Quantum Mechanics

THIRD EDITION

EUGEN MERZBACHER

University of North Carolina at Chapel Hill



JOHN WILEY & SONS, INC.

New York / Chichester / Weinheim

Brisbane / Singapore / Toronto

Acquisitions Editor Stuart Johnson
Marketing Manager Kimberly Manzi
Production Editor Sandra Russell
Designer Madelyn Lesure
Illustration Editor Edward Starr

This book was set in 10/12 Times by University Graphics, Inc. and printed and bo Hamilton Printing Company.
The cover was printed by Hamilton Printing Company.

This book is printed on acid-free paper. ♾

Copyright © 1961, 1970, 1988, 1998 John Wiley & Sons, Inc. All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (508) 750-8400, fax (508) 750-4470. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158-0012, (212) 850-6011, fax (212) 850-6008, E-Mail: PERMREQ@WILEY.COM.

Library of Congress Cataloging in Publication Data:
Merzbacher, Eugen.
Quantum mechanics / Eugen Merzbacher. — 3rd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-471-88702-1 (cloth : alk. paper)
1. Quantum theory. I. Title.
QC174.12.M47 1998
530.12—dc21 97-20756
CIP

Printed in the United States of America

10 9 8 7 6 5 4 3 2

530.12 M577
Quantum mechanics
3rd ed.
Swanston
LCN: 10798500
PO#: 200109006/0001
B/C: 31259006333202

Dedicated to our children:
Celia, Charles, Matthew, and Mary

Preface

The central role of quantum mechanics, as a unifying principle in contemporary physics, is reflected in the training of physicists who take a common course, whether they expect to specialize in atomic, molecular, nuclear, or particle physics, solid state physics, quantum optics, quantum electronics, or quantum chemistry. This book was written for such a course as a comprehensive introduction to the principles of quantum mechanics and to their application in the subfields of physics.

The first edition of this book was published in 1961, the second in 1970. At that time there were few graduate-level texts available to choose from. Now there are many, but I was encouraged by colleagues and students to embark on a further revision of this book. While this new updated edition differs substantially from its predecessors, the underlying purpose has remained the same: To provide a carefully structured and coherent exposition of quantum mechanics; to illuminate the essential features of the theory without cutting corners, and yet without letting technical details obscure the main storyline; and to exhibit wherever possible the common threads by which the theory links many different phenomena and subfields.

The reader of this book is assumed to know the basic facts of atomic and subatomic physics and to have been exposed to elementary quantum mechanics at the undergraduate level. Knowledge of classical mechanics and some familiarity with electromagnetic theory are also presupposed. My intention was to present a self-contained narrative, limiting the selection of topics to those that could be treated equitably without relying on specialized background knowledge.

The material in this book is appropriate for three semesters (or four quarters). The first 19 chapters can make up a standard two-semester (or three-quarter) course on nonrelativistic quantum mechanics. Sometimes classified as ''Advanced Quantum Mechanics'' Chapters 20–24 provide the basis for an understanding of many-body theories, quantum electrodynamics, and relativistic particle theory. The pace quickens here, and many mathematical steps are left to the exercises. It would be presumptuous to claim that every section of this book is indispensable for learning the principles and methods of quantum mechanics. Suffice it to say that there is more here than can be comfortably accommodated in most courses, and that the choice of what to omit is best left to the instructor.

Although my objectives are the same now as they were in the earlier editions, I have tried to take into account changes in physics and in the preparation of the students. Much of the first two-thirds of the book was rewritten and rearranged while I was teaching beginning graduate students and advanced undergraduates. Since most students now reach this course with considerable previous experience in quantum mechanics, the graduated three-stage design of the previous editions—wave mechanics, followed by spin one-half quantum mechanics, followed in turn by the full-fledged abstract vector space formulation of quantum mechanics—no longer seemed appropriate. In modifying it, I have attempted to maintain the inductive approach of the book, which builds the theory up from a small number of simple empirical facts and emphasizes explanations and physical connections over pure formalism. Some introductory material was compressed or altogether jettisoned to make room in the early chapters for material that properly belongs in the first half of this course without unduly inflating the book. I have also added several new topics and tried to refresh and improve the presentation throughout.

As before, the book begins with ordinary wave mechanics and wave packets moving like classical particles. The Schrödinger equation is established, the probability interpretation induced, and the facility for manipulating operators acquired. The principles of quantum mechanics, previously presented in Chapter 8, are now already taken up in Chapter 4. Gauge symmetry, on which much of contemporary quantum field theory rests, is introduced at this stage in its most elementary form. This is followed by practice in the use of fundamental concepts (Chapters 5, 6, and 7), including two-by-two matrices and the construction of a one-dimensional version of the scattering matrix from symmetry principles. Since the bra-ket notation is already familiar to all students, it is now used in these early chapters for matrix elements. The easy access to computing has made it possible to beef up Chapter 7 on the WKB method.

In order to enable the reader to solve nontrivial problems as soon as possible, the new Chapter 8 is devoted to several important techniques that previously became available only later in the course: Variational calculations, the Rayleigh-Ritz method, and elementary time-independent perturbation theory. A section on the use of nonorthogonal basis functions has been added, and the applications to molecular and condensed-matter systems have been revised and brought together in this chapter.

The general principles of quantum mechanics are now the subject of Chapters 9 and 10. Coherent and squeezed harmonic oscillator states are first encountered here in the context of the uncertainty relations. Angular momentum and the nonrelativistic theory of spherical potentials follow in Chapters 11 and 12. Chapter 13 on scattering begins with a new introduction to the concept of cross sections, for colliding and merging beam experiments as well as for stationary targets.

Quantum dynamics, with its various ''pictures'' and representations, has been expanded into Chapters 14 and 15. New features include a short account of Feynman path integration and a longer discussion of density operators, entropy and information, and their relation to notions of measurements in quantum mechanics. All of this is then illustrated in Chapter 16 by the theory of two-state systems, especially spin one-half (previously Chapters 12 and 13). From there it's a short step to a comprehensive treatment of rotations and other discrete symmetries in Chapter 17, ending on a brief new section on non-Abelian local gauge symmetry. Bound-state and time-dependent perturbation theories in Chapters 18 and 19 have been thoroughly revised to clarify and simplify the discussion wherever possible.

The structure of the last five chapters is unchanged, except for the merger of the entire relativistic electron theory in the single Chapter 24. In Chapter 20, as a bridge from elementary quantum mechanics to general collision theory, scattering is reconsidered as a transition between free particle states. Those who do not intend to cross this bridge may omit Chapter 20. The quantum mechanics of identical particles, in its ''second quantization'' operator formulation, is a natural extension of quantum mechanics for distinguishable particles. Chapter 21 spells out the simple assumptions from which the existence of two kinds of statistics (Bose-Einstein and Fermi-Dirac) can be inferred. Since the techniques of many-body physics are now accessible in many specialized textbooks, Chapter 22, which treats some sample problems, has been trimmed to focus on a few essentials.

Counter to the more usual quantization of the classical Maxwell equations, Chapter 23 starts with photons as fundamental entities that compose the electromagnetic field with its local dynamical properties like energy and momentum. The interaction between matter and radiation fields is treated only in first approximation,

leaving all higher-order processes to more advanced textbooks on field theory. The introduction to the elements of quantum optics, including coherence, interference, and statistical properties of the field, has been expanded. As a paradigm for many other physical processes and experiments, two-slit interference is discussed repeatedly (Chapters 1, 9, and 23) from different angles and in increasing depth.

In Chapter 24, positrons and electrons are taken as the constituents of the relativistic theory of leptons, and the Dirac equation is derived as the quantum field equation for charged spin one-half fermions moving in an external classical electromagnetic field. The one-particle Dirac theory of the electron is then obtained as an approximation to the many-electron-positron field theory.

Some important mathematical tools that were previously dispersed through the text (Fourier analysis, delta functions, and the elements of probability theory) have now been collected in the Appendix and supplemented by a section on the use of curvilinear coordinates in wave mechanics and another on units and physical constants. Readers of the second edition of the book should be cautioned about a few notational changes. The most trivial but also most pervasive of these is the replacement of the symbol μ for particle mass by m, or m_e when it's specific to an electron or when confusion with the magnetic quantum number lurks.

There are now almost seven hundred exercises and problems, which form an integral part of the book. The exercises supplement the text and are woven into it, filling gaps and illustrating the arguments. The problems, which appear at the end of the chapters, are more independent applications of the text and may require more work. It is assumed that students and instructors of quantum mechanics will avail themselves of the rapidly growing (but futile to catalog) arsenal of computer software for solving problems and visualizing the propositions of quantum mechanics. Computer technology (especially *MathType*@ and *Mathematica*@) was immensely helpful in preparing this new edition. The quoted references are not intended to be exhaustive, but the footnotes indicate that many sources have contributed to this book and may serve as a guide to further reading. In addition, I draw explicit attention to the wealth of interesting articles on topics in quantum mechanics that have appeared every month, for as long as I can remember, in the *American Journal of Physics*.

The list of friends, students, and colleagues who have helped me generously with suggestions in writing this new edition is long. At the top I acknowledge the major contributions of John P. Hernandez, Paul S. Hubbard, Philip A. Macklin, John D. Morgan, and especially Eric Sheldon. Five seasoned anonymous reviewers gave me valuable advice in the final stages of the project. I am grateful to Mark D. Hannam, Beth A. Kehler, Mary A. Scroggs, and Paul Sigismondi for technical assistance. Over the years I received support and critical comments from Carl Adler, A. Ajay, Andrew Beckwith, Greg L. Bullock, Alan J. Duncan, S. T. Epstein, Heidi Fearn, Colleen Fitzpatrick, Paul H. Frampton, John D. Garrison, Kenneth Hartt, Thomas A. Kaplan, William C. Kerr, Carl Lettenström, Don H. Madison, Kirk McVoy, Matthew Merzbacher, Asher Peres, Krishna Myneni, Y. S. T. Rao, Charles Rasco, G. G. Shute, John A. White, Rolf G. Winter, William K. Wootters, and Paul F. Zweifel. I thank all of them, but the remaining shortcomings are my responsibility.

Most of the work on this new edition of the book was done at the University of North Carolina at Chapel Hill. Some progress was made while I held a U.S. Senior Scientist Humboldt Award at the University of Frankfurt, during a leave of absence at the University of Stirling in Scotland, and on shorter visits to the Institute of Theoretical Physics at Santa Barbara, the Institute for Nuclear Theory in Seattle,

and TRIFORM Camphill Community in Hudson, New York. The encouragement of colleagues and friends in all of these places is gratefully acknowledged. But this long project, often delayed by other physics activities and commitments, could never have been completed without the unfailing patient support of my wife, Ann.

Eugen Merzbacher

Contents

CHAPTER 1 *Introduction to Quantum Mechanics* **1**

1. Quantum Theory and the Wave Nature of Matter **1**
2. The Wave Function and its Meaning **4**

Problems **10**

CHAPTER 2 *Wave Packets, Free Particle Motion, and the Wave Equation* **12**

1. The Principle of Superposition **12**
2. Wave Packets and the Uncertainty Relations **14**
3. Motion of a Wave Packet **18**
4. The Uncertainty Relations and the Spreading of Wave Packets **20**
5. The Wave Equation for Free Particle Motion **22**

Problems **24**

CHAPTER 3 *The Schrödinger Equation, the Wave Function, and Operator Algebra* **25**

1. The Wave Equation and the Interpretation of ψ **25**
2. Probabilities in Coordinate and Momentum Space **29**
3. Operators and Expectation Values of Dynamical Variables **34**
4. Commutators and Operator Algebra **38**
5. Stationary States and General Solutions of the Wave Equation **41**
6. The Virial Theorem **47**

Problems **49**

CHAPTER 4 *The Principles of Wave Mechanics* **51**

1. Hermitian Operators, their Eigenfunctions and Eigenvalues **51**
2. The Superposition and Completeness of Eigenstates **57**
3. The Continuous Spectrum and Closure **60**
4. A Familiar Example: The Momentum Eigenfunctions and the Free Particle **62**
5. Unitary Operators. The Displacement Operator **68**
6. The Charged Particle in an External Electromagnetic Field and Gauge Invariance **71**
7. Galilean Transformation and Gauge Invariance **75**

Problems **78**

CHAPTER 5 *The Linear Harmonic Oscillator* **79**

1. Preliminary Remarks **79**
2. Eigenvalues and Eigenfunctions **81**

3. Study of the Eigenfunctions **84**
4. The Motion of Wave Packets **89**
Problems **90**

CHAPTER 6 *Sectionally Constant Potentials in One Dimension* **92**

1. The Potential Step **92**
2. The Rectangular Potential Barrier **97**
3. Symmetries and Invariance Properties **99**
4. The Square Well **103**
Problems **111**

CHAPTER 7 *The WKB Approximation* **113**

1. The Method **113**
2. The Connection Formulas **116**
3. Application to Bound States **121**
4. Transmission Through a Barrier **125**
5. Motion of a Wave Packet and Exponential Decay **131**
Problems **134**

CHAPTER 8 *Variational Methods and Simple Perturbation Theory* **135**

1. The Calculus of Variations in Quantum Mechanics **135**
2. The Rayleigh-Ritz Trial Function **139**
3. Perturbation Theory of the Schrödinger Equation **142**
4. The Rayleigh-Ritz Method with Nonorthogonal Basis Functions **146**
5. The Double Oscillator **149**
6. The Molecular Approximation **159**
7. The Periodic Potential **165**
Problems **176**

CHAPTER 9 *Vector Spaces in Quantum Mechanics* **179**

1. Probability Amplitudes and Their Composition **179**
2. Vectors and Inner Products **186**
3. Operators **188**
4. The Vector Space of Quantum Mechanics and the Bra-Ket Notation **195**
5. Change of Basis **199**
6. Hilbert Space and the Coordinate Representation **202**
Problems **206**

CHAPTER 10 *Eigenvalues and Eigenvectors of Operators, the Uncertainty Relations, and the Harmonic Oscillator* **207**

1. The Eigenvalue Problem for Normal Operators **207**
2. The Calculation of Eigenvalues and the Construction of Eigenvectors **209**

3. Variational Formulation of the Eigenvalue Problem for a Bounded Hermitian Operator **212**
4. Commuting Observables and Simultaneous Measurements **214**
5. The Heisenberg Uncertainty Relations **217**
6. The Harmonic Oscillator **220**
7. Coherent States and Squeezed States **225**

Problems **231**

CHAPTER 11 *Angular Momentum in Quantum Mechanics* **233**

1. Orbital Angular Momentum **233**
2. Algebraic Approach to the Angular Momentum Eigenvalue Problem **238**
3. Eigenvalue Problem for L_z and $\mathbf{L}^2$. **242**
4. Spherical Harmonics **248**
5. Angular Momentum and Kinetic Energy **252**

Problems **255**

CHAPTER 12 *Spherically Symmetric Potentials* **256**

1. Reduction of the Central-Force Problem **256**
2. The Free Particle as a Central-Force Problem **257**
3. The Spherical Square Well Potential **262**
4. The Radial Equation and the Boundary Conditions **263**
5. The Coulomb Potential **265**
6. The Bound-State Energy Eigenfunctions for the Coulomb Potential **270**

Problems **275**

CHAPTER 13 *Scattering* **278**

1. The Cross Section **278**
2. The Scattering of a Wave Packet **286**
3. Green's Functions in Scattering Theory **290**
4. The Born Approximation **295**
5. Partial Waves and Phase Shifts **298**
6. Determination of the Phase Shifts and Scattering Resonances **302**
7. Phase Shifts and Green's Functions **308**
8. Scattering in a Coulomb Field **310**

Problems **314**

CHAPTER 14 *The Principles of Quantum Dynamics* **315**

1. The Evolution of Probability Amplitudes and the Time Development Operator **315**
2. The Pictures of Quantum Dynamics **319**
3. The Quantization Postulates for a Particle **323**
4. Canonical Quantization and Constants of the Motion **326**

5. Canonical Quantization in the Heisenberg Picture **331**
6. The Forced Harmonic Oscillator **335**
Problems **342**

CHAPTER 15 *The Quantum Dynamics of a Particle* **344**

1. The Coordinate and Momentum Representations **344**
2. The Propagator in the Coordinate Representation **348**
3. Feynman's Path Integral Formulation of Quantum Dynamics **355**
4. Quantum Dynamics in Direct Product Spaces and Multiparticle Systems **358**
5. The Density Operator, the Density Matrix, Measurement, and Information **363**
Problems **370**

