy. Even in that case, the perturbation method changes slightly. Applying the perturbation method one obtains a secular equation, whose roots, generally real and different, provides corrections to the unperturbed energy E_n^0.

If the degeneracy degree of E_n^0 is d, the eigenfunctions corresponding to this unperturbed energy will be denoted as $\varphi_{n_1}^0$, $\varphi_{n_2}^0$,…, and $\varphi_{n_d}^0$. In the Dirac notation these are represented by $|n_1\rangle_0$, $|n_2\rangle_0$,..., and $|n_d\rangle_0$. We have two choices to continue with the general formalism of the perturbation theory of degenerate states. One is to proceed as if we do not know that the perturbation potential would remove the degeneracy. Another is to proceed by assuming that the degeneracy will be removed when $\eta \neq 0$. Thus, the non-degenerate wave functions $|n\rangle_0 + \eta|n_i\rangle_1$ corresponding to the energy levels $E^{(1)}_{n_i}$ (with i=1,2,...,d), reduce to a combination of d orthogonal degenerate states, i.e. to $|n\rangle_0 = \sum_{j=1}^{d} a_{nj}|n_j\rangle_0$. We will follow the last alternative. Our main objective will be to determine the non-degenerate energy levels and the coefficients a_{nj} for the corresponding linear combination of degenerate states. Therefore, we can develop the solutions $|n\rangle$ and E_n as follows:

$$|n_i\rangle = \sum_{j=1}^{d} a_{ij}|nj\rangle_0 + \eta|n_i\rangle_1 + \eta^2|n_i\rangle_2 + \dots \quad i = 1, 2, \dots d, \tag{12.25}$$

$$E_{n_i} = E_n^{(0)} + \eta E_{n_i}^{(1)} + \eta^2 E_{n_i}^{(2)} + \dots \qquad i = 1, 2, \dots d. \tag{12.26}$$

If we substitute these expressions into

$$\widehat{H}|n_i\rangle = E_{n_i}|n_i\rangle, \tag{12.27}$$

we have

$$\left(\widehat{H}_o+\eta\widehat{U}\right)\left(\sum_{j=1}^{d} a_{ij}|n_j\rangle_0+\eta|n_i\rangle_1+\dots\right) = \left(E_n^{(0)}+\eta E_{n_i}^{(1)}+\dots\right) \left(\sum_{j=1}^{d} a_{ij}|n_j\rangle_0+\eta|n_i\rangle_1+\dots\right). \tag{12.28}$$

Again, equating the coefficients of equal powers of η in the left and right sides of this equation, we have

$$\widehat{H}_o|n\rangle_0 = E_n^{(0)}|n\rangle_0, \tag{12.29}$$

$$\widehat{H}_o|n_i\rangle_1 + \widehat{U}\sum_{j=1}^{d} a_{ij}|n_j\rangle_0 = E_n^{(0)}|n_i\rangle_1 + E_{n_i}^{(1)}\sum_{j=1}^{d} a_{ij}|n_j\rangle_0, \tag{12.30}$$

$$\widehat{H}_o|n_i\rangle_2 + \widehat{U}|n_i\rangle_1 = E_n^{(o)}|n_i\rangle_2 + E_{n_i}^{(1)}|n_i\rangle_1 + E_{n_i}^{(2)}\sum_{j=1}^{d} a_{ij}|n_j\rangle_0. \tag{12.31}$$

The first equation is the Schrödinger equation of the unperturbed Hamiltonian. The second yields the first order corrections $E_{n_i}^{(1)}$, that we are looking for. Let us write this equation in the form

$$\big(\widehat{H}_o - E_n^{(0)}\big)|n_i\rangle_1 = \big(-\widehat{U} + E_{n_i}^{(1)}\big)\sum_{j=1}^{d} a_{ij}|n_j\rangle_0, \tag{12.32}$$

and multiply from the left with ${}_0\langle n_l|$. Assuming that the unperturbed degenerate functions $|n_j\rangle_0$ are normalized, we obtain the following system of equations:

$$\sum_{j=1}^{2} a_{ij}\ {}_0\langle n_l|\widehat{U}|n_j\rangle_0 - E_{n_i}^{(1)}\sum_{j=1}^{d} a_{ij}\ \delta_{l,j} = 0 \qquad i,l = 1,2,\ldots d, \tag{12.33}$$

that can be written as

$$\sum_{j=1}^{d} a_{ij}\ \Big({}_0\langle n_l|\widehat{U}|n_j\rangle_0 - E_{n_i}^{(1)}\delta_{l,j}\Big) = 0 \qquad \text{with} \qquad i,l = 1,2,\ldots d. \tag{12.34}$$

It is well known that this homogeneous system of linear equations, has non trivial solutions for the coefficients a_{nj}, if

$$\begin{vmatrix} {}_0\langle n_1|\widehat{U}|n_1\rangle_0 - E_{n_i}^{(1)} & {}_0\langle n_1|\widehat{U}|n_2\rangle_0 & \ldots & {}_0\langle n_1|\widehat{U}|n_d\rangle_0 \\ {}_0\langle n_2|\widehat{U}|n_1\rangle_0 & {}_0\langle n_2|\widehat{U}|n_2\rangle_0 - E_{n_i}^{(1)} & \ldots & {}_0\langle n_2|\widehat{U}|n_d\rangle_0 \\ \vdots & & \ddots & \\ {}_0\langle n_d|\widehat{U}|n_1\rangle_0 & {}_0\langle n_d|\widehat{U}|n_2\rangle_0 & \ldots & {}_0\langle n_d|\widehat{U}|n_d\rangle_0 - E_{n_i}^{(1)} \end{vmatrix} = 0. \tag{12.35}$$

This condition gives us the energies $E_{n_i}^{(1)}$. It is clear and easy to verify that the secular equation corresponding to the energy level $E_{n_i}^{(1)}$ is identical to that of $E_{n_j}^{(1)}$. Thus all the determinants lead to the same polynomial of degree d in $E_{n_i}^{(1)}$ (or $E_{n_j}^{(1)}$). It is then sufficient to solve one of the secular equations. Generally the roots are different, and give us first order corrections $E_{n_1}^{(1)}$, $E_{n_2}^{(1)}$,... and $E_{n_d}^{(1)}$ to $E_n^{(0)}$.

To simplify the discussion we will consider $d = 2$. It is convenient to use the notation that we used in Eq. (12.14), for the matrix elements $U_{l,j}^{(0)} = {}_0\langle n_l|\widehat{U}|n_j\rangle_0$, and, to further simplify the notation, we will remove the superscript ${}^{(0)}$ which indicates evaluation of the matrix element with the unperturbed states. From the secular Eq. (12.35) we have the polynomial

$$E_{n_i}^{(1)2} - E_{n_i}^{(1)}\big(U_{1,1} + U_{2,2}\big) + U_{1,1}U_{2,2} - U_{1,2}U_{2,1} = 0, \tag{12.36}$$

with roots

$$E_{n_i}^{(1)} = \frac{1}{2}\left[U_{1,1} + U_{2,2} \mp \sqrt{(U_{1,1} - U_{2,2})^2 + 4U_{1,2}U_{2,1}}\right]. \tag{12.37}$$

Therefore, to first order of the perturbation parameter η, the non-degenerate energies are:

$$E_{n_1} = E_n^{(0)} + \frac{\eta}{2}\left[U_{1,1} + U_{2,2} - \sqrt{(U_{1,1} - U_{2,2})^2 + 4U_{1,2}U_{2,1}}\right] \tag{12.38}$$

and

$$E_{n_2} = E_n^{(0)} + \frac{\eta}{2}\left[U_{1,1} + U_{2,2} + \sqrt{(U_{1,1} - U_{2,2})^2 + 4U_{1,2}U_{2,1}}\right]. \tag{12.39}$$

These energies, to first order of the square root development, are:

$$E_{n_1} = E_n^{(0)} + \eta\left[U_{1,1} - \frac{U_{1,2}U_{2,1}}{U_{1,1} - U_{2,2}}\right] \tag{12.40}$$

and

$$E_{n_2} = E_n^{(0)} + \eta\left[U_{1,1} + \frac{U_{1,2}U_{2,1}}{U_{1,1} - U_{2,2}}\right]. \tag{12.41}$$

We can now obtain the coefficients a_{ij} of the linear combinations corresponding to $E_{n_1}^{(1)}$ and $E_{n_1}^{(2)}$. If we plug the energy corrections $E_{n_i}^{(1)}$ in the system of Eq. (12.34), for $d = 2$, and normalize the linear combination, we get for $E_{n_1}^{(1)}$ (notice that $U_{1,2} = U_{2,1}^*$)

$$a_{n_1 1} = -a_{n_1 2}\frac{U_{1,2}}{U_{1,1} - E_{n_1}^{(1)}} \tag{12.42}$$

and

$$|a_{n_1 1}|^2 = 1 - |a_{n_1 2}|^2, \tag{12.43}$$

with solutions

$$|a_{n_1 1}|^2 = \frac{1}{2}\left[1 + \frac{U_{1,1} - U_{2,2}}{\sqrt{(U_{1,1} - U_{2,2})^2 + 4U_{1,2}U_{2,1}}}\right] \tag{12.44}$$

and

$$|a_{n_1 2}|^2 = \frac{1}{2}\left[1 - \frac{U_{1,1} - U_{2,2}}{\sqrt{(U_{1,1} - U_{2,2})^2 + 4U_{1,2}U_{2,1}}}\right]. \tag{12.45}$$

One finds similar results for $E_{n_2}^{(1)}$.

12.2 Time-Dependent Perturbations

There is a large number of systems whose Hamiltonians depend on time. In those cases, the energy is, of course, not conserved. If a Hamiltonian of this kind can be written as the sum of a stationary Hamiltonian $\widehat{H}_o$ plus a time-dependent perturbation potential, i.e., if

$$\widehat{H} = \widehat{H}_o + \widehat{V}_p(t), \tag{12.46}$$

it makes sense to think of time-dependent corrections to the energies and functions of the stationary system. The aim of the perturbation theory is to determine the time evolution of the system based on the solutions of the stationary problem. To simplify the notation, let $\widehat{U}(t) \equiv \widehat{V}_p(t)$. Suppose that $|\psi_n\rangle_0$ (or $|n\rangle_0$), is a solution of the stationary problem. The solution of

$$\big(\widehat{H}_o + \widehat{U}(t)\big)|\psi\rangle = ih\frac{\partial|\psi\rangle}{\partial t}, \tag{12.47}$$

will be written as the linear combination

$$|\psi\rangle = \sum_k a_n(t)|\psi_n\rangle_0, \tag{12.48}$$

with $|\psi_n\rangle_0 = e^{-iE_nt/\hbar}\varphi_n(\mathbf{r})$ and a_n time dependent coefficients. Our purpose is to determine these coefficients. If we replace the combination (12.44) into (12.47), and take into account that

$$\widehat{H}_o|\psi_n\rangle_0 = ih\frac{\partial|\psi_n\rangle_0}{\partial t}, \tag{12.49}$$

we find the equation

$$i\hbar\sum_n |\psi_n\rangle_0\frac{\partial a_n}{\partial t} = \sum_n a_n\widehat{U}|\psi_n\rangle_0. \tag{12.50}$$

If we now multiply this equation from the left by ${}_0\langle\psi_k|$, and separate the time dependent part $|\psi_n\rangle_0$, we obtain the system of time-evolution equations

$$i\hbar\frac{\partial a_k}{\partial t} = \sum_n a_n U_{k,n}(t)e^{i\left(E_k^{(0)}-E_n^{(0)}\right)t/\hbar}. \tag{12.51}$$

To solve this system of equations we use a perturbation method as follows. If at $t = 0$ the system is in one of the eigenstates of $\widehat{H}_o$, say in the state $|\psi_n\rangle_0$, it is clear that at this time $a_n(0) = 1$, and $a_k(0) = 0$ for $k \neq n$. It is plausible to suppose that, sometime later, the coefficients can be expressed, approximately, as

$$a_l(t) = a_l(0) + a_l^{(1)}(t) \tag{12.52}$$

with $a_l(t) \ll a_n(t) \approx 1$. In this case, the system of equations takes the form

$$i\hbar\frac{da_k}{dt} = U_{k,n}(t)e^{i\left(E_k^{(0)} - E_n^{(0)}\right)t/\hbar}, \tag{12.53}$$

that can be written as

$$a_k(t) = -\frac{i}{\hbar}\int_0^t U_{k,n}(t')e^{i\omega_{kn}t'}dt' \quad k \neq n, \tag{12.54}$$

where the frequency $\omega_{kn}{=}(E_k^{(0)} - E_n^{(0)})/\hbar$ was defined. These coefficients have a very interesting physical meaning. They represent the possibility of the system, which at time $t = 0$ is in the state $|\psi_n\rangle_0$ with energy $E_n^{(0)}$, to be found at time t in the state $|\psi_k\rangle_0$ with energy $E_k^{(0)}$. Therefore, the transition probability, from the state $|\psi_n\rangle_0$ to the state $|\psi_k\rangle_0$, is:

$$|a_k(t)|^2 = \frac{1}{\hbar^2}\left|\int_0^t U_{k,n}(t')e^{i\omega_{kn}t'}dt'\right|^2, \tag{12.55}$$

and the probability that the system remains in the state $|\psi_n\rangle_0$, i.e. the survival probability, will be:

$$|a_n(t)|^2 = 1 - \sum_{k\neq n}|a_k(t)|^2 = 1 - \frac{1}{\hbar^2}\sum_{k\neq n}\left|\int_0^t U_{k,n}(t')e^{i\omega_{kn}t'}dt'\right|^2. \tag{12.56}$$