CHAPTER 16 *The Spin* **372**

1. Intrinsic Angular Momentum and the Polarization of ψ waves **372**
2. The Quantum Mechanical Description of the Spin **377**
3. Spin and Rotations **381**
4. The Spin Operators, Pauli Matrices, and Spin Angular Momentum **385**
5. Quantum Dynamics of a Spin System **390**
6. Density Matrix and Spin Polarization **392**
7. Polarization and Scattering **399**
8. Measurements, Probabilities, and Information **403**
Problems **408**

CHAPTER 17 *Rotations and Other Symmetry Operations* **410**

1. The Euclidean Principle of Relativity and State Vector Transformations **410**
2. The Rotation Operator, Angular Momentum, and Conservation Laws **413**
3. Symmetry Groups and Group Representations **416**
4. The Representations of the Rotation Group **421**
5. The Addition of Angular Momenta **426**
6. The Clebsch-Gordan Series **431**
7. Tensor Operators and the Wigner-Eckart Theorem **432**
8. Applications of the Wigner-Eckart Theorem **437**
9. Reflection Symmetry, Parity, and Time Reversal **439**
10. Local Gauge Symmetry **444**
Problems **448**

CHAPTER 18 *Bound-State Perturbation Theory* **451**

1. The Perturbation Method **451**
2. Inhomogeneous Linear Equations **453**

3. Solution of the Perturbation Equations **455**
4. Electrostatic Polarization and the Dipole Moment **459**
5. Degenerate Perturbation Theory **463**
6. Applications to Atoms **467**
7. The Variational Method and Perturbation Theory **473**
8. The Helium Atom **476**

Problems **480**

CHAPTER 19 *Time-Dependent Perturbation Theory* **482**

1. The Equation of Motion in the Interaction Picture **482**
2. The Perturbation Method **485**
3. Coulomb Excitation and Sum Rules **487**
4. The Atom in a Radiation Field **491**
5. The Absorption Cross Section **495**
6. The Photoelectric Effect **501**
7. The Golden Rule for Constant Transition Rates **503**
8. Exponential Decay and Zeno's Paradox **510**

Problems **515**

CHAPTER 20 *The Formal Theory of Scattering* **517**

1. The Equations of Motion, the Transition Matrix, the S Matrix, and the Cross Section **517**
2. The Integral Equations of Scattering Theory **521**
3. Properties of the Scattering States **525**
4. Properties of the Scattering Matrix **527**
5. Rotational Invariance, Time Reversal Symmetry, and the S Matrix **530**
6. The Optical Theorem **532**

Problems **533**

CHAPTER 21 *Identical Particles* **535**

1. The Indistinguishability of and the State Vector Space for Identical Particles **535**
2. Creation and Annihilation Operators **538**
3. The Algebra of Creation and Annihilation Operators **540**
4. Dynamical Variables **544**
5. The Continuous One-Particle Spectrum and Quantum Field Operators **546**
6. Quantum Dynamics of Identical Particles **549**

Problems **552**

CHAPTER 22 *Applications to Many-Body Systems* **555**

1. Angular Momentum of a System of Identical Particles **555**
2. Angular Momentum and Spin One-Half Boson Operators **556**

3. First-Order Perturbation Theory in Many-Body Systems **558**
4. The Hartree-Fock Method **560**
5. Quantum Statistics and Thermodynamics **564**
Problems **567**

CHAPTER 23 *Photons and the Electromagnetic Field* **569**

1. Fundamental Notions **569**
2. Energy, Momentum, and Angular Momentum of the Radiation Field **573**
3. Interaction with Charged Particles **576**
4. Elements of Quantum Optics **580**
5. Coherence, Interference, and Statistical Properties of the Field **583**
Problems **591**

CHAPTER 24 *Relativistic Electron Theory* **592**

1. The Electron-Positron Field **592**
2. The Dirac Equation **596**
3. Relativistic Invariance **600**
4. Solutions of the Free Field Dirac Equation **606**
5. Charge Conjugation, Time Reversal, and the *PCT* Theorem **608**
6. The One-Particle Approximation **613**
7. Dirac Theory in the Heisenberg picture **617**
8. Dirac Theory in the Schrödinger Picture and the Nonrelativistic Limit **621**
9. Central Forces and the Hydrogen Atom **623**
Problems **629**

APPENDIX **630**

1. Fourier Analysis and Delta Functions **630**
2. Review of Probability Concepts **634**
3. Curvilinear Coordinates **638**
4. Units and Physical Constants **640**

REFERENCES **642**

INDEX **647**

CHAPTER 1

Introduction to Quantum Mechanics

Quantum mechanics is the theoretical framework within which it has been found possible to describe, correlate, and predict the behavior of a vast range of physical systems, from particles through nuclei, atoms and radiation to molecules and condensed matter. This introductory chapter sets the stage with a brief review of the historical background and a preliminary discussion of some of the essential concepts.[1]

1. ***Quantum Theory and the Wave Nature of Matter.*** Matter at the atomic and nuclear or microscopic level reveals the existence of a variety of particles which are identifiable by their distinct properties, such as mass, charge, spin, and magnetic moment. All of these seem to be of a quantum nature in the sense that they take on only certain discrete values. This discreteness of physical properties persists when particles combine to form nuclei, atoms, and molecules.

The notion that atoms, molecules, and nuclei possess discrete energy levels is one of the basic facts of quantum physics. The experimental evidence for this fact is overwhelming and well known. It comes most directly from observations on inelastic collisions (Franck-Hertz experiment) and selective absorption of radiation, and somewhat indirectly from the interpretation of spectral lines.

Consider an object as familiar as the hydrogen atom, which consists of a proton and an electron, bound together by forces of electrostatic attraction. The electron can be removed from the atom and identified by its charge, mass, and spin. It is equally well known that the hydrogen atom can be excited by absorbing certain discrete amounts of energy and that it can return the excitation energy by emitting light of discrete frequencies. These are empirical facts.

Niels Bohr discovered that any understanding of the observed discreteness requires, above all, the introduction of *Planck's constant*, $h = 6.6261 \times 10^{-34}$ J sec $= 4.136 \times 10^{-15}$ eV sec. In the early days, this constant was often called the *quantum of action*. By the simple relation

$$\boxed{\Delta E = h\nu} \tag{1.1}$$

it links the observed spectral frequency ν to the jump ΔE between discrete energy levels. Divided by 2π, the constant $\hbar = h/2\pi$ appears as the unit of angular momentum, the discrete numbers $n\hbar$ ($n = 0, 1/2, 1, 3/2, 2, \ldots$) being the only values which a component of the angular momentum of a system can assume. All of this is true for systems that are composed of several particles, as well as for the particles themselves, most of which are no more "elementary" than atoms and nuclei. The composite structure of most particles has been unraveled by quantum theoretic

[1]Many references to the literature on quantum mechanics are found in the footnotes, and the bibliographic information is listed after the Appendix. It is essential to have at hand a current summary of the relevant empirical knowledge about systems to which quantum mechanics applies. Among many good choices, we mention Haken and Wolf (1993), Christman (1988), Krane (1987), and Perkins (1982).

analysis of "spectroscopic" information accumulated in high-energy physics experiments.

Bohr was able to calculate discrete energy levels of an atom by formulating a set of quantum conditions to which the canonical variables q_i and p_i of classical mechanics were to be subjected. For our purposes, it is sufficient to remember that in this "old quantum theory" the classical phase (or action) integrals for a conditionally periodic motion were required to be quantized according to

$$\oint p_i dq_i = n_i h \tag{1.2}$$

where the quantum numbers n_i are integers, and each contour integral is taken over the full period of the generalized coordinate q_i. The quantum conditions (1.2) gave good results in calculating the energy levels of simple systems but failed when applied to such systems as the helium atom.

Exercise 1.1. Calculate the quantized energy levels of a linear harmonic oscillator of angular frequency ω in the old quantum theory.

Exercise 1.2. Assuming that the electron moves in a circular orbit in a Coulomb field, derive the Balmer formula for the spectrum of hydrogenic atoms from the quantum condition (1.2) and the Bohr formula (1.1).

It is well known that (1.1) played an important role even in the earliest forms of quantum theory. Einstein used it to explain the photoelectric effect by inferring that light, which through the nineteenth century had been so firmly established as a wave phenomenon, can exhibit a particle-like nature and is emitted or absorbed only in quanta of energy. Thus, the concept of the photon as a particle with energy $E = h\nu$ emerged. The constant h connects the wave (ν) and particle (E) aspects of light.

Louis de Broglie proposed that the wave–particle duality is not a monopoly of light but is a universal characteristic of nature which becomes evident when the magnitude of h cannot be neglected. He thus brought out a second fundamental fact, usually referred to as the *wave nature of matter*. This means that in certain experiments beams of particles with mass give rise to interference and diffraction phenomena and exhibit a behavior very similar to that of light. Although such effects were first produced with electron beams, they are now commonly observed with slow neutrons from a reactor. When incident on a crystal, these behave very much like X rays. Heavier objects, such as entire atoms and molecules, have also been shown to exhibit wave properties. Although one sometimes speaks of *matter waves*, this term is not intended to convey the impression that the particles themselves are oscillating in space.

From experiments on the interference and diffraction of particles, we infer the very simple law that the infinite harmonic plane waves associated with the motion of a free particle of momentum $\mathbf{p}$ propagate in the direction of motion and that their (de Broglie) wavelength is given by

$$\boxed{\lambda = \frac{h}{p}} \tag{1.3}$$

This relation establishes contact between the wave and the particle pictures. The finiteness of Planck's constant is the basic point here. For if h were zero, then no

matter what momentum a particle had, the associated wave would always correspond to $\lambda = 0$ and would follow the laws of classical mechanics, which can be regarded as the short wavelength limit of wave mechanics in the same way as geometrical optics is the short wavelength limit of wave optics. A free particle would then not be diffracted but would go on a straight rectilinear path, just as we expect classically.

Let us formulate this a bit more precisely. If x is a characteristic length involved in describing the motion of a body of momentum p, such as the linear dimension of an obstacle in its path, the wave aspect of matter will be hidden from our sight, if

$$\frac{\lambda}{x} = \frac{h}{xp} \ll 1 \tag{1.4}$$

i.e., if the quantum of action h is negligible compared with xp. Macroscopic bodies, to which classical mechanics is applicable, satisfy the condition $xp \gg h$ extremely well. To give a numerical example, we note that even as light a body as an atom moving with a kinetic energy corresponding to a temperature of $T = 10^{-6}$ K still has a wavelength no greater than about a micron or 10^{-6} m! We thus expect that classical mechanics is contained in quantum mechanics as a limiting form ($h \to 0$).

Indeed, the gradual transition that we can make conceptually as well as practically from the atomic level with its quantum laws to the macroscopic level at which the classical laws of physics are valid suggests that quantum mechanics must not only be consistent with classical physics but should also be capable of yielding the classical laws in a suitable approximation. This requirement, which serves as a guide in discovering the correct quantum laws, is called the *correspondence principle*. Later we will see that the limiting process which establishes the connection between quantum and classical mechanics can be exploited to give a useful approximation for quantum mechanical problems (see WKB approximation, Chapter 7).

We may read (1.3) the other way around and infer that, generally, a wave that propagates in an infinite medium has associated with it a particle, or *quantum*, of momentum $p = h/\lambda$. If a macroscopic wave is to carry an appreciable amount of momentum, as a classical electromagnetic or an elastic wave may, there must be associated with the wave an enormous number of quanta, each contributing a very small amount of momentum. For example, the waves of the electromagnetic field are associated with quanta (*photons*) for which the Bohr-Einstein relation $E = h\nu$ holds. Since photons have no mass, their energy and momentum are according to relativistic mechanics related by $E = cp$, in agreement with the connection between energy (density) and momentum (density) in Maxwell's theory of the electromagnetic field. Reversing the argument that led to de Broglie's proposal, we conclude that (1.3) is valid for photons as well as for material particles. At macroscopic wavelengths, corresponding to microwave or radio frequency, a very large number of photons is required to build up a field of macroscopically discernible intensity. Such a field can be described in classical terms only if the photons can act coherently. As will be discussed in detail in Chapter 23, this requirement leads to the peculiar conclusion that a state of *exactly* n photons cannot represent a classical field, even if n is arbitrarily large. Evidently, statistical distributions of variable numbers of photons must play a fundamental role in the theory.

The massless quanta corresponding to elastic (e.g., sound) waves are called *phonons* and behave similarly to photons, except that c is now the speed of sound, and the waves can be longitudinal as well as transverse. It is important to remember that such waves are generated in an elastic medium, and not in free space.

2. *The Wave Function and Its Meaning.* As we have seen, facing us at the outset is the fact that matter, say an electron, exhibits both particle and wave aspects.[2] This duality was described in deliberately vague language by saying that the de Broglie relation "associates" a wavelength with a particle momentum. The vagueness reflects the fact that particle and wave aspects, when they show up in the same thing such as the electron, are incompatible with each other unless traditional concepts of classical physics are modified to a certain extent. Particle traditionally means an object with a definite position in space. Wave means a pattern spread out in space and time, and it is characteristic of a wave that it does not define a location or position sharply.

Historically, the need for a reconciliation of the two seemingly contradictory concepts of wave and particle was stressed above all by Bohr, whose tireless efforts at interpreting the propositions of quantum mechanics culminated in the formulation of a doctrine of *complementarity*. According to this body of thought, a full description and understanding of natural processes, not only in the realm of atoms but at all levels of human experience, cannot be attained without analyzing the complementary aspects of the phenomena and of the means by which the phenomena are observed. Although this epistemological view of the relationship between classical and quantal physics is no longer central to the interpretation of quantum mechanics, an appreciation of Bohr's program is important because, through stimulation and provocation, it has greatly influenced our attitude toward the entire subject.[3]

How a synthesis of the wave and particle concepts might be achieved can, for a start, perhaps be understood if we recall that the quantum theory must give an account of the discreteness of certain physical properties, e.g., energy levels in an atom or a nucleus. Yet discreteness did not first enter physics with the Bohr atom. In classical macroscopic physics discrete, "quantized," physical quantities appear naturally as the frequencies of vibrating bodies of finite extension, such as strings, membranes, or air columns. We speak typically of the (natural) *modes* of such systems. These phenomena have found their simple explanation in terms of interference between incident and reflected waves. Mathematically, the discrete behavior is enforced by boundary conditions: the fixed ends of the string, the clamping of the membrane rim, the size and shape of the enclosure that confines the air column. Similarly, it is tempting to see in the discrete properties of atoms the manifestations of bounded wave motion and to connect the discrete energy levels with standing waves. In such a picture, the bounded wave must somehow be related to the confinement of the particle to its "orbit," but it is obvious that the concept of an orbit as a trajectory covered with definite speed cannot be maintained.

A wave is generally described by its velocity of propagation, wavelength, and amplitude. (There is also the phase constant of a wave, but, as we shall see later, for one particle this is undetermined.) Since in a standing wave it is the *wavelength* (or frequency) that assumes discrete values, it is evident that if our analogy is meaningful at all, there must be a correspondence between the energy of an atom and the

[2]It will be convenient to use the generic term *electron* frequently when we wish to place equal emphasis on the particle and wave aspects of a constituent of matter. The electron has been chosen only for definiteness of expression (and historical reasons). Quantum mechanics applies equally to protons, neutrons, mesons, quarks, and so on.

[3]For a compilation of original articles on the foundations of quantum mechanics and an extensive bibliography, see Wheeler and Zurek (1985). Also see the resource letters in the *American Journal of Physics*: DeWitt and Graham (1971), and L. E. Ballentine (1987).

wavelength of the wave associated with the particle motion. For a free particle, one that is not bound in an atom, the de Broglie formula (1.3) has already given us a relationship connecting wavelength with energy (or momentum). The connection between wavelength and the mechanical quantities, momentum or energy, is likely to be much more complicated for an electron bound to a nucleus as in the hydrogen atom, or for a particle moving in any kind of a potential. Erwin Schrödinger discovered the *wave equation* that enables us to evaluate the "proper frequencies" or *eigenfrequencies* of general quantum mechanical systems.