If the potential varies slowly with time, the matrix elements $U_{k,n}$ are almost constant and can be taken out of the integral, assuming of course that t is not too large. After integration, we have the transition amplitude

$$a_k(t) = -\frac{{}_0\langle\psi_k\,|U|\psi_n\rangle_0}{E_k^{(0)} - E_n^{(0)}}\left(1 - e^{i\omega_{kn}t}\right), \quad k \neq n \tag{12.57}$$

and the transition probability per unit time, in this case the transition probability from the state $|\psi_n\rangle_0$ to the state $|\psi_k\rangle_0$, will be

$$\mathrm{W}_{kn} = \frac{|a_k(t)|^2}{t} = \frac{4\left|{}_0\langle\psi_k\,|U|\psi_n\rangle_0\right|^2}{\hbar^2}\frac{\sin^2(\omega_{kn}t/2)}{\omega_{kn}^2 t}. \tag{12.58}$$

As can be seen in Fig. 12.1, where we plot the function $\sin^2(\omega_{kn}t/2)/\omega_{kn}^2 t$, the highest maximum of this function grows with time while the width becomes narrower. The region of frequencies ω_{kn} with higher contributions is basically in the range

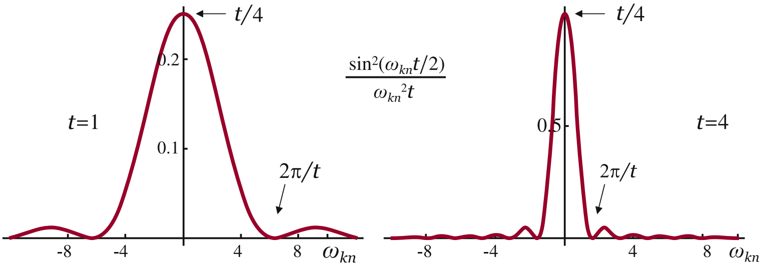


Fig. 12.1 The behavior of the function $\sin^2(\omega_{kn}t/2)/\omega_{kn}^2 t$ at t =1 and t =4. The maximum value of this function is $t/4$ and the higher amplitudes region extends from $-2\pi/t$ to $2\pi/t$. The area below the curve, from $-\infty$ to ∞, for any value of t is $\pi/2$

$|\omega_{kn}| \leq 2\pi/t = \epsilon/\hbar$. In the limit $t \to \infty$ the function tends to $(\pi/2)\delta(E_n^{(0)} - E_k^{(0)})$. We will comment on this limit at the end of this section.

Based on (12.58) we can conclude that the probability to find the system in any of the states $|\psi_k\rangle_0$ is

$$\mathrm{W}_n = \sum_{k\neq n} \mathrm{W}_{kn} = \sum_{k\neq n} \frac{4\left|{}_0\langle\psi_k\,|U|\psi_n\rangle_0\right|^2}{\hbar^2} \frac{\sin^2(\omega_{kn}t/2)}{\omega_{kn}^2 t} \tag{12.59}$$

This expression can, approximately, be evaluated if we replace the factor function $f(\omega_{kn}, t) = \sin^2(\omega_{kn}t/2)/\omega_{kn}^2 t$ by a frequency-independent function $g(t)$, defined in such a way that the integral, in the frequency region $(-\pi/t, \pi/t)$, is also $\pi/2$. That function is $t/4$ and the sum, is then restricted to the states $(-\pi\hbar/t, \pi\hbar/t)$. If $\rho(E_k^{(0)})$ is the density of states, the number of states in the energy region $(-\pi\hbar/t, \pi\hbar/t)$ is

$$\Delta N = \rho(E_k^{(0)})\frac{2\pi\hbar}{t}. \tag{12.60}$$

These considerations can lead to the approximation

$$\sum_{k\neq n} \frac{4\left|{}_0\langle\psi_k\,|U|\psi_n\rangle_0\right|^2}{\hbar^2} \frac{\sin^2(\omega_{kn}t/2)}{\omega_{kn}^2 t} \to \frac{4\left|{}_0\langle\psi_k\,|U|\psi_n\rangle_0\right|^2}{\hbar^2}\,\rho(E_k^{(0)})\frac{2\pi\hbar}{t}\,\frac{t}{4}. \tag{12.61}$$

With this approximation, it is easy to see that the transition probability per unit time is given by

$$\mathrm{W}_n = \frac{2\pi\left|{}_0\langle\psi_k\,|U|\psi_n\rangle_0\right|^2}{\hbar}\,\rho(E_k^{(0)}). \tag{12.62}$$

This is known as the *Fermi golden rule*. This formula has been widely applied for simple evaluations of quantum transitions, especially in nuclear and high energy physics.

If, on the other hand, we consider the limit $t \to \infty$ and substitute the function $\sin^2(\omega_{kn}t/2)/\omega_{kn}^2 t$ by $(\pi/2)\delta(E_n^{(0)} - E_k^{(0)})$ in Eq. (12.59), we go to the continuous limit. Changing the sum by an integral

$$\mathrm{W}_n \to \int_{-\infty}^{\infty} \rho(E_k^{(0)}) \frac{2\pi \left|{}_0\langle\psi_k\,|U|\psi_n\rangle_0\right|^2}{\hbar} \delta(E_n^{(0)} - E_k^{(0)}) dE_k^{(0)}, \tag{12.63}$$

we obtain

$$\mathrm{W}_n = \frac{2\pi \left|{}_0\langle\psi_n\,|U|\psi_n\rangle_0\right|^2}{\hbar} \rho(E_n^{(0)}). \tag{12.64}$$

According to this result, as $t \to \infty$, the transition become negligible.

12.3 The Interaction Representation

We will see now a general procedure to solve the equation

$$i\hbar \frac{\partial \psi}{\partial t} = \widehat{H}\psi, \tag{12.65}$$

when

$$\widehat{H} = \widehat{H}_o + \widehat{V}(t). \tag{12.66}$$

This method is based on the assumption that the time evolution of the wave function $\psi(\mathbf{r}, t)$ is determined by an evolution operator that depends on the time-dependent potential $\widehat{V}(t)$. If we define the function

$$\widetilde{\psi}(\mathbf{r}, t) = e^{i\widehat{H}_o t/\hbar}\psi(\mathbf{r}, t), \tag{12.67}$$

such that at $t = 0$

$$\widetilde{\psi}(\mathbf{r}, 0) = \psi(\mathbf{r}, 0), \tag{12.68}$$

it is clear that (to simplify the notation we show only the time dependence)

$$\begin{aligned}\frac{\partial\widetilde{\psi}(t)}{\partial t} &= \frac{i}{\hbar}\widehat{H}_o e^{i\widehat{H}_o t/\hbar}\ \psi(t) + e^{i\widehat{H}_o t/\hbar}\ \frac{\partial\psi(t)}{\partial t}\\ &= \frac{i}{\hbar}e^{i\widehat{H}_o t/\hbar}\widehat{H}_o\ \psi(t) + e^{i\widehat{H}_o t/\hbar}\ \left(-\frac{i}{\hbar}\widehat{H}\psi(t)\right)\\ &= -\frac{i}{\hbar}\ \left(e^{i\widehat{H}_o t/\hbar}\widehat{V}(t)e^{-i\widehat{H}_o t/\hbar}\right)\ \widetilde{\psi}(t). \end{aligned} \tag{12.69}$$

This equation can be written as

$$i\hbar\frac{\partial\widetilde{\psi}(t)}{\partial t} = \widehat{V}'(t)\widetilde{\psi}(t), \tag{12.70}$$

with $\widehat{V}'(t)$

$$\widehat{V}'(t) = e^{i\widehat{H}_o t/\hbar}\widehat{V}(t)e^{-i\widehat{H}_o t/\hbar}. \tag{12.71}$$

Formally, the solution of Eq. (12.70) is

$$\widetilde{\psi}(t) = e^{-i\widehat{V}'(t)(t-t_0)/\hbar}\widetilde{\psi}(t_0) = \widehat{U}(t,t_0)\ \widetilde{\psi}(t_0). \tag{12.72}$$

We will try to obtain a simpler representation of the unitary operator $\widehat{U}(t,t_0)$. If we take the time derivative of

$$\widehat{U}(t,t_0) = e^{-i\widehat{V}'(t)(t-t_0)/\hbar}, \tag{12.73}$$

we have the first-order differential equation

$$i\hbar\frac{d\widehat{U}(t,t_0)}{dt} = \widehat{V}'(t)\widehat{U}(t,t_0), \tag{12.74}$$

which, after integration from t_0 to t, transforms into

$$\widehat{U}(t,t_0) - \widehat{U}(t_0,t_0) = -\frac{i}{\hbar}\int\limits_{t_0}^{t}\widehat{V}'(t_1)\widehat{U}(t_1,t_0)dt_1. \tag{12.75}$$

From the definition of the operators $\widehat{V}'(t_1)$ and $\widehat{U}(t,t_0)$, it is easy to see that the above equation can be written as

$$\widehat{U}(t,t_0) = 1 - \frac{i}{\hbar}\int\limits_{t_0}^{t} e^{i\widehat{H}_o t_1/\hbar}\widehat{V}(t_1)e^{-i\widehat{H}_o t_1/\hbar}\widehat{U}(t_1,t_0)dt_1, \tag{12.76}$$

which is a time evolution operator where both parts of the Hamiltonian, $\widehat{H}_o$ and $\widehat{V}(t)$ appear explicitly. This is an integral equation for $\widehat{U}(t,t_0)$ which is solved iteratively.

At each stage of the iterative process we have, in principle, a better approximation. In the iterative process we start with $\widehat{U}^{(0)}(t,t_0)=1$. This is the zero order solution. If we use this in (12.76), we obtain the first order solution

$$\widehat{U}^{(1)}(t,t_0)=1-\frac{i}{\hbar}\int\limits_{t_0}^{t}e^{i\widehat{H}_ot/\hbar}\widehat{V}(t)e^{-i\widehat{H}_ot/\hbar}dt_1. \tag{12.77}$$

In the second step of the iterative procedure we introduce this operator in (12.76), to obtain the second order solution

$$\begin{aligned}\widehat{U}^{(2)}(t,t_0)=&1-\frac{i}{\hbar}\int\limits_{t_0}^{t}e^{i\widehat{H}_ot_1/\hbar}\widehat{V}(t_1)e^{-i\widehat{H}_ot_1/\hbar}dt_1\\ &\left(1-\frac{i}{\hbar}\int\limits_{t_0}^{t_1}e^{i\widehat{H}_ot_2/\hbar}\widehat{V}(t_2)e^{-i\widehat{H}_ot_2/\hbar}dt_2\right),\end{aligned} \tag{12.78}$$

that can be written as

$$\begin{aligned}\widehat{U}^{(2)}(t,t_0)=&1-\frac{i}{\hbar}\int\limits_{t_0}^{t}e^{i\widehat{H}_ot_1/\hbar}\widehat{V}(t_1)e^{-i\widehat{H}_ot_1/\hbar}dt_1\\ &-\frac{1}{\hbar^2}\int_{t_0}^{t}dt_1\int\limits_{t_0}^{t_1}dt_2\;e^{i\widehat{H}_ot_1/\hbar}\widehat{V}(t_1)e^{i\widehat{H}_o(t_2-t_1)/\hbar}\widehat{V}(t_2)e^{-i\widehat{H}_ot_2/\hbar}.\end{aligned} \tag{12.79}$$

This procedure continues, in principle, to any order. Given this evolution operator, we can write the wave function $\psi(\mathbf{r},t)$, for example to second order, as

$$\psi(\mathbf{r},t)=e^{-i\widehat{H}_ot/\hbar}\Big(1+\widehat{U}^{(1)}(t,t_0)+\widehat{U}^{(2)}(t,t_0)\Big)\;\psi(\mathbf{r},0), \tag{12.80}$$

with $\widehat{U}^{(1)}(t,t_0)$ and $\widehat{U}^{(2)}(t,t_0)$ as defined above.

12.4 Solved Problems

Exercise 37. If an atom is in a weak external electric field $\mathcal{E}$, the interaction of the atomic electrons with this field, can be treated as a perturbation term. Obtain the first and second order corrections to the electron energy.

Solution We know that, between the field $\mathcal{E}$ and the electric potential ϕ, holds the relation

$$\mathcal{E} = -\nabla \cdot \phi. \tag{12.81}$$

If we multiply by $d\mathbf{r}$ and integrate, with the origin of the coordinates system at the atomic center, the electron's potential energy is

$$V = q\phi = e\mathcal{E} \cdot \mathbf{r}. \tag{12.82}$$

The product $e\mathcal{E} \simeq 1.60218\,10^{-19}\mathcal{E}$ is, certainly, small in the presence of a weak electric field $\mathcal{E}$, and the effect on the energy can be treated as a perturbation. Using the matrix elements

$$\mathbf{r}_{kn}^{(0)} = {}_0\langle k|\widehat{\mathbf{r}}|n\rangle_0, \tag{12.83}$$

the atomic electron energy in the presence of an external field, up to second order in $e\mathcal{E}$, is then given by

$$E_n = E_n^0 + e\mathcal{E} \cdot \mathbf{r}_{kn}^{(0)} + e^2 \sum_{k\neq n} \frac{\left(\mathcal{E} \cdot \mathbf{r}_{nk}^{(0)}\right)\left(\mathcal{E} \cdot \mathbf{r}_{kn}^{(0)}\right)}{E_n^0 - E_k^0}. \tag{12.84}$$

The first-order correction is known as the Stark effect.