The amplitudes or wave fields, which, with their space and time derivatives, appear in the Schrödinger equation, may or may not have directional (i.e., polarization) properties. We will see in Chapter 16 that the *spin* of the particles corresponds to the polarization of the waves. However, for many purposes the dynamical effects of the spin are negligible in first approximation, especially if the particles move with nonrelativistic velocities and are not exposed to magnetic fields. We will neglect the spin for the time being, much as in a simple theory of wave optical phenomena, where we are concerned with interference and diffraction or with the geometrical optics limit, the transverse nature of light can often be neglected. Hence, we attempt to build up quantum mechanics with mass (different from zero) first by use of *scalar* waves. For particles with zero spin, for example, pions and K mesons, this gives an appropriate description. For particles with nonzero spin, such as electrons, quarks, nucleons, or muons, suitable corrections must be made later. We will also see that the spin has profound influence on the behavior of systems comprised of several, or many, identical particles.

Mathematically, the scalar waves are represented by a function $\psi(x, y, z, t)$, which in colorless terminology is called the *wave function*. Upon its introduction we immediately ask such questions as these: Is ψ a measurable quantity, and what precisely does it describe? In particular, what feature of the particle aspect of the particle is related to the wave function?

We cannot expect entirely satisfactory answers to these questions before we have become familiar with the properties of these waves and with the way in which ψ is used in calculations, but the questions can be placed in sharper focus by reexamining the de Broglie relation (1.3) between wavelength, or the wave number $k = 2\pi/\lambda$, and particle momentum:

$$\boxed{p = \hbar k = \frac{h}{2\pi}\frac{2\pi}{\lambda}} \tag{1.5}$$

Suppose that a beam of particles having momentum p in the x direction is viewed from a frame of reference that moves uniformly with velocity v, along the x axis. For nonrelativistic velocities of the particles and for $v \ll c$, the usual Galilean transformation

$$x' = x - vt \tag{1.6}$$

changes the particle momentum to

$$p' = p - mv \tag{1.7}$$

If the particles are in free space, we must assume that the de Broglie relation is valid also in the new frame of reference and that therefore

$$\lambda' = \frac{h}{p'} = \frac{h}{p - mv} \neq \lambda \tag{1.8}$$

We can easily imagine an experimental test of this relation by measuring the spacing of the fringes in a two-slit Young-type interference apparatus, which in its entirety moves at velocity v parallel to the beam. Of the outcome of such a test there can hardly be any doubt: The fringe pattern will broaden, corresponding to the increased wavelength.

When classical elastic waves, which propagate in the ''rest'' frame of the medium with speed V are viewed from the ''moving'' frame of reference, their phase, $\phi = kx - \omega t = 2\pi(x - Vt)/\lambda$, as a measure of the number of amplitude peaks and valleys within a given distance, is Galilean-invariant. The transformation (1.6) gives the connection

$$\phi = \phi' = \frac{2\pi}{\lambda}(x' + vt - Vt) = \frac{2\pi}{\lambda}(x' - V't) \tag{1.9}$$

from which we deduce the familiar Doppler shift

$$\omega' = \omega\left(1 - \frac{v}{V}\right) \tag{1.10}$$

and the unsurprising result:

$$\lambda' = \lambda \tag{1.11}$$

Although the invariance of the wavelength accords with our experience with elastic waves, it is in stark conflict with the conclusion (1.8) for de Broglie waves, the ψ waves of quantum mechanics. What has gone awry?

Two explanations come to mind to resolve this puzzle. In later chapters we will see that both are valid and that they are mutually consistent. Here, the main lesson to be learned is that the ψ waves are unlike classical elastic waves, whose amplitude is in principle observable and which are therefore unchanged under the Galilean transformation. Instead, we must entertain the possibility that, under a Galilean transformation, ψ changes into a transformed wave function, ψ', and we must ascertain the transformation law for de Broglie waves. If ψ cannot be a directly measurable amplitude, there is no compelling reason for it to be a real-valued function. We will see in Section 4.7 that by allowing ψ to be complex-valued for the description of free particles with momentum p, the conflict between Eqs. (1.8) and (1.11) can be resolved. A local gauge transformation, induced by the Galilean transformation (1.6), will then be found to provide a new transformation rule for the phase of the waves, replacing (1.11) and restoring consistency with the correct quantum relation (1.8).

An alternative, and in the final analysis equivalent, way to avoid the contradiction implied by Eqs. (1.8) and (1.11) is to realize that Lorentz, rather than Galilei, transformations may be required in spite of the assumed subluminal particle and frame-of-reference velocities. If the Lorentz transformation

$$x' = \gamma(x - vt) \quad \text{and} \quad t' = \gamma\left(t - \frac{vx}{c^2}\right), \qquad \gamma = \left(1 - \frac{v^2}{c^2}\right)^{-1/2} \tag{1.12}$$

is applied, and if Lorentz invariance (instead of Galilean invariance) of the phase of the ψ waves is assumed, the frequency and wave number must transform relativistically as

$$k' = \gamma\left(k - \frac{v\omega}{c^2}\right) \quad \text{and} \quad \omega' = \gamma(\omega - vk) \tag{1.13}$$

For low relative velocities ($\gamma \approx 1$) the second of these equations again gives the first-order Doppler frequency shift. However, the transformation law for the wave number contains a relativistic term, which was tacitly assumed to be negligible in the nonrelativistic regime. This relation becomes consistent with (1.5) and the nonrelativistic equations (1.7) only if it is assumed that the frequency ω of de Broglie waves, of which we have no direct experimental information, is related to the particle mass by

$$\hbar\omega \approx mc^2 \tag{1.14}$$

This result is eminently reasonable in a *relativistic* quantum mechanics. It implies that for nonrelativistic particles the phase velocity of the waves is

$$V = \frac{\omega}{k} = \frac{\hbar\omega}{p} \approx \frac{mc^2}{p} \tag{1.15}$$

which greatly exceeds the speed of light and which explains the need for Lorentz transformations under all circumstances. However, since, except for two final chapters on fully relativistic quantum mechanics for photons and electrons, our treatment is intended to center on *nonrelativistic* quantum mechanics, we will retain Galilean transformations and acknowledge the need to transform ψ appropriately. (See Section 4.7 for a more detailed discussion.)

Exercise 1.3. Compare the behavior of de Broglie waves for particles of mass m with the changes that the wavelength and frequency of light undergo as we look at a plane electromagnetic wave from a "moving" frame of reference.

As we progress through quantum mechanics, we will become accustomed to ψ as an important addition to our arsenal of physical concepts, in spite of its unusual transformation properties. If its physical significance remains as yet somewhat obscure to us, one thing seems certain: The wave function ψ must in some sense be a measure of the presence of a particle. Thus, we do not expect to find the particle in those regions of space where $\psi = 0$. Conversely, in regions of space where the particle may be found, ψ must be different from zero. But the function $\psi(x, y, z, t)$ itself cannot be a direct measure of the likelihood of finding the particle at position x, y, z at time t. For if it were that, it would have to have a nonnegative value everywhere. Yet, it is impossible for ψ to be nonnegative everywhere, if destructive interference of the ψ waves is to account for the observed dark interference fringes and for the instability of any but the distinguished Bohr orbits in a hydrogen atom.

In physical optics, interference patterns are produced by the superposition of waves of $\mathbf{E}$ and $\mathbf{B}$, but the *intensity* of the fringes is measured by $\mathbf{E}^2$ and $\mathbf{B}^2$. In analogy to this situation, we assume that the positive quantity $|\psi(x, y, z, t)|^2$ measures the probability of finding the particle at position x, y, z at time t. (The absolute value has been taken because it will turn out that ψ can have complex values.) The full meaning of this interpretation of ψ and its internal consistency will be discussed in detail in Chapter 3. Here we merely want to advance some general qualitative arguments for this so-called *probability interpretation* of the quantum wave function for particles with mass, which was introduced into quantum mechanics by Max Born.

From all that is known to date, it is consistent with experiment and theory to associate the wave ψ with a *single* particle or atom or other quantum system as a

representative of a statistical ensemble. Owing to their characteristic properties, such as charge and mass, particles can be identified singly in the detection devices of experimental physics. With the aid of these tools, it has been abundantly established that the interference fringes shown schematically in Figure 1.1 are the statistical result of the effect of a very large number of independent particles hitting the screen. The interference pattern evolves only after many particles have been deposited on the detection screen. Note that the appearance of the interference effects does not require that a whole beam of particles go through the slits at one time. In fact, particles can actually be accelerated and observed singly, and the interference pattern can be produced over a length of time, a particle hitting the screen now here, now there, in seemingly random fashion. When many particles have come through, a regular interference pattern will be seen to have formed. The conclusion is almost inevitable that ψ describes the behavior of single particles, but that it has an intrinsic probabilistic meaning. The quantity $|\psi|^2$ would appear to measure the chance of finding the particle at a certain place. In a sense, this conclusion was already implicit in our earlier discussion regarding an infinite plane wave as representative of a free particle with definite momentum (wavelength) but completely indefinite position. At least if ψ is so interpreted, the observations can be correlated effortlessly with the mathematical formalism. These considerations motivate the more descriptive name *probability amplitude* for the wave function $\psi(x, y, z, t)$. As a word of caution, we note that the term *amplitude*, as used in quantum mechanics, generally refers to the spacetime-dependent wave function ψ, and not merely to its extreme value, as is customary in speaking about elastic or electromagnetic waves.

The indeterminism that the probabilistic view ascribes to nature, and that still engenders discomfort in some quarters, can be illustrated by the idealized experiment shown in Figure 1.1. Single particles are subject to wave interference effects, and

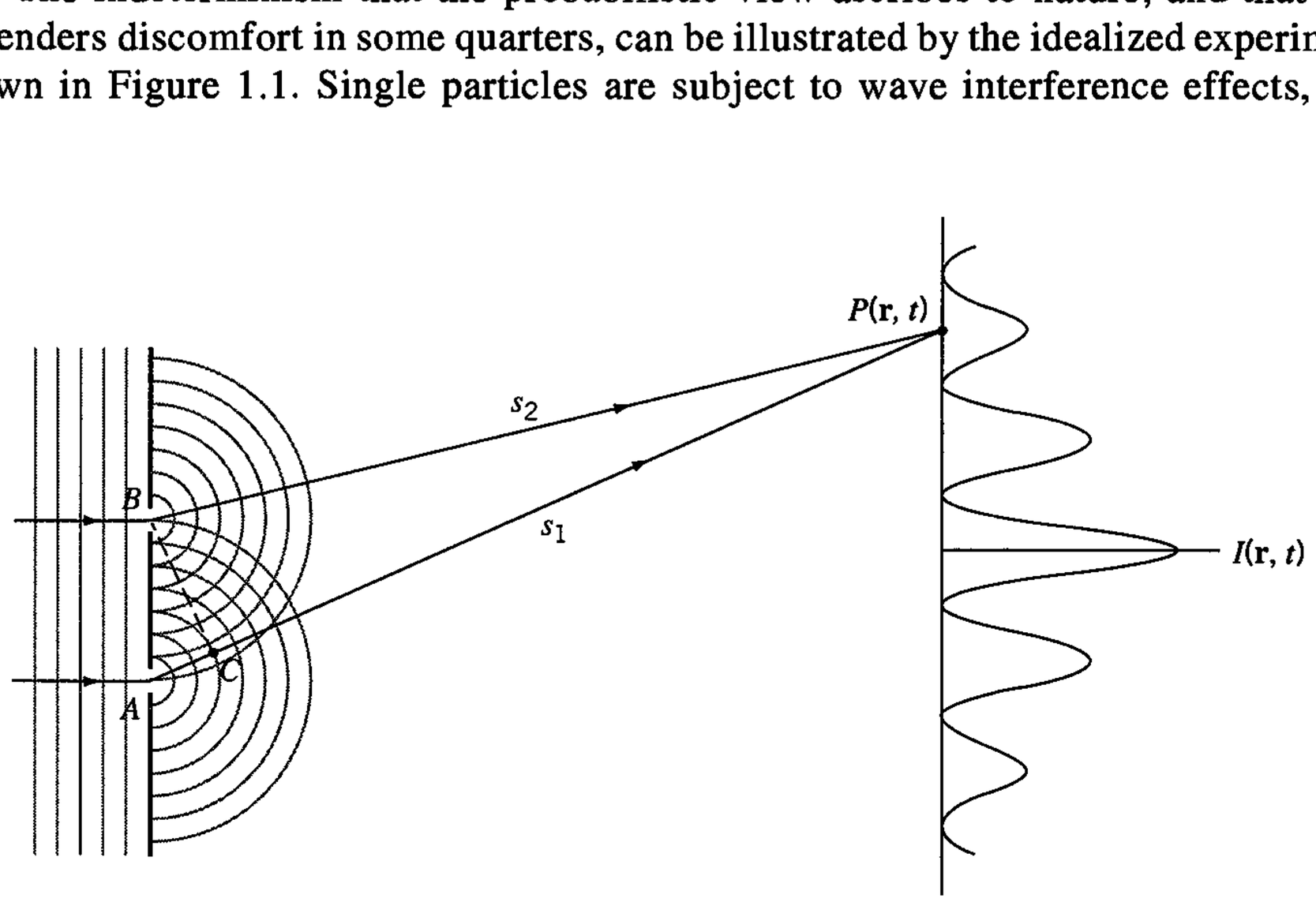


Figure 1.1. Schematic diagram of the geometry in a two-slit experiment. A plane wave, depicted by surfaces of equal phase, is incident with wavelength $\lambda = 2\pi/k$ from the left on the narrow slits A and B. The amplitude and intensity at the spacetime point $P(\mathbf{r},t)$, at a distance $s_1 = AP$ and $s_2 = BP$ from the slits, depends on the phase difference $\delta(\mathbf{r},t) = k(s_1 - s_2) = k\,(AC)$. A section of the intensity profile $I(\mathbf{r},t)$ is shown, but all scales on this figure are distorted for emphasis. Bright fringes appear at P if $|s_1 - s_2|$ equals an integral multiple of the wavelength.

some are found deposited on the screen at locations that they could not reach if they moved along classical paths through either slit. The appearance of the interference fringes depends on the passage of the wave through both slits at once, just as a light wave goes through both slits in the analogous optical Young interference experiment. If the wave describes the behavior of a single particle, it follows that we cannot decide through which of the two slits the particle has gone. If we try to avoid this consequence by determining experimentally with some subtle monitoring device through which slit the particle has passed, we change the wave drastically and destroy the interference pattern. A single particle now goes definitely through one slit or the other, and the accumulation of a large number of particles on the screen will result in two well-separated traces. Exactly the same traces are obtained if we block one slit at a time, thereby predetermining the path of a particle. In the language of the principle of complementarity, the conditions under which the interference pattern is produced forbid a determination of the slit through which the particle passes.

This impressionistic qualitative description will be put on a firmer footing in Chapters 9 and 10, and again in Chapter 23, but the basic feature should be clear from the present example: Wave aspect and particle aspect in one and the same thing are compatible only if we forego asking questions that have no meaning (such as: "Let us see the interference fringes produced by particles whose paths through an arrangement of slits we have followed!")

An alternative view of the probability interpretation maintains that the wave ψ describes only the statistical behavior of a collection or ensemble of particles, allowing, at least in principle, for a more detailed and deterministic account of the motion of the single systems that make up the ensemble than is sanctioned by quantum mechanics. For example, in the two-slit experiment some hidden property of the individual particles would then presumably be responsible for the particular trajectories that the particles follow. Generally, there would be the potential for a more refined description of the individual member systems of the ensemble. From the confrontation between experiments on an important class of quantum systems and a penetrating theoretical analysis that is based on minimal assumptions (first undertaken by John Bell), we know that such a more "complete" description—broadly referred to as *realistic* or *ontological*—cannot succeed unless it allows for something akin to action-at-a-distance between systems that are widely separated in space. On the other hand, such nonlocal features arise naturally in quantum mechanics, and no evidence for any underlying *hidden variables* has ever been found. Thus, it is reasonable to suppose that ψ fully describes the statistical behavior of single systems.