12.5 Problems

1. Prove that the energy and wave function corrections, to second-order of the perturbation theory for non-degenerate states, are:

$$E_n^{(2)} = \sum_{k\neq n} \frac{|U_{k,n}^{(0)}|^2}{\left(E_n^{(0)} - E_k^{(0)}\right)} \equiv \sum_{k\neq n} \frac{|U_{k,n}^{(0)}|^2}{\hbar\omega_{nk}}, \tag{12.85}$$

and

$$|n\rangle^{(2)} = \sum_{l,k\neq n} \frac{U_{l,k}^{(0)} U_{k,n}^{(0)}}{\hbar^2 \omega_{nk}\omega_{nl}} |l\rangle_0 - \sum_{k\neq n} \frac{U_{n,n}^{(0)} U_{k,n}^{(0)}}{\hbar^2 \omega_{nk}^2} |k\rangle_0 - \frac{1}{2} \sum_{k\neq n} \frac{|U_{k,n}^{(0)}|^2}{\hbar^2 \omega_{nk}^2} |n\rangle_0. \tag{12.86}$$

2. Show that, in the perturbation theory of degenerate states, independent of time, with degeneracy degree 2, the following identities hold:

$$\frac{|U_{1,2}|^2 + (U_{1,1} - E_{n_1}^{(1)})^2}{(U_{1,1} - E_{n_1}^{(1)})} = \sqrt{(U_{1,1} - U_{2,2})^2 + 4U_{1,2}U_{2,1}}; \tag{12.87}$$

$$\frac{|U_{1,2}|^2 + (U_{1,1} - E_{n_2}^{(1)})^2}{(U_{1,1} - E_{n_2}^{(1)})} = -\sqrt{(U_{1,1} - U_{2,2})^2 + 4U_{1,2}U_{2,1}}. \tag{12.88}$$

3. Prove that, in the perturbation theory of degenerate states, independent of time, with degeneracy degree 2, the coefficients of the normalized linear combination of degenerate states, for $E_{n_2}^{(1)}$, are:

$$|a_{n_2 1}|^2 = \frac{1}{2}\left[1 - \frac{U_{1,1} - U_{2,2}}{\sqrt{(U_{1,1} - U_{2,2})^2 + 4U_{1,2}U_{2,1}}}\right] \tag{12.89}$$

and

$$|a_{n_2 2}|^2 = \frac{1}{2}\left[1 + \frac{U_{1,1} - U_{2,2}}{\sqrt{(U_{1,1} - U_{2,2})^2 + 4U_{1,2}U_{2,1}}}\right]. \tag{12.90}$$

4. Given the first order corrections to the energy of unperturbed degenerate states and the linear combinations, obtain the energy corrections to second order of η.
5. Prove that, with the approximations described in Sect. 12.3, the transition probability in (12.59) reduces to the Fermi golden rule

$$\mathrm{W}_n = \frac{2\pi\left|{}_0\langle\psi_k\,|U|\psi_n\rangle_0\right|^2}{\hbar}\,\rho(E_k^{(0)}). \tag{12.91}$$

6. Show that replacing the function $\sin^2(\omega_{kn}t/2)/\omega_{kn}^2 t$ by $(\pi/2)\delta(E_n^{(0)} - E_k^{(0)})$ in (12.59) one obtains

$$\mathrm{W}_n = \frac{2\pi\left|{}_0\langle\psi_n\,|U|\psi_n\rangle_0\right|^2}{\hbar}\rho(E_n^{(0)}). \tag{12.92}$$

Chapter 13
Identical Particles, Bosons and Fermions

13.1 Introduction

The distiguishability or not of quantum processes under an interchange operator $\widehat{P}_{ij}$ of two particles and the interference phenomena of the wave function amplitudes, are fundamental issues of quantum theory. We will see that, for systems of identical particles, the invariance of the Hamiltonian under the action of the operator $\widehat{P}_{ij}$ implies symmetry conditions on the wave functions that are symmetric under $\widehat{P}_{ij}$ when it describes integer-spin particles (bosons), and antisymmetric, when it describes half-integer spin particles (fermions). Related with these symmetries, different and amazing properties of nature manifest when two or more identical particles with equal spins are observed. We will comment some consequences related with these symmetries: consequences and properties that can be explained within the quantum theory.

13.2 Distinguishable and Indistinguishable Quantum Processes

To visualize the properties that we want to discuss, let us suppose that we have a scattering process of two quantum particles, a and b which can be distinguished. In Fig. 13.1 we show two processes, seen from the center of mass system.[1] In the process of the left hand figure, the particle a is registered by the counter C_1 and the particle b by the counter C_2. When the counters C_1 and C_2 distinguish the particle's identity, the processes are distinguishable. We will represent the probability amplitude for the particle a to be scattered an angle θ as[2]

---

[1] An observer in the center of mass, will see that two particles, approaching or distancing, move one opposite to the other before and after the scattering, independent of the scattering angle.

[2] Generally $\varphi(\theta) = \varphi(-\theta)$ and $\varphi(\pi - \theta) = \varphi(\theta - \pi)$.

P. Pereyra, *Fundamentals of Quantum Physics*, Undergraduate Lecture Notes in Physics,
DOI: 10.1007/978-3-642-29378-8_13,

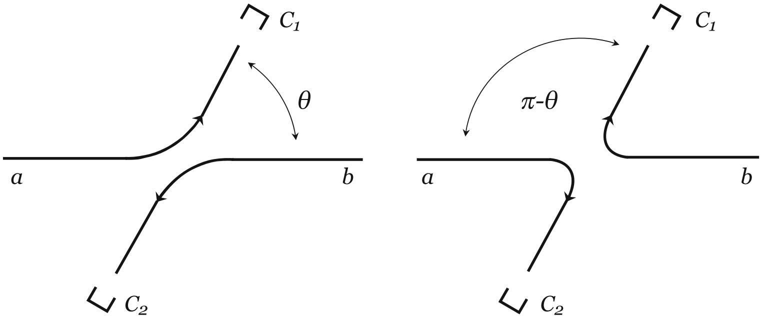


Fig. 13.1 Collision of a particle a with a particle b seen from their center of mass. The particle counters C_1 and C_2 register the scattered particles along the θ y $\pi - \theta$ directions

$$\varphi_a(\theta) = \langle 1|a\rangle. \tag{13.1}$$

This is equal to the probability amplitude for b to be scattered along the $\theta - \pi$ direction. Therefore, the probability that the counter C_1 registers the particle a, i.e., the probability for the left side process to occur, is

$$P_{a1} = |\varphi_a(\theta)|^2. \tag{13.2}$$

If the counter C_1 registers the particle b, the right hand side process occurred. If the probability amplitude for the particle a to be scattered along the $\theta - \pi$ direction, i.e. to be registered by the counter C_2, is $\varphi_a(\theta - \pi)$, the probability for the occurrence of the right hand side process is

$$P_{a2} = |\varphi_a(\pi - \theta)|^2 = |\varphi_b(\theta)|^2 = P_{b1}. \tag{13.3}$$

We could be less restrictive and ask ourselves for the probability P that the counter C_1 registers either the particle a or the particle b, i.e. for the probability that either the left or the right hand side process occurs. In this case, we have the probability for occurrence of the left hand side process, plus the probability for the occurrence of the right hand side process, i.e.

$$P = P_{a1} + P_{a2} = |\varphi_a(\theta)|^2 + |\varphi_a(\pi - \theta)|^2. \tag{13.4}$$

The occurrence of one event does not interfere with the occurrence of the other. We have a specific realization of the experiment just discussed, when we have, for example, the scattering of an α particle by a proton or any nuclei, different from α. The probability of having the particle α in C_1, and the proton in C_2, is

$$P_{\alpha 1} = |\varphi_\alpha(\theta)|^2. \tag{13.5}$$

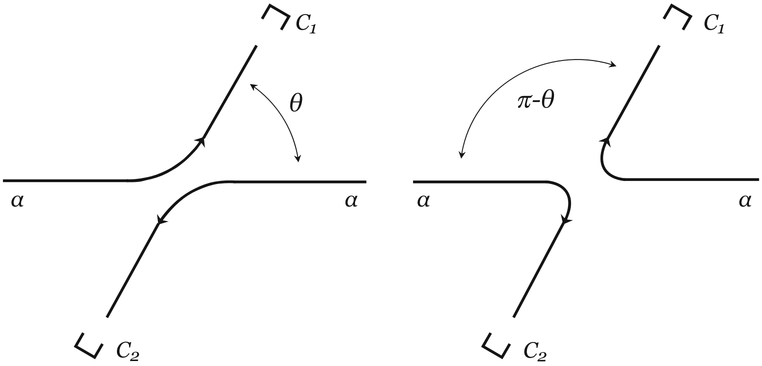


Fig. 13.2 Collision of two particles α seen from their center of mass. The particle counters C_1 and C_2 register the scattered particles along the θ and $\pi - \theta$ directions

What happens when the particles are *equal*, say both alphas as in Fig. 13.2? In this case, the processes at the left and right, are not distinguishable any more. If a particle α is detected in C_1, we do not know whether it comes from the left or from the right. Any of the two processes could have happened. The probability amplitude to detect an α in C_1 has two contributions: the probability amplitude that the alpha particle was scattered along θ and the probability amplitude that it was scattered along $\theta - \pi$, i.e.

$$\varphi_{\alpha 1} = \varphi_\alpha(\theta) + \varphi_\alpha(\pi - \theta); \tag{13.6}$$

both amplitudes *interfere* and the probability of detecting a particle α in C_1[3] is now

$$P_{\alpha 1} = |\varphi_\alpha(\theta) + \varphi_\alpha(\pi - \theta)|^2. \tag{13.7}$$

If in these experiments the scattered particles are identical, for example alphas, or electrons, etc., but they possess some quantum property like the spin projection that the counters can distinguish, the previous analysis holds. If the spin projections are different, the situation is analogous to that of particles a and b and the processes will be distinguishable. If their spin projections are equal, the particles are identical and the processes will be indistinguishable. However, it is important to notice that talking of scattering processes is not the same as talking of the occupation probability of a quantum state by identical particles with spin. The spin in this case gives rise to important physical differences. We will briefly comment those properties in the following sections.

---

[3] Which is equal to the probability of detecting an alpha particle in C_2.

13.3 Bosons and Fermions

To describe a system with N identical particles, we will use functions like

$$\begin{aligned}\phi(\xi_1,\xi_2,\ldots,\xi_N) &= \varphi(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N)\chi(s_1,s_2,\ldots,s_N)\\ &= \phi_1(\mathbf{r}_1,s_1)\phi_2(\mathbf{r}_2,s_2),\ldots,\phi_N(\mathbf{r}_N,s_N)\\ &= \phi_1(\xi_1)\phi_2(\xi_2),\ldots,\phi_N(\xi_N). \end{aligned} \tag{13.8}$$

Although it is clear that, when the particles are identical we can not label them, we can suppose (temporarily) that we have in the state 1 the particle 1, in the state 2 the particle 2,... etc. Because the particles are identical, it may happen that the particle 1 is in the state 2, the particle 2 in the state 1, the particle 3 in the state 3,... etc. This leads us to conclude that the most general representation for a system with N identical particles is obtained when we sum up all possible permutations of the coordinates ξ_j of the wave function $\phi(\xi_1,\xi_2,\ldots,\xi_N)$, i.e., the most general function to describe a system of N identical particles is

$$\Phi(\xi_1,\xi_2,\ldots,\xi_N) = C\sum_P \widehat{P}_\xi\big[\phi_1(\xi_1)\phi_2(\xi_2),\ldots,\phi_N(\xi_N)\big], \tag{13.9}$$

with P_ξ the permutation operator of the coordinates ξ_j. C is a normalization constant. When we have identical particles, and we interchange the position and spin of any two particles, the system does not change. This symmetry under the interchange of two particles must be reflected by the invariance of the function $\phi(\xi_1,\xi_2,\ldots,\xi_N)$. If $\widehat{P}_{kl}$ is the operator that interchanges the particles k and l, we have in principle the transformation

$$\widehat{P}_{kl}\phi(\xi_1,\xi_2,\ldots,\xi_k,\ldots,\xi_l,\ldots\xi_N) = e^{i\delta}\phi(\xi_1,\xi_2,\ldots,\xi_l,\ldots,\xi_k,\ldots\xi_N); \tag{13.10}$$

which returns to the original state when the operator acts twice, thus $e^{i2\delta}=1$. Hence

$$e^{i\delta} = \pm 1. \tag{13.11}$$

Combining (13.10) and (13.11), we conclude that on one side we can have

$$\phi(\xi_1,\xi_2,\ldots,\xi_k,\ldots,\xi_l,\ldots\xi_N) = \phi(\xi_1,\xi_2,\ldots,\xi_l,\ldots,\xi_k,\ldots\xi_N), \tag{13.12}$$

in which case the wave function is symmetric under the interchange of two particles, and on the other

$$\phi(\xi_1,\xi_2,\ldots,\xi_k,\ldots,\xi_l,\ldots\xi_N) = -\phi(\xi_1,\xi_2,\ldots,\xi_l,\ldots,\xi_k,\ldots\xi_N), \tag{13.13}$$

and the function is antisymmetric under the interchange of two particles. It is a fact that particles with integer spin (like photons, alpha particles, etc.) are described by symmetric functions; and those with half-integer spin (like electrons, protons, neutrons, etc.) are described by antisymmetric functions. This sign difference has enormous consequences on how the nature is built. Particles with integer spin are called generically *bosons* and those with half-integer spin, *fermions*.[4] We shall see, a little later, some important consequences as the Pauli's exclusion principle and the Bose–Einstein condensation. We know from matrix algebra, what the matrix determinant and the matrix permanent operations mean.[5] Using these operations, it is not difficult to verify that the sum of terms $\Phi(\xi_1, \xi_2, \ldots, \xi_N)$ in (13.9) can be written as the permanent (to describe bosons) or the determinant (to describe fermions) of

$$\begin{pmatrix} \phi_1(\xi_1) & \phi_1(\xi_2) & \ldots & \phi_1(\xi_N) \\ \phi_2(\xi_1) & \phi_2(\xi_2) & \ldots & \phi_2(\xi_N) \\ & & \cdot & \\ & & \cdot & \\ & & \cdot & \\ \phi_N(\xi_1) & \phi_N(\xi_2) & \ldots & \phi_N(\xi_N) \end{pmatrix}. \qquad (13.14)$$

We will briefly analyze some consequences of these symmetries.