To summarize, the single-particle probability doctrine of quantum mechanics asserts that the indetermination of which we have just given an example is a property inherent in nature and not merely a profession of our temporary ignorance from which we expect to be relieved by a future and more complete description. This interpretation thus denies the possibility of a more "complete" theory that would encompass the innumerable experimentally verified predictions of quantum mechanics but would be free of its supposed defects, the most notorious "imperfection" of quantum mechanics being the abandonment of strict classical determinism. Since the propositions of quantum mechanics are couched in terms of probabilities, the acquisition, processing, and evaluation of information are inherent in the theoretical description of physical processes. From this point of view, quantum physics can be said to provide a more comprehensive and complete account of the world than that

aspired to by classical physics. Loose talk about these issues may lead to the impression that quantum mechanics must forsake the classical goal of a wholly rational description of physical processes. Nothing could be further from the truth, as this book hopes to demonstrate.[4,5]

[4]The following list is a sample of books that will contribute to an understanding of quantum mechanics, including some by the pioneers of the field:

L. E. Ballentine, *Quantum Mechanics*, Prentice Hall, Englewood Cliffs, N.J., 1990.
A. Bohm, *Quantum Mechanics*, 2nd ed., Springer-Verlag, Berlin, 1986.
Niels Bohr, *Atomic Physics and Human Knowledge*, John Wiley, New York, 1958.
P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4th ed., Clarendon Press, Oxford, 1958.
Kurt Gottfried, *Quantum Mechanics*, Volume I, W. A. Benjamin, New York, 1966.
H. S. Green, *Matrix Mechanics*, Noordhoff, Groningen, 1965.
Walter Greiner and Berndt Müller, *Quantum Mechanics, Symmetries*, Springer-Verlag, Berlin, 1989.
Werner Heisenberg, *The Physical Principles of the Quantum Theory*, University of Chicago Press, 1930, translated by C. Eckart and C. Hoyt, Dover reprint, 1949.
Barry Holstein, *Topics in Advanced Quantum Mechanics*, Addison-Wesley, Reading, Mass., 1992.
Max Jammer, *The Conceptual Development of Quantum Mechanics*, McGraw-Hill, New York, 1966.
H. A. Kramers, *Quantum Mechanics*, Interscience, New York, 1957.
L. D. Landau and E. M. Lifshitz, *Quantum Mechanics*, Addison-Wesley, Reading, Mass., 1958, translated by J. B. Sykes and J. S. Bell.
Rubin H. Landau, *Quantum Mechanics II*, John Wiley, New York, 1990.
A. Messiah, *Quantum Mechanics*, North-Holland, Amsterdam, 1961 and 1962, Vol. I translated by G. Temmer, Vol. II translated by J. Potter.
Roland Omnès, *The Interpretation of Quantum Mechanics*, Princeton University Press, 1994.
Wolfgang Pauli, *Die allgemeinen Prinzipien der Wellenmechanik*, Vol. 5/1 of *Encyclopedia of Physics*, pp. 1–168, Springer-Verlag, Berlin, 1958.
J. J. Sakurai, with San Fu Tuan, ed., *Modern Quantum Mechanics*, revised ed., Benjamin/Cummings, New York, 1994.
B. L. van der Waerden, *Sources of Quantum Mechanics*, North-Holland, Amsterdam, 1967.
J. M. Ziman, *Elements of Advanced Quantum Theory*, Cambridge University Press, 1969.

[5]Books suitable for introductory study of quantum mechanics include:

Claude Cohen-Tannoudji, Bernard Diu, and Frank Laloë, *Quantum Mechanics*, Volumes I and II, John Wiley, New York, 1977.
R. H. Dicke and J. P. Wittke, *Introduction to Quantum Mechanics*, Addison-Wesley, Reading, Mass., 1960.
James M. Feagin, *Quantum Mechanics with Mathematica*, Springer-Verlag, 1993.
Stephen Gasiorowicz, *Quantum Physics*, 2nd ed., John Wiley, New York, 1996.
Walter Greiner, *Quantum Mechanics, An Introduction*, Springer-Verlag, Berlin, 1989.
D. Griffiths, *Introduction to Quantum Mechanics*, Prentice-Hall, Englewood Cliffs, N.J., 1995.
Thomas F. Jordan, *Quantum Mechanics in Simple Matrix Form*, John Wiley, New York, 1986.
Michael Morrison, *Understanding Quantum Mechanics*. Prentice-Hall, Englewood, N.J., 1990.
David A. Park, *Introduction to the Quantum Theory*, 3rd ed., McGraw-Hill, New York, 1992.
J. L. Powell and B. Crasemann, *Quantum Mechanics*, Addison-Wesley, Reading, Mass., 1961.
R. Shankar, *Principles of Quantum Mechanics*, Plenum, New York, 1980.
Richard W. Robinett, *Quantum Mechanics, Classical Results, Modern Systems, and Visualized Examples*, Oxford University Press, 1997.
John S. Townsend, *A Modern Approach to Quantum Mechanics*, McGraw-Hill, New York, 1992.

Problems

1. To what velocity would an electron (neutron) have to be slowed down, if its wavelength is to be 1 meter? Are matter waves of macroscopic dimensions a real possibility?
2. For the observation of quantum mechanical Bose-Einstein condensation, the interparticle distance in a gas of noninteracting atoms must be comparable to the de Broglie wavelength, or less. How high a particle density is needed to achieve these conditions if the atoms have mass number A = 100 and are at a temperature of 100 nanokelvin?

CHAPTER 2

Wave Packets, Free Particle Motion, and the Wave Equation

Building on previous experience with wave motion and simple Fourier analysis, we develop the quantum mechanical description of the motion of free particles. The correspondence between quantum and classical motion serves as a guide in the construction, by superposition of harmonic waves, of wave packets that propagate like classical particles but exhibit quantum mechanical spreading in space and time. Heisenberg's uncertainty relations and the Schrödinger wave equation make their first appearance.

1. ***The Principle of Superposition.*** We have learned that it is reasonable to suppose that a free particle of momentum $\mathbf{p}$ is associated with a harmonic plane wave. Defining a vector $\mathbf{k}$ which points in the direction of wave propagation and has the magnitude

$$k = \frac{2\pi}{\lambda} \tag{2.1}$$

we may write the fundamental de Broglie relation as

$$\mathbf{p} = \hbar\mathbf{k} \tag{2.2}$$

The symbol $\hbar$ is common in quantum physics and denotes the frequently recurring constant

$$\hbar = \frac{h}{2\pi} = 1.0545 \times 10^{-34} \text{ J sec} = 6.5819 \times 10^{-22} \text{ MeV sec} = 0.66 \times 10^{-15} \text{ eV sec}$$

It is true that diffraction experiments do not give us any direct information about the detailed dependence on space and time of the periodic disturbance that produces the alternately "bright" and "dark" fringes, but all the evidence points to the correctness of the simple inferences embodied in (2.1) and (2.2). The comparison with optical interference suggests that the fringes come about by linear superposition of two waves, a point of view that has already been stressed in Section 1.2 in the discussion of the simple two-slit interference experiment (Figure 1.1).

Mathematically, these ideas are formulated in the following fundamental assumption about the wave function $\psi(x, y, z, t)$: If $\psi_1(x, y, z, t)$ and $\psi_2(x, y, z, t)$ describe two waves, their sum $\psi(x, y, z, t) = \psi_1 + \psi_2$ also describes a possible physical situation. This assumption is known as the *principle of superposition* and is illustrated by the interference experiment of Figure 1.1. The intensity produced on the screen by opening only one slit at a time is $|\psi_1|^2$ or $|\psi_2|^2$. When both slits are open, the intensity is determined by $|\psi_1 + \psi_2|^2$. This differs from the sum of the two intensities, $|\psi_1|^2 + |\psi_2|^2$, by the interference terms $\psi_1\psi_2^* + \psi_2\psi_1^*$. (An asterisk will denote complex conjugation throughout this book.)

A careful analysis of the interference experiment would require detailed consideration of the boundary conditions at the slits. This is similar to the situation

found in wave optics, where the uncritical use of Huygens' principle must be justified by recourse to the wave equation and to Kirchhoff's approximation. There is no need here for such a thorough treatment, because our purpose in describing the idealized two-slit experiment was merely to show how the principle of superposition accounts for some typical interference and diffraction phenomena. Such phenomena, when actually observed as in diffraction of particles by crystals, are impressive direct manifestations of the wave nature of matter.

We therefore adopt the principle of superposition to guide us in developing quantum mechanics. The simplest type of wave motion to which it will be applied is an infinite harmonic plane wave propagating in the positive x direction with wavelength $\lambda = 2\pi/k$ and frequency ω. Such a wave is associated with the motion of a free particle moving in the x direction with momentum $p = \hbar k$, and its most general form is

$$\psi_1(x, t) = \cos(kx - \omega t) + \delta \sin(kx - \omega t) \tag{2.3}$$

A plane wave moving in the negative x direction would be written as

$$\begin{aligned} \psi_2(x, t) &= \cos(-kx - \omega t) + \delta \sin(-kx - \omega t) \\ &= \cos(kx + \omega t) - \delta \sin(kx + \omega t) \end{aligned} \tag{2.4}$$

The coordinates y and z can obviously be omitted in describing these one-dimensional processes.

An arbitrary displacement of x or t should not alter the physical character of these waves, which describe, respectively, a particle moving uniformly in the positive and negative x directions, nor should the phase constants of these waves have any physical significance. Hence, it must be required that

$$\cos(kx - \omega t + \varepsilon) + \delta \sin(kx - \omega t + \varepsilon) = a(\varepsilon)[\cos(kx - \omega t) + \delta \sin(kx - \omega t)]$$

for all values of x, t, and ε. Comparing coefficients of $\cos(kx - \omega t)$ and of $\sin(kx - \omega t)$, we find that this last equation leads to

$$\cos \varepsilon + \delta \sin \varepsilon = a \quad \text{and} \quad \delta \cos \varepsilon - \sin \varepsilon = a\delta$$

These equations are compatible for all ε only if $\delta^2 = -1$ or $\delta = \pm i$. We choose the solution $\delta = i$, $a = e^{i\varepsilon}$ and are thus led to the conclusion that ψ waves describing free particle motion must in general be complex. We have no physical reason for rejecting complex-valued wave functions because, unlike elastic displacements or electric field vectors, the ψ waves of quantum mechanics cannot be observed directly, and the observed diffraction pattern presumably measures only the intensity $|\psi|^2$.

Summarizing our conclusions, we see that in this scheme an infinite harmonic plane wave propagating toward increasing x is

$$\psi_1(x, t) = Ae^{i(kx - \omega t)} \tag{2.5}$$

and a wave propagating toward decreasing x is

$$\psi_2(x, t) = Be^{-i(kx + \omega t)} = Be^{i(-kx - \omega t)} \tag{2.6}$$

The initial values of these waves are

$$\psi_1(x, 0) = Ae^{ikx} \quad \text{and} \quad \psi_2(x, 0) = Be^{-ikx}$$

respectively. The (circular) frequency of oscillation is ω, which thus far we have not brought into connection with any physically observable phenomenon. Generally,

it will be a function of k (see Section 2.3). We will refer to wave functions like (2.5) and (2.6) briefly as *plane waves*.

A plane wave propagating in an arbitrary direction has the form

$$\psi(\mathbf{r}, t) = \psi(x, y, z, t) = Ae^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} = Ae^{i(k_x x+k_y y+k_z z-\omega t)} \tag{2.7}$$

Equations (2.5) and (2.6) are special cases of this with $k_y = k_z = 0$ and $k_x = \pm k$.

It should be stressed that with the acceptance of complex values for ψ, we are by no means excluding wave functions that are real-valued, at least at certain times. For example, if the initial wave function is $\psi(x, 0) = \cos kx$, this can be written as the sum of two exponential functions:

$$\psi(x, 0) = \cos kx = \frac{1}{2}e^{ikx} + \frac{1}{2}e^{-ikx} \tag{2.8}$$

According to the principle of superposition, this must be an acceptable wave function. How does it develop in time? The principle of superposition suggests a simple answer to this important question: *Each of the two or more waves, into which ψ can be decomposed at $t = 0$, develops independently, as if the other component(s) were not present.* The wave function, which at $t = 0$ satisfies the initial condition $\psi(x, 0) = \cos kx$, thus becomes for arbitrary t:

$$\psi(x, t) = \frac{1}{2}e^{i(kx-\omega t)} + \frac{1}{2}e^{-i(kx+\omega t)} \tag{2.9}$$

Note that this is not the same as $\cos(kx - \omega t)$. This rule, to which we shall adhere and which we shall generalize, ensures that $\psi(x, 0)$ determines the future behavior of the wave uniquely. If this formulation is correct, we expect that the complex ψ waves may be described by a linear differential equation which is of the first order in time.[1]

2. ***Wave Packets and the Uncertainty Relations.*** The foregoing discussion points to the possibility that by allowing the wave function to be complex, we might be able to describe the state of motion of a particle at time t completely by $\psi(\mathbf{r}, t)$. The real test of this assumption is, of course, its success in accounting for experimental observations. Strong support for it can be gained by demonstrating that the correspondence with classical mechanics can be established within the framework of this theory.

To this end, we must find a way of making $\psi(\mathbf{r}, t)$ describe, at least approximately, the classical motion of a particle that has both reasonably definite position and reasonably definite momentum. The plane wave (2.7) corresponds to particle motion with momentum, which is precisely defined by (2.2); but having amplitudes $|\psi| = \text{const.}$ for all $\mathbf{r}$ and t, the infinite harmonic plane waves (2.7) leave the position of the particle entirely unspecified. By superposition of several different plane waves, a certain degree of localization can be achieved, as the fringes on the screen in the diffraction experiment attest.

[1]Maxwell's equations are of the first order even though the functions are real, but this is accomplished using both **E** and **B** in coupled equations, instead of describing the field by just one function. The heat flow equation is also of the first order in time and describes the behavior of a real quantity, the temperature; but none of its solutions represent waves propagating without damping in one direction like (2.7).

The mathematical tools for such a synthesis of localized compact wave packets by the superposition of plane waves of different wave number are available in the form of Fourier analysis. Assuming for simplicity that only one spatial coordinate, x, need be considered, we can write

$$\boxed{\psi(x, 0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \phi(k_x) e^{ik_x x}\, dk_x} \tag{2.10}$$

and the inverse formula

$$\boxed{\phi(k_x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \psi(x, 0) e^{-ik_x x}\, dx} \tag{2.11}$$

under very general conditions to be satisfied by $\psi(x, 0)$.[2] Formulas (2.10) and (2.11) show that $\phi(k_x)$ determines the initial wave function $\psi(x, 0)$ and vice versa.

We now assume that $\phi(k_x)$ is a real positive function and that it has a shape similar to Figure 2.1, i.e., a symmetric distribution of k_x about a mean value $\bar{k}_x$. Making the change of variable

$$u = k_x - \bar{k}_x$$

we may write

$$\psi(x, 0) = \frac{1}{\sqrt{2\pi}} e^{i\bar{k}_x x} \int_{-\infty}^{+\infty} \phi(u + \bar{k}_x) e^{iux}\, du \tag{2.12}$$

This is a wave packet whose absolute value is shown in Figure 2.2. The particle is most likely to be found at a position where ψ is appreciable. It is easy to see for any number of simple examples that the width Δk_x of the amplitude ϕ and the width Δx of the wave packet ψ stand in a reciprocal relationship:

$$\Delta x\, \Delta k_x \cong 1 \tag{2.13}$$

For a proof of the uncertainty relation (2.13), note that the integral in (2.12) is an even real function of x. Let us denote by $(-x_0, x_0)$ the range of x for which ψ is appreciably different from zero (see Figure 2.2). Since ϕ is appreciably different from zero only in a range Δk_x centered at $u = 0$, where $k_x = \bar{k}_x$, the phase of e^{iux} in the integrand varies from $-x\, \Delta k_x/2$ to $+ x\, \Delta k_x/2$, i.e., by an amount $x\, \Delta k_x$ for any fixed value of x. If $x_0\, \Delta k_x$ is less than $\cong 1$, no appreciable cancellations in the integrand occur (since ϕ is positive definite). On the other hand, when $x_0\, \Delta k_x \gg 1$, the phase goes through many periods as u ranges from $-\Delta k_x/2$ to $+\Delta k_x/2$ and violent oscillations of the e^{iux} term occur, leading to destructive interference. Hence, the largest variations that the phase ever undergoes are $\pm x_0\, \Delta k_x$; this happens at the ends of the wave packet. Denoting by Δx the range $(-x_0, x_0)$, it follows that the widths of $|\psi|$ and ϕ are effectively related by (2.13).