13.3.1 Bose–Einstein Condensation and the Pauli Exclusion Principle

To simplify the analysis, we will consider a system of two particles. Let us start with two particles with integer spins, i.e. two bosons. We know that the most general wave function in this case is the superposition

$$\Phi(\xi_1, \xi_2) = \frac{1}{\sqrt{2!}}\big(\phi_1(\xi_1)\phi_2(\xi_2) + \phi_1(\xi_2)\phi_2(\xi_1)\big). \qquad (13.15)$$

This function is symmetric under the interchange operator P_{12}. The system could be one of two photons, within a space region limited by two mirrors, as shown in Fig. 13.3. Suppose also that these photons are created one after the other. When the first photon, photon $f1$, is created, the probability amplitude to find it, in some point $\mathbf{r}_1$ between the mirrors, is $\phi_1(\xi_1) = \varphi_1(\mathbf{r}_1)\chi(s_1)$, and the probability density is

$$p_{1f} = |\varphi_1(\mathbf{r}_1)\chi(s_1)|^2 = |\phi(\mathbf{r}_1, s_1)|^2. \qquad (13.16)$$

---

[4] W. Pauli, using field theory dealt with this issue.

[5] The matrix permanent develops like the matrix determinant, but without changes of sign.

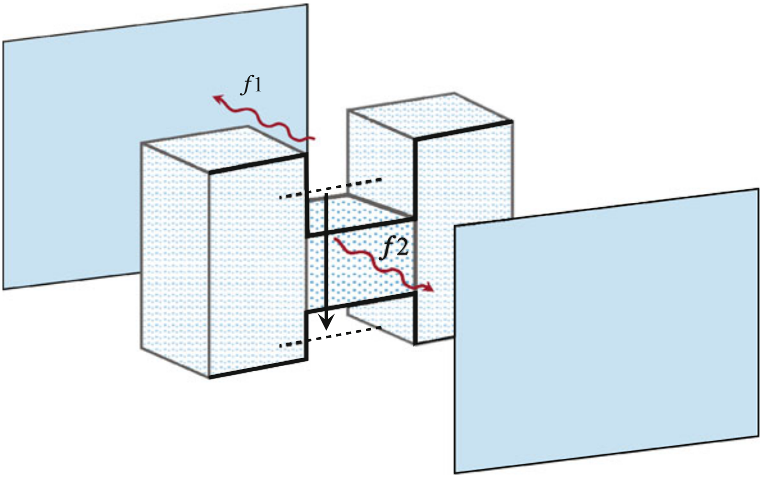


Fig. 13.3 In a system like the one shown here, photons are created by the recombination of conduction-band electrons with holes (absences) of the valence band of a quantum well. The excited electrons in the conduction band occupy the quantum well states according to the Pauli exclusion principle and the Fermi–Dirac statistics. Photons moving between mirrors fulfill the Bose–Einstein statistics

We will interpret this, as the probability to create the first photon. When the second photon is created, the function that describes the two photons is the function $\Phi(\xi_1, \xi_2)$ given in (13.15). If the particles are identical,

$$\phi_1(\xi_1) = \phi_2(\xi_2). \tag{13.17}$$

The wave function or probability amplitude for the two photons $\Phi(\xi_1, \xi_2)$ is then written as

$$\Phi(\xi_1, \xi_2) = \frac{2}{\sqrt{2!}}\phi_1^2(\xi_1) = \sqrt{2!}\varphi_1^2(\mathbf{r}_1)\chi^2(s_1). \tag{13.18}$$

Therefore, the probability of finding the photon $f1$ at $\mathbf{r}_1$ and the photon $f2$ at $\mathbf{r}_2$ is

$$p_{2f} = 2|\varphi_1(\mathbf{r}_1)\chi(s_1)|^2|\varphi_1(\mathbf{r}_1)\chi(s_1)|^2 = 2p_{1f}p_{1f}. \tag{13.19}$$

This result can be interpreted as the product of the probability $p_{f1} = p_{1f}$ to create the photon $f1$ (in the empty state) times the probability $p_{f2} = 2p_{1f}$ to create the photon $f2$, when the photon $f1$ already exists. This probability, p_{f2}, is twice the probability p_{f1}. If we generalize this argument[6] for an arbitrary number of photons, we conclude that the probability density for creating one photon when $n-1$ photons

---

[6] A detailed discussion on this issue, can be found in Richard P. Feynman, Robert Leighton and Matthew Sands, *Lectures on Physics* (Addison-Wesley, 1964).

exist already, is

$$p_{nf} = np_{(n-1)f} = n!p_{1f}^{n}. \tag{13.20}$$

This probability is n times larger than the probability of having one more photon when $n-2$ already exist, and so on. This, without doubt, is an interesting result that is behind the Bose–Einstein condensation phenomenon[7] and the superfluidity phenomenon discovered in 1938 by Kapitsa, Allen and Misener.[8]

Suppose now that we have two fermionic particles, for example, two electrons in the bound energy levels of the quantum well in Fig. 13.3. Suppose also, for a moment, that they were not identical. Due to the symmetry properties of the fermion wave functions, the two electrons would be described by the antisymmetric function

$$\Phi(\xi_1, \xi_2) = \frac{1}{\sqrt{2!}}\big(\phi_1(\xi_1)\phi_2(\xi_2) - \phi_1(\xi_2)\phi_2(\xi_1)\big). \tag{13.21}$$

If the particles are identical, the function $\Phi(\xi_i, \xi_j)$ becomes zero. This means that *it is not possible to have two identical fermions in the same quantum state*. This is precisely the Pauli exclusion principle, a fundamental property of nature; the recipe to build and distinguish atoms. The exclusion principle allows quantum theory to explain the electronic configuration and the periodic table. Let us now introduce a final comment on the normalization constant of the general solution (13.14). If we have N bosons, with N_1 particles in the state 1, N_2 in the state 2, etc., the wave function $\Phi(\xi_1, \xi_2, \ldots, \xi_N)$ in (13.9) contains $N!/N_1!N_2!\ldots$ different terms. If we instead have N fermions, the wave function $\Phi(\xi_1, \xi_2, \ldots, \xi_N)$ in (13.9) will be the determinant of (13.14) with $N!$ terms. If the wave functions $\phi_k(\xi)$ are normalized, the normalization constant C in (13.9) will be $C = \sqrt{N_1!N_2!\ldots/N!}$ for bosons and $C = 1/\sqrt{N!}$ for fermions.

13.3.2 Bose–Einstein and Fermi–Dirac Statistics

One of the variables that one needs quite frequently in applications is the occupation probability or the average number of particles $\overline{n}$ in quantum states. In a system of many particles, described by a Hamiltonian $\widehat{H}$, the stationary quantum states φ_j and their corresponding energies E_j are solutions of

$$\widehat{H}\varphi_j = E_j\varphi_j, \quad j = 1, 2, 3, \ldots \tag{13.22}$$

---

[7] Eric Cornell and Carl Wieman, in 1995, have shown that cooling at 0.17 μK, the rubidium gas undergoes to the Bose–Einstein condensate state.

[8] Pyotr Kapitsa, John Allen and Don Misener discovered that He^4 passes to the superfluid state when the temperature reaches 2.17 K.

In a system of identical particles, it is fundamental to determine the number of particles in each of the available quantum states, i.e. the set of occupation numbers

$$\{n_j\} = (n_1, n_2, \ldots, n_j, \ldots). \tag{13.23}$$

With this information it is possible to evaluate a number of physical quantities, for example, the energy of the whole system

$$E = \sum_j E_j n_j. \tag{13.24}$$

The Bose–Einstein and the Fermi–Dirac statistics differ precisely in the occupation numbers. We will derive the average occupation numbers $\overline{n}$ in both cases. The reader can find, in the standard statistical theory courses, different derivations of $\overline{n}$. When the system is in contact with a heat bath, at temperature T, that works as a reservoir of particles, characterized by a chemical potential μ, one uses for the statistical evaluation the partition function

$$Z_N = \sum_{\{n_j\}} e^{-(E_1 n_1 + E_2 n_2 + \ldots)/k_B T}; \quad \text{with} \quad N = \sum_j n_j. \tag{13.25}$$

The sum here extends over all sets of occupation numbers $\{n_j\}$, compatible with $N = \sum_j n_j$. With this function one can obtain the grand partition function (of the grand canonical ensemble)

$$\Xi = \sum_{N=0} e^{-N\mu/k_B T} Z_N = \sum_{n_1} e^{(\mu - E_1)n_1/k_B T} \sum_{n_2} e^{(\mu - E_2)n_2/k_B T} \ldots \tag{13.26}$$

To perform the sums we have to take into account all possible values of n_j, in each case. In the Bose–Einstein statistics n_j can be any natural number and zero; this means that for each factor of the last equation we have

$$\sum_{n=0}^{\infty} e^{(\mu - E_j)n_j/k_B T} = \frac{1}{1 - e^{(\mu - E_j)/k_B T}}, \tag{13.27}$$

but in the Fermi–Dirac statistics, n_j can be only 0 or 1. As a consequence

$$\sum_{n=0}^{\infty} e^{(\mu - E_j)n_j/k_B T} = 1 + e^{(\mu - E_j)/k_B T}. \tag{13.28}$$

If we assign the sign $-$ to Bose–Einstein statistics and the sign $+$ to Fermi–Dirac statistics, we can write the previous results in compact form as

$$\sum_{n=0}^{\infty} e^{(\mu-E_j)n_j/k_BT} = \left[1 \mp e^{(\mu-E_j)/k_BT}\right]^{\mp 1}, \tag{13.29}$$

thus the grand partition function becomes

$$\Xi = \prod_j \left[1 \mp e^{(\mu-E_j)/k_BT}\right]^{\mp 1}. \tag{13.30}$$

Using this partition function it has been possible to derive, in the statistical theory, the occupation probability

$$\wp(n_j) = \frac{e^{(\mu-E_j)n_j/k_BT}}{\sum_{n_j} e^{(\mu-E_j)n_j/k_BT}}, \tag{13.31}$$

of the state φ_j. Therefore, the average number $\overline{n}_j$ of particles in the state φ_j is given by

$$\begin{aligned} \overline{n}_j &= \sum_{n_j} n_j \wp(n_j) = \frac{\sum_{n_j} n_j e^{(\mu-E_j)n_j/k_BT}}{\sum_{n_j} e^{(\mu-E_j)n_j/k_BT}} \\ &= k_BT \frac{\partial}{\partial\mu} \sum_{n_j} e^{(\mu-E_j)n_j/k_BT}. \end{aligned} \tag{13.32}$$

If we use the relation (13.29), the average occupation number is given also by

$$\overline{n}_j = k_BT \frac{\partial}{\partial\mu}\left[1 \mp e^{(\mu-E_j)/k_BT}\right]^{\mp 1}, \tag{13.33}$$

which leads, on one side (for bosons) to

$$\overline{n}_j = \frac{1}{e^{(E_j-\mu)/k_BT} - 1}, \tag{13.34}$$

a distribution similar to the Planck distribution, and on the other side (for fermions) to

$$\overline{n}_j = \frac{1}{e^{(E_j-\mu)/k_BT} + 1}. \tag{13.35}$$

Both distributions reduce, in the classical limit of high temperatures, to the Boltzmann distribution. Another important limit is the continuum limit, in which the occupied volume is so large that the single particle states are more and more densely distributed. For a large fermion system, the number of states $\Delta\mathcal{N}$, with energies between E and $E + \Delta E$, can be expressed as

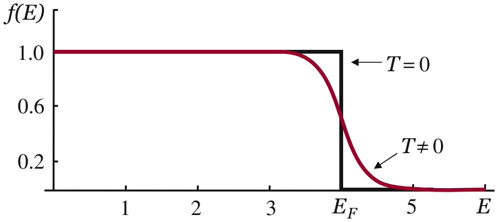


Fig. 13.4 The Fermi distribution for $T = 0$ and for $T \neq 0$. When the system temperature is $T = 0$, all levels below the Fermi energy are occupied with probability 1. Increasing the temperature, the occupation probability for energy levels above the Fermi energy E_F increases, while for energy levels below E_F diminishes. The occupation probability at the *Fermi level* is $1/2$. At $T = 0$ the Fermi energy and the Fermi level coincide. The energy interval in which the distribution falls down from 1 to 0 is proportional to the thermal energy $k_B T$

$$\Delta\mathcal{N} = \mathcal{D}(E)\Delta E, \tag{13.36}$$

where $\mathcal{D}(E)$ is the density of states. The occupation probability of the energy level E is given, in this limit, by the Fermi–Dirac distribution

$$f(E) = \frac{1}{e^{(E-E_F)/k_B T} + 1}. \tag{13.37}$$

In this expression the chemical potential μ was substituted by the "Fermi level" E_F, a concept most used in physics. This distribution function depends also on the temperature T. In Fig. 13.4 we plot the function $f(E)$ for $T = 0$ and for $T \neq 0$. It is easy to verify that when $T = 0$, the function $f(E) = 1$ for energies $E < E_F$, and $f(E) = 0$ when $E > E_F$. It is also clear that when the energy E equals the Fermi energy[9], $f(E) = 1/2$. This means that, at temperature $T = 0$, all states with energies less than E_F are occupied and those with higher energies are unoccupied.

If the temperature is different from zero, the distribution function shown in Fig. 13.4 has a continuous variation close to the Fermi level. In an energy interval of the order of the thermal energy $k_B T$, it changes from an occupation probability that is almost 1 to an occupation probability that tends to 0 as the energy E increases. At the Fermi energy the occupation probability is $1/2$. The behavior of the Fermi–Dirac distribution function, when $T \neq 0$, shows that the occupation probability for the energy levels close to E_F is less than 1. This behavior of the occupation probability is a consequence of the thermal excitation of fermions above the Fermi level.

Using the density of states $\mathcal{D}(E)$ and the Fermi distribution $f(E)$ one can evaluate the mean energy and the number of particles

---

[9] When $T = 0$, it is common to call Fermi energy to Fermi level.

$$\overline{E} = \int E f(E)\mathcal{D}(E)dE, \tag{13.38}$$

$$N = \int f(E)\mathcal{D}(E)dE. \tag{13.39}$$

These are important functions of statistical physics, the solid state physics and the semiconductor physics which, in general, deal with systems of many particles, especially with many electron systems.

13.4 The Effect of the Statistics on the Energy

Explaining the electronic configuration of atoms was one of the problems where the spin has been relevant. In Chap. 10, we mentioned the need of using the Pauli exclusion principle and the electron's spin. To understand a bit more the effect of the statistics on the behavior of a system of particles, and, on the system's energy, we will consider the Helium atom and the problem of two electrons in a one-dimensional potential well. In both systems, the Hamiltonian is independent of the spin, and the wave function for two electrons can be factored in a coordinates-dependent function and a spin-dependent factor, as follows:

$$\phi_{n_1,n_2}(\xi_1, \xi_2) = \varphi_{n_1,n_2}(x_1, x_2)\chi_{m_{s1},m_{s2}}. \tag{13.40}$$

Prior to discussing the Schrödinger equation, let us consider the spin-dependent factor for a two-particle system.

13.4.1 Spin States for Two Spin 1/2 Particles

We saw before that for a particle with spin s_1, and spin projection m_{s1} along the z axis, we need to introduce, besides the operator $\widehat{\mathbf{S}}_1$, the operators $\widehat{S}_1^2$ and $\widehat{S}_{1z}$, which satisfy the eigenvalues equations

$$\widehat{S}_1^2|s_1 m_{s1}\rangle = \hbar^2 s_1(s_1+1)|s_1 m_{s1}\rangle, \tag{13.41}$$

$$\widehat{S}_{1z}|s_1 m_{s1}\rangle = \hbar m_{s_1}|s_1 m_{s1}\rangle, \quad \text{with} \quad |m_{s1}| \le s_1. \tag{13.42}$$

In a two-particle system with spins s_1 and s_2 and spin operators $\widehat{\mathbf{S}}_1$ and $\widehat{\mathbf{S}}_2$, we can determine the spin $\widehat{\mathbf{S}}$ of the whole system using the angular momenta addition rules, that were sketched in Chap. 11. In this case we have

$$\widehat{\mathbf{S}} = \widehat{\mathbf{S}}_1 + \widehat{\mathbf{S}}_2; \tag{13.43}$$

with the corresponding sum for the spin projections

$$\widehat{S}_z = \widehat{S}_{z1} + \widehat{S}_{z2}. \tag{13.44}$$

If we denote the spin state of the two-particle system as

$$|s_1 s_2 m_{s1} m_{s2}\rangle = |s_1 m_{s1}\rangle |s_2 m_{s2}\rangle, \tag{13.45}$$

it is clear that

$$m_s = m_{s1} + m_{s2}. \tag{13.46}$$

The operators $\widehat{\mathbf{S}}$, $\widehat{S}^2$ and $\widehat{S}_z$ satisfy the analogous commutation relations as those for $\widehat{\mathbf{L}}$, $\widehat{L}^2$ and $\widehat{L}_z$. Moreover, the operators $\widehat{S}^2$, $\widehat{S}_z$, $\widehat{S}_1^2$ and $\widehat{S}_2^2$ commute among them. The operators $\widehat{S}^2$ and $\widehat{S}_z$ commute also with the interchange operator $\widehat{P}_{1,2}$, and with the Hamiltonian. This means that, for well defined values of s_1 and s_2, the eigenfunctions should be represented by functions characterized by the quantum numbers s, m_s, s_1 and s_2, or alternatively by states defined by the quantum numbers s_1, m_{s1}, s_2 and m_{s2}, that make reference to the electrons' spin projections.[10] In the Dirac notation we have, in one case,

$$\psi_{sm_s} = |sm_s s_1 s_2\rangle, \tag{13.47}$$

and

$$\chi_{m_{s1},m_{s2}} = |s_1 s_2 m_{s1} m_{s2}\rangle, \tag{13.48}$$

in the other. For particles with spin 1/2 it is possible to make the notation even lighter. We can use, for example, just the signs of the spin projections $m_{s1} = \pm 1/2$ and $m_{s2} = \pm 1/2$. In this case we have

$$\chi_{\pm,\pm} = |s_1 s_2 \pm 1/2 \pm 1/2\rangle = \chi_\pm \chi_\pm. \tag{13.49}$$

---

[10] Both representations are possible and one can establish the relation between them. This relation is a particular case of the transformation

$$\begin{aligned} |sm_s s_1 s_2\rangle &= \sum_{m_{s1}+m_{s2}=m_s} |s_1 s_2 m_{s1} m_{s2}\rangle \langle s_1 s_2 m_{s1} m_{s2} || sm_s s_1 s_2\rangle \\ &= \sum_{m_{s1}+m_{s2}=m_s} C^{s_1,s_2,s}_{m_{s1},m_{s2},m_s} |s_1 s_2 m_{s1} m_{s2}\rangle. \end{aligned}$$

The coefficients $C^{s_1,s_2,s}_{m_{s1},m_{s2},m_s}$ are known as the Clebsch-Gordan coefficients. The interested reader can find more details, for example, in A.R. Edmonds *Angular Momentum in Quantum Mechanics*, (Princeton University Press, Princeton, N. J. 1974).

If we represent the eigenfunctions of $\widehat{S}^2$ and $\widehat{S}_z$ with $|sm_sm_{s1}m_{s2}\rangle$, the eigenvalue equations of $\widehat{S}^2$ and $\widehat{S}_z$ will read

$$\widehat{S}^2|sm_sm_{s1}m_{s2}\rangle = \hbar^2 s(s+1)|sm_sm_{s1}m_{s2}\rangle, \tag{13.50}$$

$$\widehat{S}_z|sm_sm_{s1}m_{s2}\rangle = \hbar m_s|sm_sm_{s1}m_{s2}\rangle, \quad \text{with} \quad |m_s| \leq s. \tag{13.51}$$

As mentioned before, the largest value that the spin s can take is $s_1 + s_2$ and the smallest $|s_1 - s_2|$. This means that in our two particle system, with spins $s_1 = s_2 = 1/2$, the spin quantum numbers are

$$s = 0, 1 \quad \text{and} \quad m_s = -s, -s+1, \ldots, s. \tag{13.52}$$

When the spins are parallel, the total spin is $s = 1$, while $s = 0$ when the spins are antiparallel. For the state with $s = 1$, we have three possible spin components along the z axis: $m_s = -1, 0, 1$. For this reason, this spin state is called a *triplet*. Since $m_s = m_{s1} + m_{s2}$, we can infer that the state ψ_{11} is equivalent to the state χ_{++} with m_{s1} =1/2 and m_{s2} =1/2. Therefore

$$\psi_{1,1} = \chi_{++}. \tag{13.53}$$

In the same way

$$\psi_{1,-1} = \chi_{--}. \tag{13.54}$$

These states are symmetric under the interchange of particles. The relation of χ_{+-} and χ_{-+} with the state $\psi_{1,0}$ is different. If $m_s = 0$, it is clear that m_{s1} should be $= 1/2$ and that m_{s2} =–1/2 or, alternatively, m_{s1} =–1/2 and m_{s2} =1/2. This means that the states χ_{+-} and χ_{-+} contribute to $\psi_{1,0}$. In this case we say that they interfere. Since the state must be symmetric, we write it as

$$\psi_{1,0} = \frac{1}{\sqrt{2}}\left(\chi_{+-} + \chi_{-+}\right). \tag{13.55}$$

We will see, at the end of this chapter, that an alternative procedure to obtain this state is leaving the lowering operator $\widehat{S}_-$ to act on the state $\psi_{1,1}$. Finally, what can we say on the state ψ_{00}, that corresponds to spin $s = 0$? The spin has only one component along the z axis, the state ψ_{00} must be antisymmetric and orthogonal to the triplet state with spin $s = 1$. From this condition, and since the spin projection is also zero, we have the antisymmetric state

$$\psi_{0,0} = \frac{1}{\sqrt{2}}\left(\chi_{+-} - \chi_{-+}\right), \tag{13.56}$$

called a *singlet* state.

13.4.2 Two Electrons in a Quantum Well

Let us suppose now that we have two electrons in an infinite quantum well. The Hamiltonian of this system can be written as

$$\widehat{H} = \frac{\widehat{p}_1^2}{2m} + \frac{\widehat{p}_2^2}{2m} + V(x_1, x_2) + \frac{e^2}{|x_2 - x_1|} = H_o + \frac{e^2}{|x_2 - x_1|}, \tag{13.57}$$

with $V(x_1, x_2) = 0$, for $0 < x_1, x_2$ and for $x_1, x_2 < L$, and $V(x_1, x_2) = \infty$ elsewhere. The Schrödinger equation inside the well is then

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_1^2} - \frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_2^2} + \frac{e^2}{|x_2 - x_1|}\right]\varphi(x_1, x_2) = E\varphi(x_1, x_2). \tag{13.58}$$

The electron–electron interaction is relatively weak and can be treated as a perturbation to H_o. In this case, the Schrödinger equation for the unperturbed Hamiltonian, reads

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_1^2} - \frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_2^2}\right]\varphi(x_1, x_2) = E\varphi(x_1, x_2). \tag{13.59}$$

It is easy to verify that this equation is formally equal to that of an electron in a two-dimensional well. The Eq. (13.59) is separable and the solutions of the separate equations are the same as for the one-dimensional infinite quantum well. Therefore, the unperturbed eigenfunctions of the two electrons system, inside an infinite quantum well, are functions like

$$\varphi^{(0)}_{n_1 n_2}(x_1, x_2) = \sin\left(\frac{n_1\pi}{L}x_1\right)\sin\left(\frac{n_2\pi}{L}x_2\right), \quad \text{with} \quad n_1, n_2 = 1, 2, 3, \ldots, \tag{13.60}$$

with eigenvalues given by

$$E^{(0)}_{n_1 n_2} = \frac{\hbar^2\pi^2}{2mL^2}\left(n_1^2 + n_2^2\right). \tag{13.61}$$

The two-electron system is a fermionic system and we have to include the spin effect. Thus, the wave function $\phi(\xi_1, \xi_2)$ should be written as

$$\phi_{n_1 n_2 s m_s}(\xi_1, \xi_2) = \varphi_{n_1 n_2, s}(x_1, x_2)\psi_{s m_s}. \tag{13.62}$$

From the first part of this chapter, we know that this function must be anti-symmetric under the particles-interchange operator. This symmetry is fulfilled if one of the functions $\varphi_{n_1 n_2 s}$ or $\psi_{s m_s}$ is symmetric and the other one is anti-symmetric, i.e., we must have either the function

$$\phi_{n_1 n_2 s m_s}(\xi_1, \xi_2) = \varphi^S_{n_1 n_2, s}(x_1, x_2)\psi^A_{s m_s}, \tag{13.63}$$

or the function

$$\phi_{n_1 n_2 s m_s}(\xi_1, \xi_2) = \varphi^A_{n_1 n_2, s}(x_1, x_2)\psi^S_{s m_s}. \tag{13.64}$$

This means that the independence of H_o from the spin variables is just apparent. The symmetry properties must, after all, be taken into account. For this reason we include the quantum number s to label the functions $\varphi_{n_1 n_2 s}(x_1, x_2)$. As will be seen lines below, the energies also depend on the spins, even though the Hamiltonian is spin-independent. In the lowest energy level $E_{1,1} = 2\hbar^2\pi^2/2mL^2$, only exists the symmetric function

$$\varphi_{11,s} = \sin\frac{\pi}{L}x_1 \sin\frac{\pi}{L}x_2, \tag{13.65}$$

since the anti-symmetric function vanishes. Hence, the spinor $\psi_{s m_s}$ associated to $\phi_{11 s m_s}(\xi_1, \xi_2)$ should be the singlet state

$$\psi_{s m_s} = \psi_{00} = \frac{1}{\sqrt{2}}\big(\chi_{+-} - \chi_{-+}\big). \tag{13.66}$$

Therefore, the eigenfunction corresponding to the first energy level E_{11} is

$$\phi_{1100}(\xi_1, \xi_2) = \frac{1}{\sqrt{2}} \sin\frac{\pi}{L}x_1 \sin\frac{\pi}{L}x_2\big(\chi_{+-} - \chi_{-+}\big). \tag{13.67}$$

In the second level the energy is $E_{12} = E_{21} = 5\hbar^2\pi^2/2mL^2$. The function that depends on the space coordinates can be symmetric or anti-symmetric. We plot in Fig. 13.5 the functions

$$\varphi^S_{12,s}(x_1, x_2) = \tfrac{1}{\sqrt{2}}\big[\varphi_1(x_1)\varphi_2(x_2) + \varphi_2(x_1)\varphi_1(x_2)\big], \tag{13.68}$$

and

$$\varphi^A_{12,s}(x_1, x_2) = \frac{1}{\sqrt{2}}\big[\varphi_1(x_1)\varphi_2(x_2) - \varphi_2(x_1)\varphi_1(x_2)\big]. \tag{13.69}$$

We know that the spinor that multiplies the function $\varphi^S_{12,s}(x_1, x_2)$ should be a singlet, with $s = 0$, while the spinor that multiplies to $\varphi^A_{12,s}(x_1, x_2)$ can be any of the triplet state spinors: ψ_{11}, ψ_{10} or ψ_{1-1}. It is easy to verify that the first energy level is a non-degenerate energy level, while the second is a degenerate energy level, with degeneracy degree equal to 4.