[2]For a review of Fourier analysis, see any text in mathematical physics, e.g., Bradbury (1984), Arfken (1985), and Hassani (1991). For Eq. (2.11) to be the inverse of (2.10), it is sufficient that $\psi(x, 0)$ be a function that is sectionally continuous. The basic formulas are summarized in Section 1 of the Appendix.

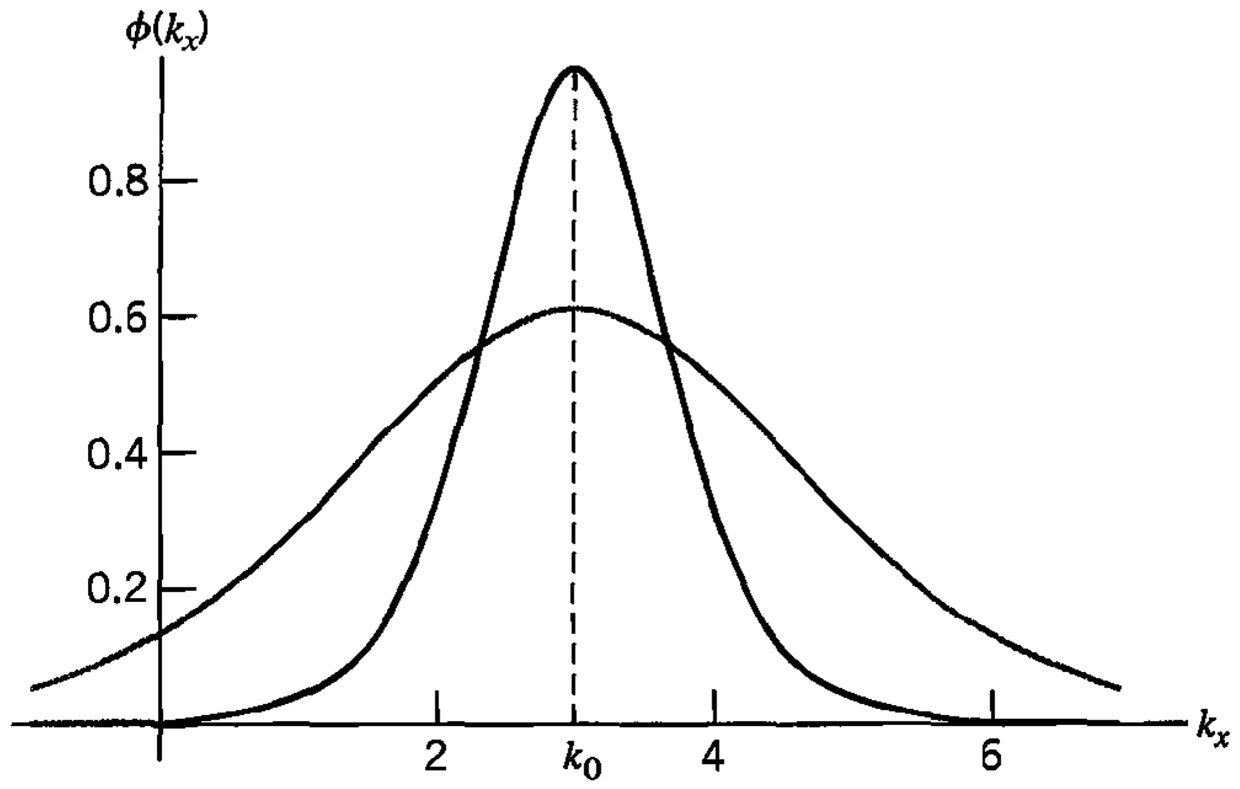


Figure 2.1. Example of a one-dimensional momentum distribution. The function $\phi(k_x) = \sqrt{a/2}\,[\cosh a\,(k_x - k_0)]^{-1}$ represents a wave packet moving in the positive x direction with mean wave number $k_0 = 3$. The function $\phi(k_x)$ is normalized so that $\int_{-\infty}^{+\infty} |\phi(k_x)|^2\, dk_x = 1$. The two functions correspond to $a = 0.72$ and $\sqrt{\pi}$ (heavy line).

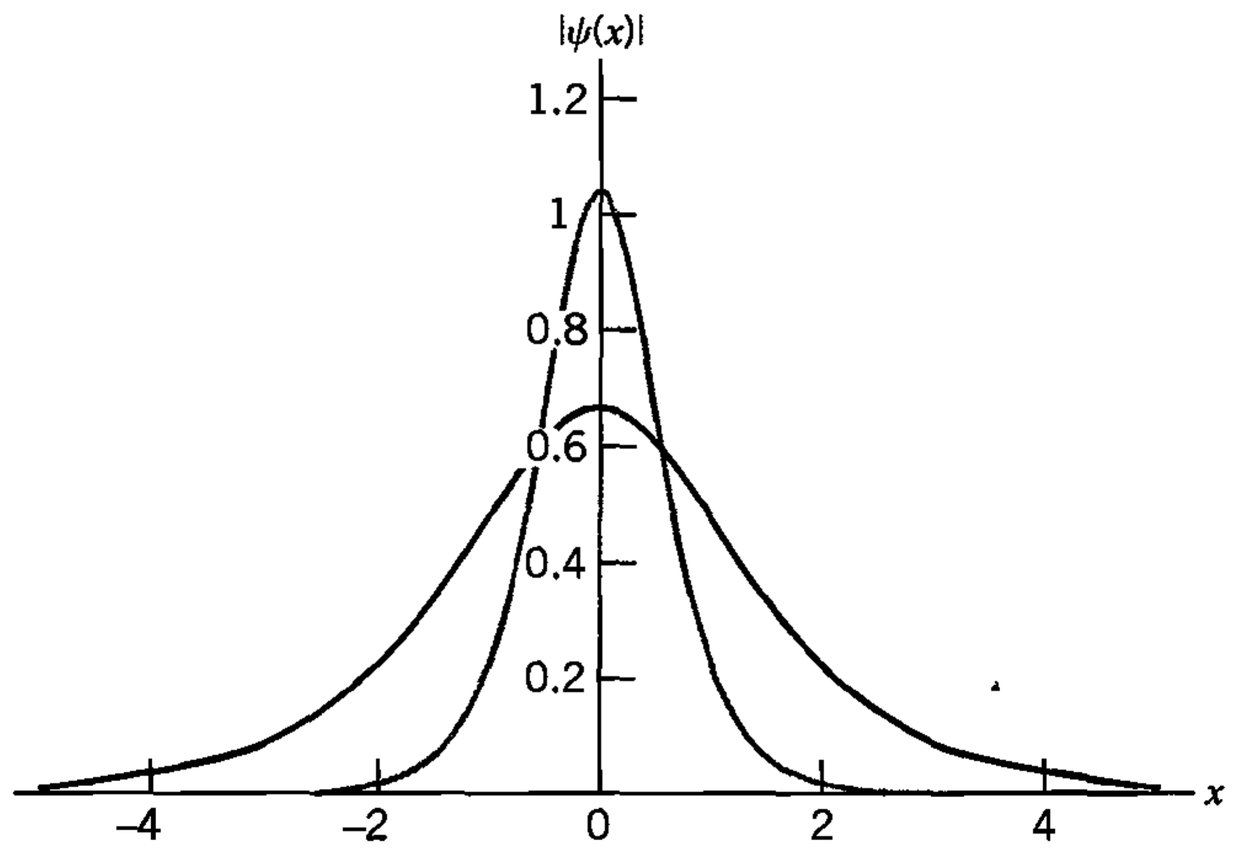


Figure 2.2. Normalized wave packet corresponding to the momentum distributions of Figure 2.1. The plotted amplitude, $|\psi(x)| = \sqrt{\dfrac{\pi}{4a}}\left(\cosh\dfrac{\pi x}{2a}\right)^{-1}$, modulates the plane wave e^{ik_0x}. Note the reciprocity of the widths of the corresponding curves in Figures 2.1 and 2.2.

Exercise 2.1. Assume $\phi(k_x) = \sqrt{2\pi}$ for $\bar{k}_x - \delta \leq k_x \leq \bar{k}_x + \delta$, and $\phi(k_x) = 0$ for all other values of k_x. Calculate $\psi(x, 0)$, plot $|\psi(x, 0)|^2$ for several values of δ, and show that (2.13) holds if Δx is taken as the width at half maximum.[3]

Exercise 2.2. Assume $\psi(x, 0) = e^{-\lambda|x|}$ for $-\infty < x < +\infty$. Calculate $\phi(k_x)$ and show that the widths of ψ and ϕ, reasonably defined, satisfy the reciprocal relation (2.13).

[3]Note that if $(\Delta x)^2$ is taken as the variance of x, as in Section 10.5, its value will be infinite for the wave packets defined in Exercises 2.1 and 2.3.

The initial wave function $\psi(x, 0)$ thus describes a particle that is localized within a distance Δx about the coordinate origin. This has been accomplished at the expense of combining waves of wave numbers in a range Δk_x about $\bar{k}_x$. The relation (2.13) shows that we can make the wave packet define a position more sharply only at the cost of broadening the spectrum of k_x-values which must be admitted. Any hope that these consequences of (2.13) might be averted by choosing $\phi(k_x)$ more providentially, perhaps by making it asymmetric or by noting that in general it is complex-valued, has no basis in fact. On the contrary, we easily see (and we will rigorously prove in Chapter 10) that in general

$$\Delta x \, \Delta k_x \geq \frac{1}{2} \tag{2.14}$$

The product $\Delta x \, \Delta k_x$ assumes a value near its minimum of 1/2 only if the absolute value of ϕ behaves as in Figure 2.1 and its phase is either constant or a linear function of k_x.

Exercise 2.3. Assume

$$\phi(k_x) = \sqrt{2\pi} \text{ for } \bar{k}_x - \frac{\delta}{2} \leq k_x \leq \bar{k}_x + \frac{\delta}{2}$$

$$= -\sqrt{2\pi} \text{ for } \bar{k}_x - \delta \leq k_x \leq \bar{k}_x - \frac{\delta}{2} \text{ and } \bar{k}_x + \frac{\delta}{2} \leq k_x \leq \bar{k}_x + \delta$$

and $\phi = 0$ for all other values of k_x. Calculate $\psi(x, 0)$ and show that, although the width of the k_x-distribution is the same as in Exercise 2.1, the width of $\psi(x, 0)$ is greater in accordance with the inequality (2.14).

The fact that in quantum physics both waves and particles appear in the description of the same thing has already forced us to abandon the classical notion that position and momentum can be defined with perfect precision simultaneously. Equation (2.14) together with the equation $p_x = \hbar k_x$ expresses this characteristic property of quantum mechanics quantitatively:

$$\Delta x \, \Delta p_x \geq \frac{\hbar}{2} \tag{2.15}$$

This *uncertainty relation* suggests that when ψ is the sum of plane waves with different values of k_x, the wave function cannot represent a particle of definite momentum (for if it did ψ would have to be a harmonic plane wave). It also suggests that its spread-out momentum distribution is roughly pictured by the behavior of $|\phi|^2$. If $|\phi|^2$ has its greatest magnitude at $\bar{k}_x$, the particle is most likely to have the momentum $\hbar k_x$. Obviously, we will have to make all these statements precise and quantitative as we develop the subject.

Exercise 2.4. By choosing reasonable numerical values for the mass and velocity, convince yourself that (2.15) does not in practice impose any limitations on the precision with which the position and momentum of a macroscopic body can be determined.

For the present, it suffices to view a peak in ψ as a crudely localized particle and a peak in ϕ as a particle moving with an approximately defined velocity. The *uncertainty relation* (2.15) limits the precision with which position and momentum

can be simultaneously ascribed to the particle. Generally, both quantities are fuzzy and indeterminate (*Heisenberg uncertainty principle*).

We have already discussed in Chapter 1 what this means in terms of experiments. The function $|\psi|^2$ is proportional to the probability of finding the particle at position x. Yet upon measurement the particle will always be found to have a definite position. Similarly, quantum mechanics does not deny that precision measurements of momentum are feasible even when the particle is not represented by a plane wave of sharp momentum. Rather, $|\phi|^2$ is proportional to the probability of finding the particle to have momentum $\hbar k_x$.

Under these circumstances, must we admit that the particle described by ϕ (or ψ) *has* no definite momentum (or position) when we merely are unable to determine a single value consistently in making the same measurement on identically prepared systems? Could the statistical uncertainties for the individual systems be reduced below their quantum mechanical values by a greater effort? Is there room in the theory for supplementing the statistical quantum description by the specification of further (''hidden'') variables, so that two systems that are in the same quantal state (ψ or ϕ) may be found to be distinct in a more refined characterization?

These intriguing questions have been hotly debated since the advent of quantum mechanics.[4] We now know (*Bell's theorem*) that the most natural kinds of hidden-variable descriptions are incompatible with some of the subtle predictions of quantum mechanics. Since these predictions have been borne out experimentally to high accuracy, we adopt as the central premise of quantum mechanics that

> *For any given state* (ψ or ϕ) *the measurement of a particular physical quantity results with calculable probability in a numerical value belonging to a range of possible measured values.*

No technical or mathematical ingenuity can presumably devise the means of giving a sharper and more accurate account of the physical state of a single system than that permitted by the wave function and the uncertainty relation. These claims constitute a *principle* which by its very nature cannot be proved, but which is supported by the enormous number of verified consequences derived from it.

Bohr and Heisenberg were the first to show in detail, in a number of interesting thought experiments, how the finite value of $\hbar$ in the uncertainty relation makes the coexistence of wave and particle both possible and necessary. These idealized experiments demonstrate explicitly how any effort to design a measurement of the momentum component p_x with a precision set by Δp_x inescapably limits the precision of a simultaneous measurement of the coordinate x to $\Delta x \geq \hbar/2\Delta p_x$, and vice versa.

Illuminating as Bohr's thought experiments are, they merely illustrate the important discovery that underlies quantum mechanics, namely, that the behavior of a material particle is described completely by its wave function ψ or ϕ (suitably modified to include the spin and other degrees of freedom, if necessary), but generally not by its precise momentum or position. Quantum mechanics contends that the wave function contains the maximum amount of information that nature allows us concerning the behavior of electrons, photons, protons, neutrons, quarks, and the like. Broadly, these tenets are referred to as the *Copenhagen interpretation* of quantum mechanics.

3. ***Motion of a Wave Packet.*** Having placed the wave function in the center of our considerations, we must attempt to infer as much as possible about its properties.

[4]See Wheeler and Zurek (1985).

In three dimensions an initial wave packet $\psi(\mathbf{r}, 0)$ is represented by the Fourier integral:[5]

$$\psi(x, y, z, 0) = \psi(\mathbf{r}, 0) = \frac{1}{(2\pi)^{3/2}} \int \phi(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}}\, d^3k \tag{2.16}$$

and the inverse formula

$$\phi(\mathbf{k}) = \frac{1}{(2\pi)^{3/2}} \int \psi(\mathbf{r}, 0) e^{-i\mathbf{k}\cdot\mathbf{r}}\, d^3r \tag{2.17}$$

If the particle which the initial wave function (2.16) describes is free, then by the rule of Section 2.1 (the superposition principle) each component plane wave contained in (2.16) propagates independently of all the others according to the prescription of (2.7). The wave function at time t becomes

$$\boxed{\psi(\mathbf{r}, t) = \frac{1}{(2\pi)^{3/2}} \int \phi(\mathbf{k}) e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)}\, d^3k} \tag{2.18}$$

Formula (2.18) is, of course, incomplete until we determine and specify the dependence of ω on $\mathbf{k}$. This determination will be made on physical grounds.