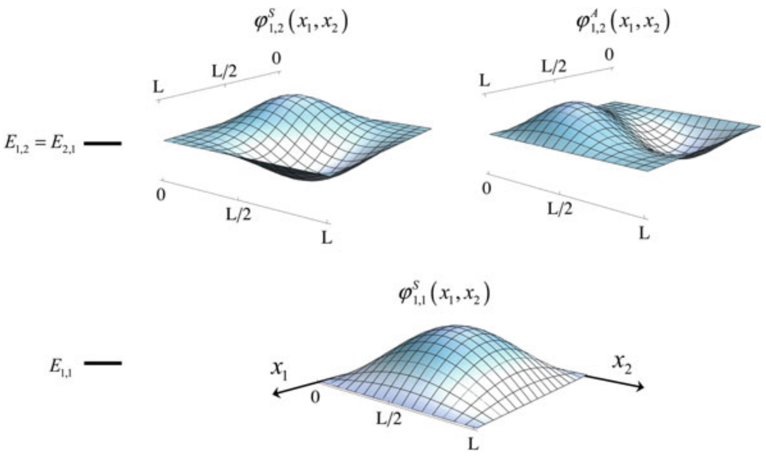


Fig. 13.5 The first two energy levels of two electrons inside an infinite quantum well. In the first level, the coordinates-dependent function is symmetric, thus the electrons coexist in this level only if their spins are antiparallel. In the second energy level, we plot the symmetric and the anti-symmetric functions, under the interchange of particles

13.4.3 The Helium Atom and the Exchange Energy

Another system with two electrons is the Helium atom. If we neglect the interactions that depend explicitly on the spins, and assume that the Helium atom nuclei is at rest, the two-electron Hamiltonian can be written in the form

$$\widehat{H} = \frac{\widehat{p}_1^2}{2m} + \frac{\widehat{p}_2^2}{2m} - \left(\frac{e^2}{r_1} + \frac{e^2}{r_2}\right) + \frac{e^2}{r_{12}}, \tag{13.70}$$

where $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$. This Hamiltonian remains invariant under the interchange operator $\widehat{P}_{12}$, but as mentioned before, the two-fermion system should be anti-symmetric. Thus it imposes symmetry requirements on the solution of the Schrödinger equation

$$\widehat{H}\varphi(\mathbf{r}_1, \mathbf{r}_2) = E\varphi(\mathbf{r}_1, \mathbf{r}_2). \tag{13.71}$$

If we treat the electron–electron interaction term e^2/r_{12} as a perturbation, we are left with a separable differential equation, and the resulting equations are similar to that of the Hydrogen atom with eigenfunctions

$$\phi_{n_1 l_1 m_l} \quad \text{and} \quad \phi_{n_2 l_2 m_l}, \tag{13.72}$$

and eigenvalues

$$E_{n_1} = -\frac{me^4}{2\hbar^2 n_1^2} \quad \text{and} \quad E_{n_2} = -\frac{me^4}{2\hbar^2 n_2^2}. \tag{13.73}$$

The angular momentum and spin operators $\widehat{\mathbf{L}}_i$ and $\widehat{\mathbf{S}}_i$, with $i = 1, 2$, do not commute with the interchange operator $\widehat{P}_{12}$. But the operators

$$\widehat{\mathbf{L}}^2 = \left(\widehat{\mathbf{L}}_1 + \widehat{\mathbf{L}}_2\right)^2, \quad \widehat{\mathbf{S}}^2 = \left(\widehat{\mathbf{S}}_1 + \widehat{\mathbf{S}}_2\right)^2, \tag{13.74}$$

and their corresponding projections $\widehat{L}_z$ and $\widehat{S}_z$, do commute with $\widehat{P}_{12}$. For this reason we use the quantum numbers n_1, n_2, l, s, m_l and m_s to label the states $\phi(\xi_1, \xi_2)$ of the Helium atom which, like in the previous example, we write in the form

$$\phi_{n_1 n_2 l s m_l m_s}(\xi_1, \xi_2) = \varphi_{n_1 n_2 l m_l s}(\mathbf{r}_1, \mathbf{r}_2)\psi_{s m_s}. \tag{13.75}$$

From our discussion in the first part of this chapter, we know that the function $\psi_{s m_s}$ is anti-symmetric if $s = 0$ and symmetric if $s = 1$. Therefore, to describe the two-fermion system we will have either

$$\phi_{n_1 n_2 l 0 m_l 0}(\xi_1, \xi_2) = \varphi^S_{n_1 n_2 l m_l}(\mathbf{r}_1, \mathbf{r}_2)\psi^A_{00}; \tag{13.76}$$

or

$$\phi_{n_1 n_2 l 1 m_l m_s}(\xi_1, \xi_2) = \varphi^A_{n_1 n_2 l m_l}(\mathbf{r}_1, \mathbf{r}_2)\psi^S_{1 m_s}; \tag{13.77}$$

with

$$\varphi^S_{n_1 n_2 l m_l}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}\Big(\varphi_{n_1 l_1 m_{l1}}(\mathbf{r}_1)\varphi_{n_2 l_2 m_{l2}}(\mathbf{r}_2) + \varphi_{n_1 l_1 m_{l1}}(\mathbf{r}_2)\varphi_{n_2 l_2 m_{l2}}(\mathbf{r}_1)\Big), \tag{13.78}$$

and

$$\varphi^A_{n_1 n_2 l m_l}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}\Big(\varphi_{n_1 l_1 m_{l1}}(\mathbf{r}_1)\varphi_{n_2 l_2 m_{l2}}(\mathbf{r}_2) - \varphi_{n_1 l_1 m_{l1}}(\mathbf{r}_2)\varphi_{n_2 l_2 m_{l2}}(\mathbf{r}_1)\Big), \tag{13.79}$$

respectively. For a given set of values of the quantum numbers n_1, n_2, l, m_l, all four spin states are degenerate. This degeneracy is removed when the interaction term e^2/r_{12} is taken into account.

To evaluate the first order correction, according to the perturbation theory we can use the functions $\phi_{n_1 n_2 l 0 m_l 0}(\xi_1, \xi_2)$ of the unperturbed Hamiltonian. If we calculate the matrix elements

$$\langle\phi_{n'_1 n'_2 l' s' m'_l m'_s}|\frac{e^2}{r_{12}}|\phi_{n_s n_2 l 1 m_l m_s}\rangle \tag{13.80}$$

we find that all of them are zero except those in the diagonal. It is easy to verify that these matrix elements can be written as

$$\langle \phi_{n_1'n_2'l's'm_l'm_s'}|\frac{e^2}{r_{12}}|\phi_{n_s n_2 l 1 m_l m_s}\rangle = E_D \pm E_I; \tag{13.81}$$

where

$$E_D = \int |\varphi_{n_1 l_1 m_{l1}}(\mathbf{r}_1)|^2 |\varphi_{n_2 l_2 m_{l2}}(\mathbf{r}_2)|^2 \frac{e^2}{r_{12}} d^3\mathbf{r}_1 d^3\mathbf{r}_2 \tag{13.82}$$

is the contribution known as the *direct energy*, and the energy term

$$E_I = \int \varphi^*_{n_1 l_1 m_{l1}}(\mathbf{r}_1)\varphi^*_{n_2 l_2 m_{l2}}(\mathbf{r}_2)\frac{e^2}{r_{12}}\varphi_{n_1 l_1 m_{l1}}(\mathbf{r}_2)\varphi_{n_2 l_2 m_{l2}}(\mathbf{r}_1) d^3\mathbf{r}_1 d^3\mathbf{r}_2, \tag{13.83}$$

that is known as the *exchange or interchange energy*. The signs ± come from $\varphi^S_{n_1 n_2 l m_l}(\mathbf{r}_1, \mathbf{r}_2)$ and $\varphi^A_{n_1 n_2 l m_l}(\mathbf{r}_1, \mathbf{r}_2)$, respectively. Since the integral E_I is generally positive, the state with $s = 0$ has higher energy than those with $s = 1$. This is an important statistics effect on the energy levels and on the electronic configuration.

To conclude this point, let us recall the coupling of two angular momentum vectors at the end of Chap. 11. Equation (11.84) applied to $\widehat{\mathbf{S}}_1$ and $\widehat{\mathbf{S}}_2$ reads

$$\widehat{S}^2 = \widehat{S}_1^2 + \widehat{S}_2^2 + 2\widehat{\mathbf{S}}_1 \cdot \widehat{\mathbf{S}}_2. \tag{13.84}$$

Since the signs of the exchange energy depends on the spin orientations, one can use the expectation value

$$\langle s s_1 s_2 m_s|\widehat{\mathbf{S}}_1 \cdot \widehat{\mathbf{S}}_2|s s_1 s_2 m_s\rangle = \hbar^2 \frac{s(s+1) - s_1(s_1+1) - s_2(s_2+1)}{2}, \tag{13.85}$$

with $s_1 = s_2 = 1/2$ to write the exchange term $\pm E_I$ as

$$\pm E_I = -\frac{1}{2}\Big(1 + \frac{4\langle \widehat{\mathbf{S}}_1 \cdot \widehat{\mathbf{S}}_2\rangle}{\hbar^2}\Big)E_I, \tag{13.86}$$

which makes explicit the relation between spin orientations and the signs. This representation was suggested by P.A.M. Dirac and is known as *Dirac's exchange operator*.

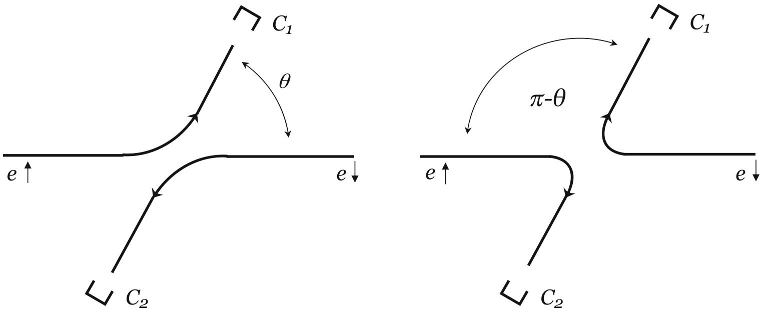


Fig. 13.6 Collision of two electrons with opposite spin projections seen from their center of mass. The particle counters C_1 and C_2 register the scattered particles along the directions θ and $\pi - \theta$

13.5 Solved Problems

Exercise 38 Suppose that we have the scattering of two electrons, whose spin projections may be known or not. Show that the probability $P_{e\uparrow i1}$ to scatter them along the direction $\theta = \pi/2$, when they are identical (indistinguishable) and the spins are known, is four times the probability $P_{e\uparrow d1}$ to scatter them, in the same direction, but when they are not identical (distinguishable). However, if the spins do not matter, the probability P_{ei1} to scatter one indistinguishable electron along $\theta = \pi/2$ is twice the probability P_{ed1} to scatter one distinguishable electron, along the same direction.

Solution Consider the processes shown in Fig. 13.6. It is supposed that: (1) the spin projections along the z axis are different and, (2) the particle counters can recognize the orientation of the spin projections. In this case, the processes are distinguishable and the situation is similar to that discussed lines above for particles a and b. If C_1 detects one electron with spin $\downarrow$, the probability amplitude and probability that the process on the right hand side occurred are

$$\varphi_{e\downarrow}(\pi - \theta), \quad \text{and} \quad P_{e\downarrow d1}(\pi - \theta) = |\varphi_{e\downarrow}(\pi - \theta)|^2, \tag{13.87}$$

respectively. We are using the letters d and i to indicate the distinguishability and indistinguishability of the processes. Similarly, if C_1 detects one electron with spin $\uparrow$, we know that the process on the left hand side occurred with probability amplitude and probability

$$\varphi_{e\uparrow}(\theta) \quad \text{and} \quad P_{e\uparrow d1} = |\varphi_{e\uparrow}(\theta)|^2. \tag{13.88}$$

Even though the processes are distinguishable, we can also ask for the probability of having one electron in C_1, independently of the spin orientation. In this case, we can have anyone of the two processes, thus the probability of having one electron in C_1 will be:

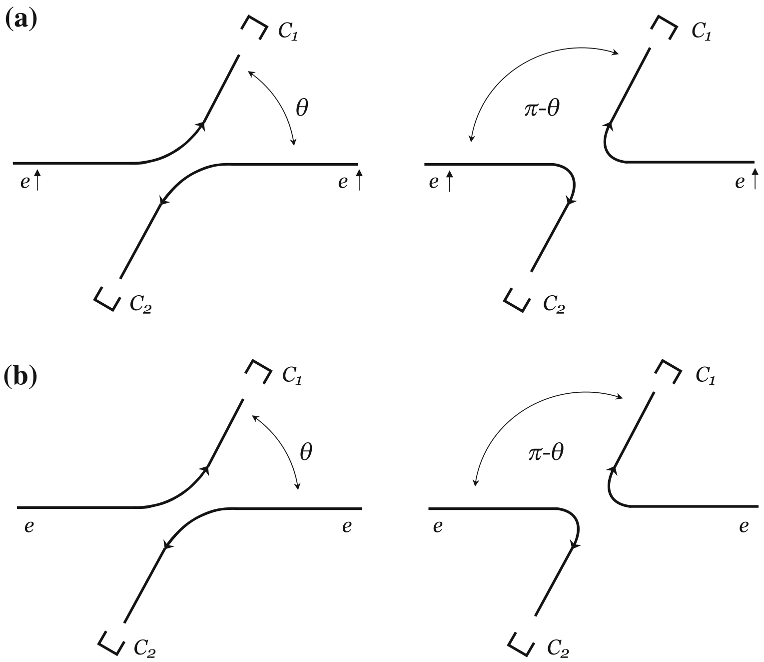


Fig. 13.7 Collision of two electrons seen from their center of mass. In (**a**) both particles have the same spin projection $m_s = 1/2$ and in (**b**) the spin projections are unknown. The particle counters C_1 and C_2 register the scattered particles along the directions θ and $\pi - \theta$

$$P_{ed1}(\theta) = |\varphi_e(\theta)|^2 + |\varphi_e(\pi - \theta)|^2 = P_{e\uparrow d1} + P_{e\downarrow d1}. \tag{13.89}$$

Lets us see now how these probability amplitudes and the probabilities change when the particles are identical, i.e., when all the observable quantum numbers are equal. If we have the processes in Fig. 13.7a, the electron entering into C_1 has, in any case, spin ↑ and we can not know whether it comes from the left or from the right. The processes are now indistinguishable and the probability amplitudes *interfere*, this means that the probability amplitude to detect one electron in C_1 is

$$\varphi_{e\uparrow}(\theta) + \varphi_{e\uparrow}(\pi - \theta). \tag{13.90}$$

A similar situation we have in a two-electron collision when their spin projections are unknown, as we show in Fig. 13.7b. In this case, the probability amplitude to detect one electron in C_1 will be also

$$\varphi_e(\theta) + \varphi_e(\pi - \theta). \tag{13.91}$$

In both cases the amplitudes interfere because the particles participating in the scattering processes, up to the level that we know them, are identical. The collision processes in the left and in the right of the Fig. 13.7a and b are *indistinguishable*. The probability of detecting one electron in C_1 of Fig. 13.7a is

$$P_{e\uparrow i1}(\theta) = |\varphi_{e\uparrow}(\theta) + \varphi_{e\uparrow}(\pi - \theta)|^2 \tag{13.92}$$

$$= |\varphi_{e\uparrow}(\theta)|^2 + |\varphi_{e\uparrow}(\pi - \theta)|^2 + 2\Re e\{\varphi^*_{e\uparrow}(\theta)\varphi_{e\uparrow}(\pi - \theta)\}, \tag{13.93}$$

while the probability to detect one electron in C_1 of Fig. 13.7b is

$$P_{ei1}(\theta) = |\varphi_e(\theta) + \varphi_e(\pi - \theta)|^2 \tag{13.94}$$

$$= |\varphi_e(\theta)|^2 + |\varphi_e(\pi - \theta)|^2 + 2\Re e\{\varphi^*_e(\theta)\varphi_e(\pi - \theta)\}. \tag{13.95}$$

The difference between $P_{e\uparrow i1}$ and P_{ei1} is that, in the first case, we know that the electron that enters into C_1 has spin ↑, while in the second we do not know.

With this results we can compare the probabilities of having one electron with spin ↑, in the direction $\theta = \pi/2$, in an indistinguishable process like in Fig. 13.7a, where

$$P_{e\uparrow i1}(\pi/2) = |\varphi_{e\uparrow}(\pi/2) + \varphi_{e\uparrow}(\pi/2)|^2 = |2\varphi_{e\uparrow}(\pi/2)|^2, \tag{13.96}$$

with the probability of having one electron with spin ↑, in the direction $\theta = \pi/2$, but for a distinguishable process like in the Fig. 13.6, where

$$P_{e\uparrow d1}(\pi/2) = |\varphi_{e\uparrow}(\pi/2)|^2, \tag{13.97}$$

The relation between these probabilities is, certainly:

$$P_{e\uparrow i1}(\pi/2) = 4P_{e\uparrow d1}(\pi/2). \tag{13.98}$$

We can equally compare the probability $P_{ei1}(\pi/2)$ to register one electron in the direction $\theta = \pi/2$ of a process like the one shown in Fig. 13.7b, with the probability $P_{ed1}(\pi/2)$ to register one electron in a process like that in Fig. 13.6. It can be verified that in this case

$$P_{ei1}(\pi/2) = 2P_{ed1}(\pi/2). \tag{13.99}$$

These differences show that in the collision of two electrons, the probability to detect one electron in C_1 is higher when the electrons are indistinguishable than when they are distinguishable.

Exercise 39 Using the relations (11.33) in the form

$$\widehat{S}_- \psi_{s,m_s} = \hbar\sqrt{(s+m_s)(s-m_s+1)}\ \psi_{s,m_s-1}, \tag{13.100}$$

and the definition $\widehat{S}_- = \widehat{S}_x - i\widehat{S}_y$, show that

$$\psi_{10} = \frac{1}{\sqrt{2}}\Big(\chi_{-+} + \chi_{+-}\Big). \tag{13.101}$$

Solution From 13.102 we have

$$\widehat{S}_{-}\chi_{++} = \widehat{S}_{-}\phi_{11} = \hbar\sqrt{2}\;\psi_{10}. \tag{13.102}$$

On the other hand, the addition

$$\widehat{\mathbf{S}} = \widehat{\mathbf{S}}_1 + \widehat{\mathbf{S}}_2, \tag{13.103}$$

means also that

$$\widehat{S}_i = \widehat{S}_{i1} + \widehat{S}_{i2}, \quad \text{for} \quad i = x, y, z. \tag{13.104}$$

The rising and lowering operators can be expressed as

$$\widehat{S}_{\pm} = \widehat{S}_x \pm i\widehat{S}_x = \big(\widehat{S}_{x1} + \widehat{S}_{x2}\big) \pm i\big(\widehat{S}_{y1} + \widehat{S}_{y2}\big). \tag{13.105}$$

We know also that

$$\widehat{S}_x\chi_{+} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{\hbar}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{\hbar}{2}\chi_{-}; \tag{13.106}$$

$$i\widehat{S}_y\chi_{+} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = -\frac{\hbar}{2}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\frac{\hbar}{2}\chi_{-}. \tag{13.107}$$

Using these relations, it is easy to verify that

$$\begin{aligned} \widehat{S}_{-}\chi_{++} &= \big(\widehat{S}_{-1} + \widehat{S}_{-2}\big)\chi_{+}\chi_{+} \\ &= \frac{\hbar}{2}\left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}_1 + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}_2 - \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}_1 - \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}_2\right]\chi_{+}\chi_{+} \\ &= \hbar\big(\chi_{-}\chi_{+} + \chi_{+}\chi_{-}\big). \end{aligned} \tag{13.108}$$

Therefore

$$\psi_{10} = \frac{1}{\sqrt{2}}\Big(\chi_{-+} + \chi_{+-}\Big). \tag{13.109}$$

13.6 Problems

1. Show that when two particles are identical, we have

$$\Phi^S(\xi_1, \xi_2) = \varphi^A(\mathbf{r}_1, \mathbf{r}_2)\chi^A(s_1, s_2) = 0. \tag{13.110}$$

2. Show that

$$\widehat{S}_-|10\rangle = \hbar\sqrt{2}|1-1\rangle. \tag{13.111}$$

3. Deduce the relation (13.20)

$$p_{nf} = np_{(n-1)f} = n!p_{1f}^n. \tag{13.112}$$

4. Deduce the bosons and fermions occupation numbers

$$\overline{n}_j = \frac{1}{e^{(E_j-\mu)/k_BT} - 1}, \tag{13.113}$$

and

$$\overline{n}_j = \frac{1}{e^{(E_j-\mu)/k_BT} + 1}, \tag{13.114}$$

respectively.
5. Plot the Fermi–Dirac distribution function for $T = 76\,\mathrm{K}$ and for $T = 300\,\mathrm{K}$.
6. Determine the thermal energy k_BT for $T = 300\,\mathrm{K}$.
7. Use the Hydrogen atom functions $\phi_{n_1lm_l}(\xi_1)$ to obtain the solutions $|\phi_{n_sn_2l1m_lm_s}\rangle$ of the unperturbed Helium atom and evaluate the matrix elements

$$\langle\phi_{n'_1n'_2l's'm'_lm'_s}|\frac{e^2}{r_{12}}|\phi_{n_sn_2l1m_lm_s}\rangle. \tag{13.115}$$

Verify that these matrix elements can be written as

$$\langle\phi_{n'_1n'_2l's'm'_lm'_s}|\frac{e^2}{r_{12}}|\phi_{n_sn_2l1m_lm_s}\rangle = E_D \pm E_I, \tag{13.116}$$

where

$$E_D = \int |\varphi_{n_1l_1m_{l1}}(\mathbf{r}_1)|^2|\varphi_{n_2l_2m_{l2}}(\mathbf{r}_2)|^2\frac{e^2}{r_{12}}d^3\mathbf{r}_1d^3\mathbf{r}_2, \tag{13.117}$$

and

$$E_I = \int \varphi^*_{n_1l_1m_{l1}}(\mathbf{r}_1)\varphi^*_{n_2l_2m_{l2}}(\mathbf{r}_2)\frac{e^2}{r_{12}}\varphi_{n_1l_1m_{l1}}(\mathbf{r}_2)\varphi_{n_2l_2m_{l2}}(\mathbf{r}_1)d^3\mathbf{r}_1d^3\mathbf{r}_2. \tag{13.118}$$

8. Show that the exchange operator in (13.86), can be written as

$$\pm K = -\frac{1}{2}\left(1 + \frac{4\langle \widehat{\mathbf{S}}_1 \cdot \widehat{\mathbf{S}}_2 \rangle}{\hbar^2}\right) K. \tag{13.119}$$

9. Verify the relations (13.98) and (13.99).

Appendix A
Time Reversal Invariance

Let us briefly discuss the time reversal property. In Chap. 4, we introduced the scattering and transfer matrices S and M. We have seen also that in a scattering process as the one shown in Fig. A.1, these matrices satisfy the following relations

$$\begin{pmatrix} \varphi_{ol}(x_1) \\ \varphi_{or}(x_2) \end{pmatrix} = S \begin{pmatrix} \varphi_{il}(x_1) \\ \varphi_{ir}(x_2) \end{pmatrix} \quad \begin{pmatrix} \varphi_{or}(x_2) \\ \varphi_{ir}(x_2) \end{pmatrix} = M \begin{pmatrix} \varphi_{il}(x_1) \\ \varphi_{ol}(x_1) \end{pmatrix}, \tag{A.1}$$

the matrix S relates the incoming waves with the outgoing ones, while the matrix M connects the state vector at x_1 with the state vector at x_2. If we take the complex conjugate of the previous relations, we have

$$\begin{pmatrix} \varphi^*_{ol}(x_1) \\ \varphi^*_{or}(x_2) \end{pmatrix} = S^* \begin{pmatrix} \varphi^*_{il}(x_1) \\ \varphi^*_{ir}(x_2) \end{pmatrix} \quad \begin{pmatrix} \varphi^*_{or}(x_2) \\ \varphi^*_{ir}(x_2) \end{pmatrix} = M^* \begin{pmatrix} \varphi^*_{il}(x_1) \\ \varphi^*_{ol}(x_1) \end{pmatrix}. \tag{A.2}$$

Let us now consider the time inversion operator T. The action of this operator changes the sign of time, thus the physical quantities depending linearly on time as the velocity and the wave number $k = p/\hbar$, change sign. Under the time inversion operation, particles that were moving to the right, move towards the left and vice versa. A wave function like e^{ikx}, under T, gets transformed into e^{-ikx}. This is equivalent to complex conjugate. Therefore

$$T\varphi(x) = \varphi^*(x). \tag{A.3}$$

If the time inversion operator T acts on the system of Fig. A.1, we will have

$$T\varphi_{il}(x_1) = \varphi^*_{il}(x_1) = \phi_{ol}(x_1), \quad T\varphi_{ir}(x_2) = \varphi^*_{ir}(x_2) = \phi_{or}(x_2), \tag{A.4}$$

$$T\varphi_{ol}(x_1) = \varphi^*_{ol}(x_1) = \phi_{il}(x_1), \quad T\varphi_{or}(x_2) = \varphi^*_{or}(x_2) = \phi_{ir}(x_2). \tag{A.5}$$

As one can see in Fig. A.2, the wave functions that were incoming functions transform into outgoing functions and viceversa. Reversing time in the scattering system implies reversing time every where. This includes the interactions that define the scattering and transfer matrices. If the system remains invariant under time inversion, the matrices S and M should also remain invariant. Therefore, if a

P. Pereyra, *Fundamentals of Quantum Physics*, Undergraduate Lecture Notes in Physics, DOI: 10.1007/978-3-642-29378-8, © Springer-Verlag Berlin Heidelberg 2012

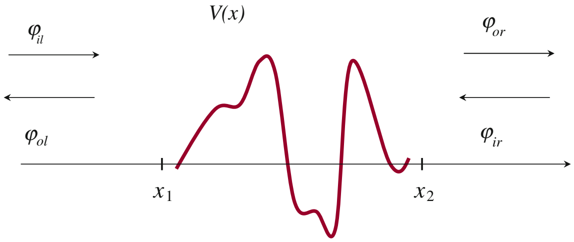


Fig. A.1 The incident wave functions, and the reflected and transmitted wave functions by a dispersion potential $V(x)$