If ψ is of appreciable magnitude only in coordinate ranges Δx, Δy, Δz, then ϕ is sensibly different from zero only in k_x lies in a range Δk_x, k_y in a range Δk_y, and k_z in a range Δk_x, such that

$$\Delta x\, \Delta k_x \geq 1, \qquad \Delta y\, \Delta k_y \geq 1, \qquad \Delta z\, \Delta k_z \geq 1 \tag{2.19}$$

These inequalities are generalizations of the one-dimensional case. By (2.2), $\Delta\mathbf{p} = \hbar\Delta\mathbf{k}$; hence, in three dimensions the uncertainty relations are

$$\boxed{\Delta x\, \Delta p_x \geq \hbar, \qquad \Delta y\, \Delta p_y \geq \hbar, \qquad \Delta z\, \Delta p_z \geq \hbar} \tag{2.20}$$

Let us consider a wave packet whose Fourier inverse $\phi(\mathbf{k})$ is appreciably different from zero only in a limited range $\Delta\mathbf{k}$ near the mean wave vector $\hbar\bar{\mathbf{k}}$. In coordinate space, the wave packet $\psi(\mathbf{r}, t)$ must move approximately like a classical free particle with mean momentum $\hbar\bar{\mathbf{k}}$. To see this behavior we expand $\omega(\mathbf{k})$ about $\bar{\mathbf{k}}$:

$$\begin{aligned} \omega(\mathbf{k}) &= \omega(\bar{\mathbf{k}}) + (\mathbf{k} - \bar{\mathbf{k}}) \cdot (\nabla_{\mathbf{k}}\omega)_{\mathbf{k}=\bar{\mathbf{k}}} + \cdots \\ &= \bar{\omega} + (\mathbf{k} - \bar{\mathbf{k}}) \cdot \nabla_{\bar{\mathbf{k}}}\bar{\omega} + \cdots \end{aligned} \tag{2.21}$$

with obvious abbreviations.

If we substitute the first two terms of this expansion into (2.18), we obtain

$$\psi(\mathbf{r}, t) = \exp(-i\bar{\omega}\, t + i\bar{\mathbf{k}} \cdot \nabla_{\bar{\mathbf{k}}}\bar{\omega}\, t)\psi(\mathbf{r} - \nabla_{\bar{\mathbf{k}}}\bar{\omega}\, t, 0) \tag{2.22}$$

Ignoring the phase factor in front, we see that this describes a wave packet performing a uniform translational motion, without any change of shape and with the group velocity

$$\mathbf{v}_g = \nabla_{\bar{\mathbf{k}}}\bar{\omega} \tag{2.23}$$

Our interpretation, based on the correspondence between classical and quantum mechanics, leads us to identify the group velocity of the ψ wave with the mean particle velocity $\hbar\mathbf{k}/m$, where m is the mass of the particle and nonrelativistic particle motion

[5] $d^3k = dk_x\, dk_y\, dk_z$ denotes the volume element in $\mathbf{k}$-space, and $d^3r = dx\, dy\, dz$ is the volume element in (position) coordinate space.

has been assumed. Since this relation must hold for an arbitrary choice of $\mathbf{k}$, we may omit the averaging bars and require

$$m\nabla_{\mathbf{k}}\omega = \hbar\mathbf{k} \tag{2.24}$$

Integrating this, we obtain

$$\omega(\mathbf{k}) = \frac{\hbar}{2m}(k_x^2 + k_y^2 + k_z^2) + \text{const.} \tag{2.25}$$

or

$$\hbar\omega(\mathbf{k}) = \frac{\hbar^2 k^2}{2m} + V = \frac{p^2}{2m} + V \tag{2.26}$$

where V is a constant, which may be thought of as an arbitrary constant potential energy. We may assume that $\hbar\omega$ represents the particle energy,

$$E = \hbar\omega \tag{2.27}$$

A similar equation has already been found to hold for photons. This is not as surprising as it may seem, for (2.2) and (2.27), although obtained in this chapter for nonrelativistic particle mechanics, are Lorentz-invariant and thus remain valid in the relativistic case. Indeed, they merely express the proportionality of two four-vectors:

$$\left(\frac{E}{c}, \mathbf{p}\right) = \hbar\left(\frac{\omega}{c}, \mathbf{k}\right) \tag{2.28}$$

Hence their great universality.

4. ***The Uncertainty Relations and the Spreading of Wave Packets.*** We must examine the conditions under which it is legitimate to neglect the quadratic and higher terms in (2.21). Without these corrections the wave packet moves uniformly without change of shape, and we should ask what happens in time when the neglect of the quadratic terms is no longer justified.

For the nonrelativistic case considered here, the expansion (2.21) actually terminates because (with $V = 0$)

$$\omega(\mathbf{k}) = \frac{\hbar k^2}{2m} = \frac{\hbar\bar{k}^2}{2m} + \frac{\hbar\bar{\mathbf{k}}}{m}\cdot(\mathbf{k} - \bar{\mathbf{k}}) + \frac{\hbar}{2m}(\mathbf{k} - \bar{\mathbf{k}})^2 \tag{2.29}$$

Thus, if quadratic terms are retained, the behavior of the wave packet is given by an exact formula.

For simplicity, let us consider a one-dimensional wave packet for which $\phi(\mathbf{k}) = 0$ unless $\mathbf{k} = k\hat{\mathbf{x}}$ ($k_y = k_z = 0$), in which case the neglected term in (2.21), multiplied by t, is

$$\frac{\hbar}{2m}(k_x - \bar{k}_x)^2 t = \frac{\hbar}{2m}(\Delta k_x)^2 t = \frac{(\Delta p_x)^2}{2m\hbar} t \tag{2.30}$$

This term generally contributes to the exponent in the integrand of (2.18). An exponent can be neglected only if it is much less than unity in absolute value. Thus, for nonrelativistic particles, the wave packet moves without appreciable change of shape for times t such that

$$|t| \ll \frac{m\hbar}{(\Delta p_x)^2}$$

or

$$|t|\ \Delta v_x = \frac{\Delta p_x}{m}\,|t| \ll \frac{\hbar}{\Delta p_x} \cong \Delta x \tag{2.31}$$

The product $|t|\ \Delta v_x$ represents an uncertainty in the position at time t, over and beyond the initial uncertainty Δx, and is attributable to the spread in initial velocity. If $|t|$ grows too large, condition (2.31) is violated, the second-order contribution to the phase can no longer be neglected, and the wave packet is broadened so that for very large $|t|$,

$$\Delta x(t)\Delta k_x > 1 \tag{2.32}$$

This means that as t increases from the distant past ($t \to -\infty$) to the remote future ($t \to +\infty$), the wave packet first contracts and eventually spreads.

Exercise 2.5. Consider a wave packet satisfying the relation $\Delta x\ \Delta p_x \approx \hbar$. Show that if the packet is not to spread appreciably while it passes through a fixed position, the condition $\Delta p_x \ll p_x$ must hold.

Exercise 2.6. Can the atoms in liquid helium at 4 K (interatomic distance about 0.1 nanometer = 1Å) be adequately represented by nonspreading wave packets, so that their motion can be described classically?

Exercise 2.7. Make an estimate of the lower bound for the distance Δx, within which an object of mass m can be localized for as long as the universe has existed ($\approx 10^{10}$ years). Compute and compare the values of this bound for an electron, a proton, a one-gram object, and the entire universe.

The uncertainty relation (2.15) has a companion that relates uncertainties in time and energy. The kinetic energy is $E = p_x^2/2m$, and it is uncertain by an amount

$$\Delta E = \frac{p_x}{m}\,\Delta p_x = v_x\,\Delta p_x \tag{2.33}$$

If a monitor is located at a fixed position, he or she ''sees'' the wave packet sweeping through his or her location. The determination of the time at which the particle passes the monitor must then be uncertain by an amount:

$$\Delta t = \frac{\Delta x}{v_x} \geq \frac{\hbar}{v_x\,\Delta p_x} = \frac{\hbar}{\Delta E}$$

Hence,

$$\Delta E\ \Delta t \geq \hbar \tag{2.34}$$

In this derivation, it was assumed that the wave packet does not spread appreciably in time Δt while it passes through the observer's position. According to (2.31), this is assured if

$$\frac{\hbar}{\Delta E} \leq \Delta t \ll \frac{m\hbar}{(\Delta p_x)^2}$$

With (2.33) this condition is equivalent to $\Delta E \ll E$. But the latter condition must be satisfied if we are to be allowed to speak of *the* energy of the particles in a beam at all (rather than a distribution of energies).

The derivation of the time-energy uncertainty relation given here is narrow and precise. More generally, a similar inequality relates the indeterminancy Δt in the time of occurrence of an event at the quantum level (e.g., the decay of a nucleus or particle) to the spread ΔE in the energy associated with the process, but considerable care is required to establish valid quantitative statements (see Chapter 19).

Exercise 2.8. Assume that in Figures 2.1 and 2.2 the units for the abscissa are Å^{-1} (for k_x) and Å (for x), respectively. If the particles described by the wave packets in these figures are neutrons, compute their mean velocity (in m/sec), the mean kinetic energy (in meV), the corresponding temperature (in Kelvin), and the energy spread (also in meV). Estimate the time scale for the spatial spread of this wave packet.

5. ***The Wave Equation for Free Particle Motion.*** Although the plane wave (Fourier) representation (2.18) gives the most general form of a *free particle* wave function, other representations are often useful. These are most conveniently obtained as solutions of a wave equation for this motion. We must find a linear partial differential equation that admits (2.18) as its general solution, provided that the relation between ω and k is given by the dispersion formula (2.26). To accomplish this, we need only establish the wave equation for the plane waves (2.7), since the linearity of the wave equation ensures that the superposition (2.18) satisfies the same equation.

The wave equation for the plane waves $e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$, with

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

with constant V, is obviously

$$i\hbar\frac{\partial\psi(\mathbf{r}, t)}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi(\mathbf{r}, t) + V\psi(\mathbf{r}, t) \tag{2.35}$$

This quantum mechanical wave equation is also known as the *time-dependent Schrödinger equation* for the free particle. As expected, it is a first-order differential equation in t, requiring knowledge of the initial wave function $\psi(\mathbf{r}, 0)$, but not its derivative, for its solution. That the solutions must, in general, be complex-valued is again manifest from the appearance of i in the differential equation. For $V = 0$, the quantum mechanical wave equation and the diffusion (or heat flow) equation become formally identical, but the presence of i ensures that the quantum mechanical wave equation has solutions with wave character.

The details of solving Eq. (2.35), using various different coordinate systems for the spatial variables, will be taken up in later chapters. Here we merely draw attention to an interesting alternative form of the wave equation, obtained by the transformation

$$\psi(\mathbf{r}, t) = e^{iS(\mathbf{r},t)/\hbar} \tag{2.36}$$

If this expression for the wave function is substituted in Eq. (2.35), the equation

$$\left\{\frac{\partial S(\mathbf{r}, t)}{\partial t} + \frac{[\nabla S(\mathbf{r}, t)]^2}{2m} - \frac{i\hbar}{2m}\nabla^2 S(\mathbf{r}, t) + V\right\}e^{iS(\mathbf{r},t)/\hbar} = 0 \tag{2.37}$$

is derived. The function $S(\mathbf{r}, t)$ is generally complex, but in special cases it may be real, as for instance when $\psi(\mathbf{r}, t)$ represents a plane wave, in which case we have

$$S(\mathbf{r}, t) = \hbar\mathbf{k}\cdot\mathbf{r} - \hbar\omega t = \mathbf{p}\cdot\mathbf{r} - Et \tag{2.38}$$

Equation (2.37) informs us that wherever and whenever the wave function $\psi(\mathbf{r}, t)$ does not vanish, the nonlinear differential equation

$$\frac{\partial S(\mathbf{r}, t)}{\partial t} + \frac{[\nabla S(\mathbf{r}, t)]^2}{2m} - \frac{i\hbar}{2m}\nabla^2 S(\mathbf{r}, t) + V = 0 \tag{2.39}$$

must hold. If the term in this equation containing i and $\hbar$, the hallmarks of quantum mechanics, were negligible, (2.39) would be the Hamilton-Jacobi equation for the Hamilton Principal Function of classical mechanics.[6]

To further explore the connection between classical and quantum mechanics for a free particle, we again consider the motion of a wave packet. Subject to the restrictions imposed by the uncertainty relations (2.19), we assume that both ψ and ϕ are fairly localized functions in coordinate and $\mathbf{k}$-space, respectively. In one dimension, Figures 2.1 and 2.2 illustrate such a state. The most popular prototype, however, is a state whose ψ wave function at the initial time $t = 0$ is a plane wave of mean momentum $\hbar\mathbf{k}$, modulated by a real-valued Gaussian function of width Δx and centered at x_0:

$$\psi(x, 0) = C\exp\left\{-\frac{(x - x_0)^2}{4(\Delta x)^2} + i\bar{k}_x x\right\} \tag{2.40}$$

(The dynamics of this particular wave packet is the content of Problem 1 at the end of this chapter.)

Somewhat more generally, we consider an initial one-dimensional wave packet that can be written in the form

$$\psi(x, 0) = e^{iS(x,0)/\hbar} = \exp[-\chi(x) + i\bar{k}_x x] \tag{2.41}$$

where $\chi(x)$ is a smooth real-valued function that has a minimum at x_0 and behaves as $\chi(\pm\infty) \to +\infty$ at large distances. In the approximation that underlies (2.22), for times short enough to permit neglect of the changing shape of the wave packet, the time development of $\psi\,(x, t) = e^{iS(x,\,t)/\hbar}$ is given by

$$\frac{S(x, t)}{\hbar} \approx -\bar{\omega}t + \bar{k}_x\frac{\partial\bar{\omega}}{\partial\bar{k}_x}t + \bar{k}_x\left(x - \frac{\partial\bar{\omega}}{\partial\bar{k}_x}t\right) + i\chi\left(x - \frac{\partial\bar{\omega}}{\partial\bar{k}_x}t\right) \tag{2.42}$$

If the localized wave packet is broad enough so that $\chi(x)$ changes smoothly and slowly over the distance of a mean de Broglie wavelength, we obtain from (2.42) for the partial derivatives of $S(x, t)$, to good approximation:

$$\frac{\partial S(x, t)}{\partial t} \approx -\hbar\bar{\omega} = -\bar{E} \tag{2.43}$$

and

$$\frac{\partial S(x, t)}{\partial x} \approx \hbar\bar{k}_x = \bar{p}_x \tag{2.44}$$

[6]To review the elements of classical Hamiltonian mechanics, see Goldstein (1980).

These equations hold for values of x near the minimum of χ, hence near the peak $|\psi|$. From them we deduce that in this approximation $S(x, t)$ satisfies the *classical* Hamilton-Jacobi equation

$$\frac{\partial S(x, t)}{\partial t} + \frac{1}{2m}\left[\frac{\partial S(x, t)}{\partial x}\right]^2 + V = 0 \tag{2.45}$$

This shows that for a broad class of wave packets a *semiclassical approximation* for the wave function may be used by identifying $S(x, t)$ with the Hamilton Principal Function $S(x, t)$ which corresponds to *classical* motion of the wave packet.

The present chapter shows that our ideas about wave packets can be made quantitative and consistent with the laws of classical mechanics when the motion of free particles is considered. We must now turn to an examination of the influence of forces and interactions on particle motion and wave propagation.

Problems

1. A one-dimensional initial wave packet with a mean wave number $\bar{k}_x$ and a Gaussian amplitude is given by

$$\psi(x, 0) = C \exp\left[-\frac{x^2}{4(\Delta x)^2} + i\bar{k}_x x\right]$$

Calculate the corresponding $\bar{k}_x$-distribution and $\psi(x, t)$, assuming free particle motion. Plot $|\psi(x, t)|^2$ as a function of x for several values of t, choosing Δx small enough to show that the wave packet spreads in time, while it advances according to the classical laws. Apply the results to calculate the effect of spreading in some typical microscopic and macroscopic experiments.

2. Express the spreading Gaussian wave function $\psi(x, t)$ obtained in Problem 1 in the form $\psi(x, t) = \exp[i\, S(x, t)/\hbar]$. Identify the function $S(x, t)$ and show that it satisfies the quantum mechanical Hamilton-Jacobi equation.

3. Consider a wave function that initially is the superposition of two well-separated narrow wave packets:

$$\psi_1(x, 0) + \psi_2(x, 0)$$

chosen so that the absolute value of the overlap integral

$$\gamma(0) = \int_{-\infty}^{+\infty} \psi_1^*(x)\psi_2(x)\, dx$$

is very small. As time evolves, the wave packets move and spread. Will $|\gamma(t)|$ increase in time, as the wave packets overlap? Justify your answer.

4. A high-resolution neutron interferometer narrows the energy spread of thermal neutrons of 20 meV kinetic energy to a wavelength dispersion level of $\Delta\lambda/\lambda = 10^{-9}$. Estimate the length of the wave packets in the direction of motion. Over what length of time will the wave packets spread appreciably?

CHAPTER 3

The Schrödinger Equation, the Wave Function, and Operator Algebra

In this chapter we study the properties and behavior of the Schrödinger wave function as a function of space and time, identifying it as a probability amplitude. As the main tool for extracting information from the wave function, Hermitian operators and their algebraic properties are characterized. The advanced reader can regard this chapter and the next as a refresher of the concepts and principles of wave mechanics.[1]

1. ***The Wave Equation and the Interpretation of ψ.*** A straightforward generalization of the free particle wave equation (2.35) to the case of motion of a particle of mass m in the field of force represented by a potential energy function $V(x, y, z, t)$, which is dependent on the position $\mathbf{r}$ and possibly also on the time t, is the equation

$$\boxed{i\hbar \frac{\partial \psi(\mathbf{r}, t)}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{r}, t) + V(x, y, z, t)\psi(\mathbf{r}, t)} \tag{3.1}$$

Schrödinger advanced this equation on the grounds that the same reasoning that led from (2.35) to (2.39) for a constant potential—setting $\psi = e^{iS/\hbar}$—now implies the equation

$$\frac{\partial S}{\partial t} + \frac{[\nabla S]^2}{2m} - \frac{i\hbar}{2m} \nabla^2 S + V(x, y, z, t) = 0 \tag{3.2}$$

This looks like the Hamilton-Jacobi equation of classical mechanics in the presence of forces, supplemented again by a quantum mechanical term proportional to $\hbar$. The existence of the factor i in this equation is of crucial importance and generally requires that the function S be complex-valued. More will be said in Chapter 15 about the relation between wave mechanics and classical Hamiltonian mechanics.

We will adopt (3.1) as the fundamental equation of nonrelativistic quantum mechanics for particles without spin and call it the *wave equation* or the *time-dependent Schrödinger equation.* In subsequent chapters, this theory will be generalized to encompass many different and more complicated systems, but wave mechanics, which is the branch of quantum mechanics with a dynamical law in the explicit form (3.1), remains one of the simplest paradigms of the general theory.

The ultimate justification for choosing (3.1) must, of course, come from agreement between predictions and experiment. Hence, we must examine the properties of this equation. Before going into mathematical details, it would seem wise, however, to attempt to say precisely what ψ is. We are in the paradoxical situation of having obtained an equation that we believe is satisfied by this quantity, but of

[1]An excellent modern textbook on wave mechanics is Bialynicki-Birula, Cieplak, and Kaminski (1992). It includes an extensive bibliography with many references to the original quantum mechanics literature.

having so far given only a deliberately vague interpretation of its physical significance. We have regarded the wave function as a "measure of the probability" of finding the particle at time t at the position $\mathbf{r}$. How can this statement be made precise?

Obviously, the quantity ψ itself cannot be a probability. All hopes we might have entertained in that direction vanished when ψ became a complex function, since probabilities are real and positive. In the face of this dilemma, as was already explained in Chapter 1, the next best guess is that the probability is proportional to $|\psi|^2$, the square of the amplitude of the wave function.

Of course, we were careless when we used the phrase "probability of finding the particle at position $\mathbf{r}$." Actually, all we can speak of is the probability that the particle is in a volume element d^3r which contains the point $\mathbf{r}$. Hence, we now try the interpretation that $|\psi|^2$ is proportional to the probability that upon measurement of its position the particle will be found in the given volume element. The probability of finding the particle in some finite region of space is then proportional to the integral of $\psi^*\psi$ over this region.

The consistency of this probabilistic interpretation of the wave function requires that if the probability of finding the particle in some bounded region of space *decreases* as time goes on, then the probability of finding it outside of this region must *increase* by the same amount. The probability interpretation of the ψ waves can be made consistently only if this *conservation of probability* is guaranteed. This requirement is fulfilled, owing to Gauss' integral theorem, if it is possible to define a *probability current density* $\mathbf{j}$, which together with the *probability density* $\rho = \psi^*\psi$ satisfies a continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0 \tag{3.3}$$

exactly as in the case of the conservation of matter in hydrodynamics or the conservation of charge in electrodynamics.

A relation of the form (3.3) can easily be deduced from the wave equation for real-valued V. Multiply (3.1) on the left by ψ^*, and the complex conjugate of (3.1) by ψ on the right; subtract the two equations and make a simple transformation. The resulting equation

$$\frac{\partial(\psi^*\psi)}{\partial t} + \frac{\hbar}{2mi} \nabla \cdot [\psi^*\nabla\psi - (\nabla\psi^*)\psi] = 0 \tag{3.4}$$

has the form of (3.3) if the identification

$$\rho = C\psi^*\psi \tag{3.5}$$

$$\mathbf{j} = C\frac{\hbar}{2mi}[\psi^*\nabla\psi - (\nabla\psi^*)\psi] \tag{3.6}$$

is made. Here C is a constant.

Exercise 3.1. Derive (3.4). Note that it depends on V being real.

Exercise 3.2. Generalizing (3.4), prove that, again for real V,

$$\frac{\partial(\psi_1^*\psi_2)}{\partial t} + \frac{\hbar}{2mi} \nabla \cdot [\psi_1^*\nabla\psi_2 - (\nabla\psi_1^*)\psi_2] = 0 \tag{3.7}$$

if ψ_1 and ψ_2 are two solutions of (3.1).

We have emphasized that V must be real-valued if conservation of probability is to hold. Complex-valued potentials are used to simulate the interaction of a particle (e.g., a nucleon) with a complicated many-body system (nucleus), without having to give a detailed account of the dynamic process. The imaginary part of such an *optical* potential serves as a shorthand description of the absorption of the particle by the many-body system, but in principle the particle's motion can be tracked even after it is "absorbed." The true creation or annihilation of a particle, as in nuclear beta decay, requires generalization of the quantum mechanical formalism (see Chapter 21).

Exercise 3.3. Show that if $S(\mathbf{r}, t)$ is defined by $\psi = e^{iS/\hbar}$ as in (2.36), the probability current density can be expressed as

$$\mathbf{j} = \rho \nabla(\text{Re } S)/m \tag{3.8}$$

Exercise 3.4. Assume that the function $\phi(\mathbf{k})$ in expression (2.18) is real and symmetric with respect to $\bar{\mathbf{k}}$. Show that

$$\psi(\mathbf{r}, 0) = e^{i\bar{\mathbf{k}}\cdot\mathbf{r}}\psi_0(\mathbf{r})$$

where $\psi_0(\mathbf{r})$ is a real amplitude modulating the plane wave. Calculate ρ and $\mathbf{j}$ for a wave packet of this form and show that for free particle motion in the approximation (2.22) the relation

$$\mathbf{j} = \frac{\bar{\mathbf{p}}}{m}\rho = \bar{\mathbf{v}}\rho$$

is obtained. Is this a reasonable result?

So far, we have only assumed that ρ and $\mathbf{j}$ are proportional to ψ and ψ^*. Since (3.1) determines ψ only to within a multiplicative constant, we may set $C = 1$.

$$\boxed{\rho = \psi^*\psi} \tag{3.5a}$$

$$\boxed{\mathbf{j} = \frac{\hbar}{2mi}[\psi^*\nabla\psi - (\nabla\psi^*)\psi]} \tag{3.6a}$$

and normalize the wave function by requiring

$$\boxed{\int_{\text{all space}} \rho\, d^3r = \int_{\text{all space}} \psi^*\psi\, d^3r = 1} \tag{3.9}$$

provided that the integral of $\psi^*\psi$ over all space exists (or is, in mathematical terminology, an element of the space L^2). From here on, all spatial integrations will be understood to extend over *all space*, unless otherwise stated, and the limits of integration will usually be omitted. Equation (3.9) expresses the simple fact that the probability of finding the particle anywhere in space at all is unity. Whenever this integral exists, we will assume that this normalization has been imposed. Functions ψ for which $\int \psi^*\psi\, d^3r$ exists are often called *quadratically integrable*.

If the integral (3.9) exists, and if the probability current density falls to zero faster than r^{-2} as r becomes very large, then Gauss' theorem applied to (3.3) gives

$$\frac{\partial}{\partial t}\int \rho \, d^3r = 0 \tag{3.10}$$

which simply tells us that the wave equation guarantees the conservation of normalization: If ψ was normalized at $t = 0$, it will remain normalized at all times.

In light of the foregoing remarks, it is well to emphasize that throughout this book all integrals that are written down will be assumed to exist. In other words, whatever conditions ψ must satisfy for a certain integral to exist will be assumed to be met by the particular wave function used. This understanding will save us laborious repetition. Examples will occur presently.

When the integral (3.9) does *not* exist, we may use only (3.5) and (3.6), and we must then speak of *relative* rather than *absolute* probabilities. Wave functions that are not normalizable to unity in the sense of (3.9) are as important as normalizable ones. The former appear when the particle is unconfined, and in principle they can always be avoided by the use of finite wave packets. However, rigid insistence on dealing only with finite wave packets would prevent us from using such simple wave functions as the infinite plane wave $e^{i\mathbf{k}\cdot\mathbf{r}}$, which describes a particle that is equally likely to be found anywhere in space at all—hence, of course, with zero absolute probability in any finite volume. Owing to the simple properties of Fourier integrals, mathematically this patently unphysical object is extremely convenient. It is an idealization of a finite, quadratically integrable, but very broad wave packet that contains very many wavelengths. If these limitations are recognized, the infinite plane matter wave is no more objectionable than an infinite electromagnetic wave that is also unphysical, because it represents an infinite amount of energy. We shall depend on the use of such wave functions. Alternative methods for normalizing them are discussed in Sections 4.3 and 4.4.

One further comment on terminology: Since $\psi(\mathbf{r}, t)$ is assumed to contain all information about the state of the physical system at time t, the terms *wave function* ψ, and *state* may be used interchangeably. Thus, we may speak without danger of confusion of a *normalized state* if (3.9) holds.

It is tempting to use the definition (3.5a) of the probability density and the definition of the function S to express the wave function as

$$\psi = e^{-\mathrm{Im}S/\hbar}\, e^{i\mathrm{Re}S/\hbar} = \sqrt{\rho}\, e^{i\mathrm{Re}S/\hbar} \tag{3.11}$$

The coupled equations of motion for ρ and the phase function ReS are easily obtained as the real and imaginary parts of (3.2). It follows from (3.8) that one of these is simply the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \left[\rho \frac{\nabla(\mathrm{Re}\, S)}{m}\right] = 0 \tag{3.12}$$

This equation can be interpreted as stating that probability is similar to a fluid that flows with velocity $\mathbf{v} = \nabla(\mathrm{Re}\, S)/m$. This irrotational flow ($\nabla \times \mathbf{v} = 0$) is known as *Madelung* flow.

The second equation of motion is

$$\frac{\partial(\mathrm{Re}\, S)}{\partial t} + \frac{[\nabla(\mathrm{Re}\, S)]^2}{2m} + V + \frac{\hbar^2}{4m}\left[\frac{(\nabla\rho)^2}{2\rho^2} - \frac{\nabla^2\rho}{\rho}\right] = 0 \tag{3.13}$$

or

$$\frac{\partial(\text{Re } S)}{\partial t} + \frac{\{\nabla(\text{Re } S)\}^2}{2m} + V - \frac{\hbar^2}{2m}\frac{\nabla^2\sqrt{\rho}}{\sqrt{\rho}} = 0 \tag{3.14}$$

Since everything in Eqs. (3.12) and (3.13) is real-valued, it is not unreasonable to think of Eq. (3.13) or (3.14) as a classical Hamilton-Jacobi equation for the dynamics of a particle, provided that the last term is regarded as a *quantum potential* that must be added to the conventional potential energy V. Taking his cue from these nonlinear equations, which are so classical in appearance, David Bohm and others advocated a more realistic and deterministic, or *ontological*, interpretation of quantum mechanics than the one adopted in the present text. Whatever conceptual appeal Eqs. (3.12) and (3.13) may have is offset by the abstruseness of the quantum potentials and the complexities of the Bohm approach for many-particle and relativistic systems.[2]

Since the interpretation of the wave function is in terms of probabilities, the elements of probability theory are reviewed in Appendix A, Section 2. There it is assumed that the events are discrete, whereas in the application to quantum mechanics continuous probability distributions such as $\rho = |\psi(x, y, z, t)|^2$ are common, as we will see in Section 3.2. The summations in the formulas for expectation (or average or mean) values must then be replaced by integrations. In quantum mechanics the term *expectation value* is preferred when it is desirable to emphasize the predictive nature of the theory and the fact that the behavior of a single particle or physical system is involved. Although the wave function must be regarded as describing an ensemble representing a single system, the predictions for the state can be compared with experimental trials only by the use of a large number of identical and identically prepared systems. The situation is analogous to the one encountered customarily in the use of probability concepts. For instance, the probability of tossing heads with a slightly unsymmetrical coin might be 0.48 and that of tossing tails 0.52. These probabilities refer to the particular coin and to the individual tossings, but the experimental verification of the statement "the probability for tossing heads is 48 percent" requires that we toss many identical coins, or the same coin many times in succession, and 48 percent of a very large sequence of trials are expected to yield "heads."

2. ***Probabilities in Coordinate and Momentum Space.*** For a state normalized according to condition (3.9), the average or expectation value of the coordinate x, which is a random variable, is

$$\langle x \rangle = \int x\,|\psi|^2\,d^3r \tag{3.15}$$

The expectation value of the position vector $\mathbf{r}$, or the center of the wave packet, is defined by

$$\boxed{\langle \mathbf{r} \rangle = \int \mathbf{r}\,|\psi|^2\,d^3r} \tag{3.16}$$

[2]For a fuller and more sympathetic assessment of David Bohm's approach to quantum mechanics, see Cushing (1994).

An arbitrary function of **r** has the expectation value

$$\langle f(\mathbf{r})\rangle = \int f(\mathbf{r})\,|\psi|^2\,d^3r = \int \psi^* f(\mathbf{r})\psi\,d^3r \tag{3.17}$$

The reason for writing $\langle f\rangle$ in the clumsy form $\int \psi^* f\psi\,d^3r$ will become clear when less elementary expectation values make their appearance.

At this point, it is not obvious how to set up the calculation of expectation values of momentum or kinetic energy or any other physical quantity that is not expressed simply in terms of the position coordinates. Since we are assuming that the physical state of a particle at time t is described as fully as possible by the normalized wave function $\psi(\mathbf{r}, t)$, the discussion in Chapter 2 suggests that knowledge about the momentum of the particle is contained in the Fourier transform $\phi(\mathbf{p}, t)$, which is related to $\psi(\mathbf{r}, t)$ by the reciprocal Fourier transforms (see Appendix, Section 1):

$$\psi(\mathbf{r}, t) = \frac{1}{(2\pi\hbar)^{3/2}} \int \phi(\mathbf{p}, t)\exp\left(\frac{i}{\hbar}\,\mathbf{p}\cdot\mathbf{r}\right)d^3p \tag{3.18}$$

and

$$\phi(\mathbf{p}, t) = \frac{1}{(2\pi\hbar)^{3/2}} \int \psi(\mathbf{r}, t)\exp\left(-\frac{i}{\hbar}\,\mathbf{p}\cdot\mathbf{r}\right)d^3r \tag{3.19}$$

In these equations we have used **p** rather than $\mathbf{k} = \mathbf{p}/\hbar$ as the conjugate Fourier variable, but the analogous relations (2.16), (2.17), and (2,18) in **k**-space are recovered by the substitutions

$$\mathbf{p} \to \hbar\mathbf{k}, \qquad d^3p \to \hbar^3 d^3k, \qquad \phi(\mathbf{p}) \to \hbar^{-3/2}\phi(\mathbf{k})$$

Rather than introducing two different symbols for the functions describing the state in momentum space (**p**) and in **k**-space, we rely on the explicit notation $\phi(\mathbf{p})$ and $\phi(\mathbf{k})$ to remind us of the difference between them.