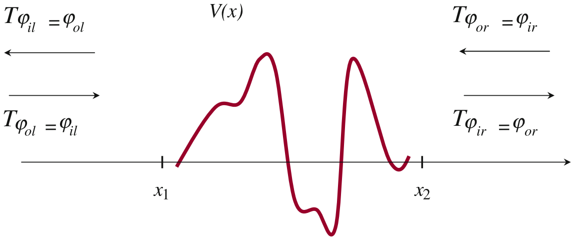


Fig. A.2 The incident wave functions, and the wave functions that are reflected and transmitted by the dispersion potential $V(x)$

system is time reversal invariant, we must have

$$\begin{pmatrix} \phi_{ol}(x_1) \\ \phi_{or}(x_2) \end{pmatrix} = S \begin{pmatrix} \phi_{il}(x_1) \\ \phi_{ir}(x_2) \end{pmatrix} \qquad \begin{pmatrix} \phi_{or}(x_2) \\ \phi_{ir}(x_2) \end{pmatrix} = M \begin{pmatrix} \phi_{il}(x_1) \\ \phi_{ol}(x_1) \end{pmatrix}. \tag{A.6}$$

with S and M that we had in (A.1). These relations, written in terms of the wave functions φ, take the form

$$\begin{pmatrix} \varphi^*_{il}(x_1) \\ \varphi^*_{ir}(x_2) \end{pmatrix} = S \begin{pmatrix} \varphi^*_{ol}(x_1) \\ \varphi^*_{or}(x_2) \end{pmatrix} \qquad \begin{pmatrix} \varphi^*_{ir}(x_2) \\ \varphi^*_{or}(x_2) \end{pmatrix} = M \begin{pmatrix} \varphi^*_{ol}(x_1) \\ \varphi^*_{il}(x_1) \end{pmatrix}. \tag{A.7}$$

If we multiply the first of these equations, from the left by S^{-1}, and we observe that

$$\begin{pmatrix} \varphi^*_{ir}(x_2) \\ \varphi^*_{or}(x_2) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \varphi^*_{or}(x_2) \\ \varphi^*_{ir}(x_2) \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \varphi^*_{ol}(x_1) \\ \varphi^*_{il}(x_1) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \varphi^*_{il}(x_1) \\ \varphi^*_{ol}(x_1) \end{pmatrix}, \tag{A.8}$$

it becomes clear, after comparing with (A.2), that

$$S^{-1} = S^* \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} M \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = M^*. \tag{A.9}$$

For the transfer matrix, this means that

$$\begin{pmatrix} \delta & \gamma \\ \beta & \alpha \end{pmatrix} = \begin{pmatrix} \alpha^* & \beta^* \\ \gamma^* & \delta^* \end{pmatrix}. \tag{A.10}$$

Therefore we conclude that, for systems that are time reversal invariant, the transfer matrices have the structure

$$M = \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix}. \tag{A.11}$$

Notice that this holds when the wave number k on the right side is the same as on the left side of the scattering region. If that is not the case, these structures may change but (A.9) still holds.

Appendix B
Laguerre's Polynomials

With respect to the Laguerre polynomials, there is in the literature such a great diversity of definitions and notations that we feel justified to include this appendix, and deduce some basic relations and properties. The polynomials that Laguerre studied are the polynomials $L_\nu(x) = L_\nu^0(x)$ defined by

$$L_\nu(x) = \frac{1}{\nu!} e^x \frac{d^\nu}{dx^\nu} x^\nu e^{-x}, \tag{B.1}$$

or as the expansion coefficients of the generating function

$$G_L(t,x) = \exp\left\{\frac{xt}{t-1}\right\} = (1-t)\sum_{\nu=0}^{\infty} L_\nu(x) t^\nu, \tag{B.2}$$

and they are solutions of the differential equation

$$\frac{d^2y}{dx^2} + \left(\frac{1}{x} - 1\right)\frac{dy}{dx} + \frac{\nu}{x} y = 0. \tag{B.3}$$

The Laguerre polynomials of lower orders are

$$\begin{aligned}
L_0(x) &= 1,\\
L_1(x) &= \frac{1}{1!}\big(1-x\big),\\
L_2(x) &= \frac{1}{2!}\big(1-4x+x^2\big),\\
L_3(x) &= \frac{1}{3!}\big(6-18x+9x^2-x^3\big),\\
L_4(x) &= \frac{1}{4!}\big(24-96x+72x^2-16x^3+x^4\big).
\end{aligned}$$

Related to these polynomials we have the generalized (or associated) Laguerre polynomials $L_{\nu-\mu}^{\mu}(x)$ defined by

P. Pereyra, *Fundamentals of Quantum Physics*, Undergraduate Lecture Notes in Physics,
DOI: 10.1007/978-3-642-29378-8, © Springer-Verlag Berlin Heidelberg 2012

$$L_v^k(x) = (-1)^k \frac{1}{(v+k)!} \frac{d^k}{dx^k} e^x \frac{d^{v+k}}{dx^{v+k}} x^{v+k} e^{-x} = (-1)^k \frac{d^k}{dx^k} L_{v+k}(x). \tag{B.4}$$

These polynomials are also the coefficients of a generating function, and one can easily verify that a particular solution of the differential equation

$$x\frac{d^2y}{dx^2} + (k+1-2v)\frac{dy}{dx} + \left(n + \frac{k+1}{2} - \frac{x}{4} + \frac{v(v-k)}{x}\right)y = 0 \tag{B.5}$$

is

$$y_n = e^{-x/2} x^v L_n^k. \tag{B.6}$$

In this way, if $v = k/2$, the function

$$y_n = e^{-x/2} x^{k/2} L_n^k, \tag{B.7}$$

is a solution of

$$xy_n'' + y_n' + \left(n + \frac{k+1}{2} - \frac{x}{4} - \frac{k^2}{4x}\right)y_n = 0, \tag{B.8}$$

that can be written as

$$\frac{d}{dx} x \frac{d}{dx} y_n + \left(n + \frac{k+1}{2} - \frac{x}{4} - \frac{k^2}{4x}\right)y_n = 0. \tag{B.9}$$

A very useful representation of $L_v^k(x)$ is the Rodrigues formula

$$L_v^k(x) = \frac{1}{v!} e^x x^{-k} \frac{d^v}{dx^v} x^{v+k} e^{-x} = \sum_{r=0}^{v} (-1)^r \frac{(v+k)!\, x^r}{(v-r)!\, r!\, (r+k)!}. \tag{B.10}$$

Some generalized Laguerre polynomials are:

$$\begin{aligned}
&L_0^0(x) = 1,\\
&L_1^0(x) = 1 - x, \quad L_1^1(x) = 2 - x,\\
&L_2^0(x) = \frac{1}{2}\left(2 - 4x + x^2\right), \quad L_2^1(x) = \frac{1}{2}\left(6 - 6x + x^2\right), \quad L_2^2(x) = \frac{1}{2}\left(12 - 8x + x^2\right).
\end{aligned}$$

Like other orthogonal polynomials, the Laguerre polynomials satisfy a three-term recurrence relation. The relation that the Laguerre polynomials satisfy is

$$(n+1)L_{n+1}^k(x) + (x - k - 2n - 1)L_n^k(x) + (n+k)L_{n-1}^k(x) = 0 \quad n = 1, 2, \ldots \tag{B.11}$$

This recurrence relation can also be written as

$$nL_n^k(x)+(x-k-2n+1)L_{n-1}^k(x)+(n+k-1)L_{n-2}^k(x)=0 \quad n=1,2,\ldots \tag{B.12}$$

With the help of these definitions and relations we will obtain two very useful results for the evaluation of normalization integrals and the calculation of expectation values for the Hydrogen atom.

Orthogonality of Laguerre's Polynomials

If we multiply equation (B.9) by y_m, subtract the same equation (with the indices n and m interchanged), and integrate from 0 to ∞, we have

$$\begin{aligned}&\int_0^\infty \left(y_m\frac{dy}{dx}xy'_n - y_n\frac{dy}{dx}xy'_m\right)dx + (n-m)\int_0^\infty y_m y_n dx = 0,\\ &(y_m xy'_n - y_n xy'_m)\big|_0^\infty + (n-m)\int_0^\infty y_m y_n dx = 0.\end{aligned} \tag{B.13}$$

As $y_n = e^{-x/2}x^{k/2}L_n^k$, and as $L_n^k(x) = \frac{1}{n!}e^x x^{-k}\frac{d^n}{dx^n}x^{n+k}e^{-x}$ is a polynomial of degree n, the first term of the previous equation is zero at $x=0$ if $k>-1$. Thus

$$(n-m)\int_0^\infty e^{-x}x^k L_n^k L_m^k dx = 0. \tag{B.14}$$

From here we conclude that

$$\int_0^\infty e^{-x}x^k L_n^k L_m^k dx = 0 \quad \text{if} \quad n \neq m.$$

Normalization of Laguerre's Polynomials

To obtain the normalization constant of the Laguerre polynomials we need the recurrence relations (B.11) and (B.12). If we consider the relation (B.11) multiplied by $L_{n-1}^k(x)$, subtract (B.12) multiplied by $L_n^k(x)$, multiply the result by $e^{-x}x^k$ and integrate from 0 an ∞ (using the orthogonality condition), we will obtain

$$\int_0^\infty e^{-x}x^k\left[L_n^k(x)\right]^2 dx = \frac{n+k}{n}\int_0^\infty e^{-x}x^k\left[L_{n-1}^k(x)\right]^2 dx. \tag{B.16}$$

One can repeat this relation $n-1$ times to obtain

$$\int_0^\infty e^{-x}x^k\left[L_n^k(x)\right]^2dx = \frac{n+k}{n}\frac{n-1+k}{n-1}\cdots\frac{2+k}{2}\int_0^\infty e^{-x}x^k\left[L_1^k(x)\right]^2dx. \tag{B.17}$$

Since $L_1^k(x) = 1 + k - x$ it is possible to show that

$$\int_0^\infty e^{-x}x^k\left[L_1^k(x)\right]^2dx = (k+1)! \tag{B.18}$$

Therefore

$$\int_0^\infty e^{-x}x^k\left[L_n^k(x)\right]^2dx = \frac{(n+k)!}{n!} \quad \text{for} \quad n = 2, 3, \ldots \tag{B.19}$$

It is possible to explicitly show that this formula is also valid for $n = 0$ and $n = 1$. Thus

$$\int_0^\infty e^{-x}x^k\left[L_n^k(x)\right]^2dx = \frac{(n+k)!}{n!} \quad \text{for} \quad n = 0, 1, 2, 3, \ldots \tag{B.20}$$

This integral and the orthogonality condition are summarized in the following equation

$$\frac{\nu!}{(\nu+k)!}\int_0^\infty x^k e^{-x}L_\nu^k(x)L_{\nu'}^k(x)dx = \delta_{\nu,\nu'}. \tag{B.21}$$

Finally, if we consider the recurrence relation (B.11), we multiply it by $x^k e^{-x}L_n^k(x)$ and integrate from 0 to ∞ we get:

$$\int_0^\infty x^{k+1}e^{-x}L_n^k(x)L_n^k(x)dx = (k+2n+1)\frac{(n+k)!}{n!}. \tag{B.22}$$

References

1. G. Arfken, *Mathematical Methods for Physicists* (Academic Press, New York, 1970)
2. D. Bohm, *Quantum Theory* (Prentice-Hall, New York, 1951)
3. C. Cohen-Tannoudji, B. Din, F. Laloe, *Quantum Mechanics* (Wiley Interscience, New York, 1977)
4. R.H. Dicke, J.P. Wittke, *Introduction to Quantum Mechanics* (Addison-Wesley, Boston, 1960)
5. P.A.M. Dirac, *The Principles of Quantum Mechanics*, 2nd edn. (Oxford University Press, Oxford, 1935)
6. A.R. Edmonds, *Angular Momentum in Quantum Mechanics* (Princeton University Press, Princeton, 1957)
7. R.P. Feynman, R. Leyghton, M. Sands, *Lectures on Physics* vol. 3, (Addison-Wesley, Boston, 1963)
8. S. Gasiorowics, *Quantum Physics* (Wiley, New York, 1974)
9. L.D. Landau, E.M. Lifshitz, *Quantum Mechanics Non-Relativistic Theory* (Butterworth-Heinemann Limited, Oxford, 1981)
10. E. Merzbacher, *Quantum Mechanics* (Wiley, New York, 1961)
11. W. Pauli, *Pauli Lectures on Physics* vol. 5, (MIT Press, Cambridge, 1977)
12. L. de la Peña, *Introducción a la Mecánica Cuántica* (Fondo de Cultura Económica and UNAM, México, 1991)
13. L.I. Schiff, *Quantum Mechanics* (McGraw-Hill, New York, 1949)
14. D.S. Saxon, *Elements of Quantum Mechanics* (Holden-Day, San Francisco, 1968)
15. E. Wigner, *Group Theory and Its Application to the Quantum Mechanics of Atomic Spectra* (Academic Press, New York, 1959)

P. Pereyra, *Fundamentals of Quantum Physics*, Undergraduate Lecture Notes in Physics,
DOI: 10.1007/978-3-642-29378-8,

Index

P. Pereyra, *Fundamentals of Quantum Physics*, Undergraduate Lecture Notes in Physics,
DOI: 10.1007/978-3-642-29378-8, © Springer-Verlag Berlin Heidelberg 2012