We emphasize once more that the functions ψ in coordinate space (or position space) and ϕ in momentum space are both equally valid descriptions of the state of the system. Given either one of them, the other can be calculated from (3.18) or (3.19). Both ψ and ϕ depend on the time t, which is a parameter. For the special case of a *free* particle we know from Chapter 2 how a plane wave of momentum $\mathbf{p} = \hbar\mathbf{k}$ develops in time and that

$$\phi(\mathbf{p}, t) = \phi(\mathbf{p}, 0)e^{-i\omega(\mathbf{p})t} \tag{3.20}$$

[In Eqs. (2.16), (2.17), and (2.18), $\phi(\mathbf{k}, 0)$ was denoted simply as $\phi(\mathbf{k})$.]

For a general motion in a potential, the time development of $\phi(\mathbf{p}, t)$ is more complicated than (3.20) and must be worked out from the wave equation or its partner in momentum space. The latter equation is obtained from the wave equation in coordinate space (3.1), most easily by differentiating (3.19) with respect to t:

$$i\hbar\,\frac{\partial\phi(\mathbf{p}, t)}{\partial t} = \frac{1}{(2\pi\hbar)^{3/2}} \int i\hbar\,\frac{\partial\psi(\mathbf{r}, t)}{\partial t}\exp\left(-\frac{i}{\hbar}\,\mathbf{p}\cdot\mathbf{r}\right)d^3r$$

or

$$i\hbar\,\frac{\partial\phi(\mathbf{p}, t)}{\partial t} = \frac{1}{(2\pi\hbar)^{3/2}} \int \exp\left(-\frac{i}{\hbar}\,\mathbf{p}\cdot\mathbf{r}\right)\left[-\frac{\hbar^2}{2m}\,\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r}, t)\,d^3r$$

The first term of this last expression can be transformed by integrating by parts twice and assuming that ψ is subject to boundary conditions, vanishing sufficiently fast at large distances, so that the surface terms can be neglected. If the Fourier transforms (3.18) and (3.19) are then used again, after a little rearranging we obtain

$$\boxed{\begin{aligned} i\hbar \frac{\partial \phi(\mathbf{p}, t)}{\partial t} &= \frac{p^2}{2m} \phi(\mathbf{p}, t) \\ &+ \frac{1}{(2\pi\hbar)^{3/2}} \int \exp\left(-\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{r}\right) V(\mathbf{r}) \exp\left(\frac{i}{\hbar} \mathbf{p}' \cdot \mathbf{r}\right) \phi(\mathbf{p}', t) d^3r \, d^3p' \end{aligned}} \tag{3.21}$$

This mixed integral-differential equation for ϕ is the wave equation in momentum space, and it is fully equivalent to (3.1). It can be transformed into a differential equation if V is an analytic function of x, y, z, and if we can write

$$\exp\left(-\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{r}\right) V(\mathbf{r}) = V(i\hbar \nabla_{\mathbf{p}}) \exp\left(-\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{r}\right) \tag{3.22}$$

Substituting this into (3.21) and removing the differential operator with respect to $\mathbf{p}$ from the integrals, we can now perform the $\mathbf{r}$-integration, and we are left with a delta function. This in turn makes the $\mathbf{p}$-integration a trivial operation, and the result is a partial differential equation in momentum space:

$$\boxed{i\hbar \frac{\partial \phi(\mathbf{p}, t)}{\partial t} = \frac{p^2}{2m} \phi(\mathbf{p}, t) + V(i\hbar \nabla_{\mathbf{p}}) \phi(\mathbf{p}, t)} \tag{3.23}$$

Exercise 3.5. Work through the steps leading to (3.23), using the Appendix, Section 1.

If we multiply (3.21) by ϕ^*, and the complex conjugate of (3.21) by ϕ, subtract the two equations, and integrate the resulting equation over the entire momentum space ($\mathbf{p}$), we obtain, after renaming some variables of the integration,

$$\frac{\partial}{\partial t} \int \phi^*(\mathbf{p}, t) \phi(\mathbf{p}, t) d^3p = 0 \tag{3.24}$$

This result should not be surprising because it follows directly from the Fourier transforms (3.18) and (3.19) that

$$\int |\phi(\mathbf{p}, t)|^2 \, d^3p = \int |\psi(\mathbf{r}, t)|^2 \, d^3r \tag{3.25}$$

This equation implies that if ψ is normalized to unity for any t, ϕ is automatically also so normalized. Equation (3.10) or (3.24) then guarantees that this normalization is preserved at all times.

Since $\exp\left(\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{r}\right)$ represents a state with definite sharp momentum $\mathbf{p}$, $\phi(\mathbf{p})$ according to (3.18) is the amplitude with which the momentum $\mathbf{p}$ is represented in the wave function $\psi(\mathbf{r})$. If ϕ is strongly peaked near a particular value of $\mathbf{p}$, the state tends to be one of rather definite momentum and similar to a plane wave. This observation, in conjunction with the conservation property (3.24), leads us to the assumption that $|\phi(\mathbf{p}, t)|^2 \, d^3p$ is the *probability of finding the momentum of the*

particle in the volume element d^3p in the neighborhood of $\mathbf{p}$ at time t, and that $|\phi(\mathbf{p}, t)|^2$ is the *probability density in momentum space*.

The symmetry of $\mathbf{r}$ and $\mathbf{p}$ which the present section exhibits is also evident in classical mechanics, especially in the Hamiltonian form, where (generalized) coordinates q and momenta p stand in a similar reciprocal relationship. However, in classical analytical dynamics the symmetry is primarily a formal one, for q and p together determine the state at time t, and classical physics tacitly assumes that they can both be determined simultaneously and independently with perfect precision. In quantum mechanics, the statistical properties of these two physical quantities are determined by the functions ψ and ϕ, respectively. Since these functions are one another's Fourier transforms, q and p are correlated and can no longer be measured or chosen entirely independently. In general, the covariance (Appendix, Section 2) of the random variables representing position and momentum does not vanish. The probability distributions $|\psi(\mathbf{r}, t)|^2$ and $|\phi(\mathbf{p}, t)|^2$ are subject to the Heisenberg uncertainty relation, a precise statement and mathematical proof of which will be given in Section 10.5.

Wave mechanics in coordinate space (ψ) and wave mechanics in momentum space (ϕ) are thus seen to be two equivalent descriptions of the same thing. As Chapters 9 and 10 will show, it is advantageous to regard both of these descriptions as two special representations of a general and abstract formulation of quantum mechanics. The terms *coordinate representation* and *momentum representation* will then be seen to be apt for referring to these two forms of the theory, but among the infinitely many possible representations no special status will be accorded them. Although most of the equations in wave mechanics will be written explicitly in the coordinate representation, any equation can be transcribed into the momentum representation at any desired state of a calculation.

The probability interpretation of the momentum wave function ϕ allows us to construct the expectation value of the momentum for any state:

$$\boxed{\langle \mathbf{p} \rangle = \int \mathbf{p} |\phi(\mathbf{p}, t)|^2 \, d^3p} \tag{3.26}$$

In order to derive an expression for this expectation value in terms of ψ, we substitute the complex conjugate of (3.19) into (3.26) and write

$$\begin{aligned} \langle \mathbf{p} \rangle &= \int \int \psi^*(\mathbf{r}, t) \exp\left(\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{r}\right) \mathbf{p} \, \phi(\mathbf{p}, t) \, d^3r \, d^3p \\ &= \int \int \psi^*(\mathbf{r}, t) \frac{\hbar}{i} \nabla \exp\left(\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{r}\right) \phi(\mathbf{p}, t) \, d^3r \, d^3p \end{aligned}$$

After liberal interchange of differentiations and integrations, and use of (3.18), the desired result

$$\boxed{\langle \mathbf{p} \rangle = \int \psi^*(\mathbf{r}, t) \frac{\hbar}{i} \nabla \psi(\mathbf{r}, t) \, d^3r} \tag{3.27}$$

is obtained. More generally, the same technique gives for any (analytic) function of momentum, $g(\mathbf{p})$,

$$\langle g(\mathbf{p}) \rangle = \int g(\mathbf{p}) \, |\phi(\mathbf{p}, t)|^2 \, d^3p = \int \psi^*(\mathbf{r}, t) g\left(\frac{\hbar}{i} \nabla\right) \psi(\mathbf{r}, t) \, d^3r \tag{3.28}$$

where products of ∇'s or $\partial/\partial x$'s are to be understood as successive operations of differentiation of the function on the right, on which they act. Conversely, the expectation value of any (analytic) function $f(\mathbf{r})$ of position is given by

$$\langle f(\mathbf{r})\rangle = \int f(\mathbf{r})\,|\psi(\mathbf{r}, t)|^2\,d^3r = \int \phi^*(\mathbf{p}, t) f(i\hbar\nabla_{\mathbf{p}})\phi(\mathbf{p}, t)\,d^3p \tag{3.29}$$

If the physical system under consideration can be described in the language of classical mechanics, all quantities associated with it, such as kinetic and potential energy, angular momentum, and the virial, can be expressed in terms of coordinates *and* momenta. The question that then arises is how to compute the expectation value of an arbitrary function $F(\mathbf{r}, \mathbf{p})$. The expression for this expectation value can be derived to a certain extent from our previous work.

For example, if we wish to calculate the expectation value of a mixed function like $f(x, p_y, z)$, we carry out the Fourier analysis only part of the way and define a mixed coordinate-momentum wave function by the integral

$$\chi(x, p_y, z) = \frac{1}{(2\pi)^{1/2}} \int \psi(\mathbf{r}, t)\exp\left(-\frac{i}{\hbar}p_y y\right)dy \tag{3.30}$$

The square of the absolute value of this function, $|\chi(x, p_y, z)|^2$, can easily be seen to be the probability density of finding the particle to have coordinates x and z, but indeterminate y, whereas the y-component of its momentum is equal to p_y. From this, it follows readily that the expectation value of any function of x, p_y, z is

$$\langle f(x, p_y, z)\rangle = \int f(x, p_y, z)\,|\chi(x, p_y, z)|^2\,dx\,dp_y\,dz$$

Similar considerations hold for other combinations of components of $\mathbf{r}$ and $\mathbf{p}$, provided that the simultaneous specification of conjugate variables, such as x and p_x, is avoided. As a result, we may conclude that

$$\boxed{\langle F(\mathbf{r}, \mathbf{p})\rangle = \begin{cases} \displaystyle\int \psi^*(\mathbf{r}, t) F\left(\mathbf{r}, \frac{\hbar}{i}\nabla\right)\psi(\mathbf{r}, t)\,d^3r \\ \displaystyle\int \phi^*(\mathbf{p}, t) F(i\hbar\nabla_{\mathbf{p}}, \mathbf{p})\phi(\mathbf{p}, t)\,d^3p \end{cases}} \tag{3.31}$$

However, we have not proved (3.31) generally, because no account has been taken of as simple a function of coordinates and momenta as $F = \mathbf{r}\cdot\mathbf{p} = xp_x + yp_y + zp_z$ (which will be encountered in Section 3.6). The trouble here is that our analysis can never give us a wave function the square of whose absolute value would represent the probability density of finding the particle with coordinate x *and* momentum component p_x. Heisenberg's uncertainty principle is an articulation of this inability. If products like xp_x do appear, we postulate that (3.31) holds, but we must keep in mind that classically the product xp_x can be equally well written as $p_x x$, whereas the quantum mechanical expectation values of xp_x and $p_x x$ are generally different (and, in fact, not even real-valued).

The equality of the two expressions on the right side of Eq. (3.31) for any pair of coordinate and momentum wave functions is equivalent to the following theorem:

If $\psi(\mathbf{r})$ and $\phi(\mathbf{p})$ are Fourier transforms of each other in the sense of (3.18) and (3.19), then for any function $F(\mathbf{r}, \mathbf{p})$, the coordinate function $F\left(\mathbf{r}, \frac{\hbar}{i}\nabla\right)\psi(\mathbf{r})$ and

the momentum function $F(i\hbar\nabla_{\mathbf{p}}, \mathbf{p})\phi(\mathbf{p})$ are Fourier transforms of each other in the same sense, that is,

$$F\left(\mathbf{r}, \frac{\hbar}{i}\nabla\right)\psi(\mathbf{r}) = \frac{1}{(2\pi\hbar)^{3/2}}\int e^{i/\hbar\mathbf{p}\cdot\mathbf{r}}\, F(i\hbar\nabla_{\mathbf{p}}, \mathbf{p})\phi(\mathbf{p})\, d^3p \tag{3.32}$$

and

$$F(i\hbar\nabla_{\mathbf{p}}, \mathbf{p})\phi(\mathbf{p}) = \frac{1}{(2\pi\hbar)^{3/2}}\int e^{-i/\hbar\mathbf{p}\cdot\mathbf{r}}\, F\left(\mathbf{r}, \frac{\hbar}{i}\nabla\right)\psi(\mathbf{r})\, d^3r \tag{3.33}$$

These formulas hold if ψ and ϕ go to zero sufficiently fast for large values of their arguments, so that surface terms, resulting from integration by parts, vanish.

When $F\left(\mathbf{r}, \frac{\hbar}{i}\nabla\right)$ acts on $\psi(\mathbf{r})$, a new function of the coordinates results, and similarly $F(i\hbar\nabla_{\mathbf{p}}, \mathbf{p})$ turns $\phi(\mathbf{p})$ into a new function of momentum. Thus, F is an example of an operator that maps one function into another.

Exercise 3.6. Derive the equality of the two expressions on the right side of (3.31) from the transforms (3.32) and (3.33).

Exercise 3.7. Prove that if $\psi_1(\mathbf{r})$ and $\psi_2(\mathbf{r})$ are Fourier transforms of $\phi_1(\mathbf{p})$ and $\phi_2(\mathbf{p})$, respectively, we have the identity

$$\int \psi_1^*(\mathbf{r})\psi_2(\mathbf{r})d^3r = \int \phi_1^*(\mathbf{p})\phi_2(\mathbf{p})\, d^3p \tag{3.34}$$

which is a generalization of (3.25).

Exercise 3.8. Show that for a linear harmonic oscillator with $V = m\omega^2x^2/2$ the wave equations in coordinate and momentum space have the same structure and that every normalized solution $\psi_1(x, t)$ of the wave equation in coordinate space can be related to a solution

$$\phi_2(p_x, t) = \frac{e^{i\alpha}}{\sqrt{m\omega}}\,\psi_1\left(\frac{p_x}{m\omega}, t\right) \tag{3.35}$$

of the wave equation in momentum space. [The different subscripts are meant as reminders that, generally, ψ_1 and ϕ_2 in (3.35) are not Fourier transforms of each other and do not represent the same state.]

3. ***Operators and Expectation Values of Dynamical Variables.*** An operator is a rule that maps every function ψ (or ϕ) into a function $F\psi$ (or $F\phi$). An operator is said to be *linear* if its action on any two functions ψ_1 and ψ_2 is such that

$$F(\lambda\psi_1 + \mu\psi_2) = \lambda F\psi_1 + \mu F\psi_2 \tag{3.36}$$

where λ and μ are arbitrary complex numbers. Derivatives are obviously linear operators, as are mere multipliers. Most of the operators that are relevant to quantum mechanics are linear. Unless it is specifically stated that a given operator is not linear, the term *operator* will henceforth be reserved for linear operators.

The only other category of operators important in quantum mechanics is the *antilinear* variety, characterized by the property

$$F(\lambda\psi_1 + \mu\psi_2) = \lambda^* F\psi_1 + \mu^* F\psi_2 \tag{3.37}$$

Complex conjugation itself is an example of an antilinear operator. Antilinear operators are important in quantum mechanics when time-reversed physical processes are considered.

Exercise 3.9. Construct some examples of linear and antilinear operators and of some operators that are neither linear nor antilinear.

From Eq. (3.31) we see that with any physical quantity, which is a function of coordinates and linear momenta, there is associated a linear operator which, when interposed between the wave function and its complex conjugate, upon integration gives the expectation value of that physical quantity. The explicit form of the operator evidently depends on whether the ψ or ϕ function is used to represent the state. The expectation value of the physical quantity F is given by the formula

$$\boxed{\langle F \rangle = \int \psi^* F \psi \, d^3r = \int \phi^* F \phi \, d^3p} \tag{3.38}$$

The student of quantum mechanics must get used to the potentially confusing practice of frequently denoting the mathematical operators, even in different representations, by the same symbols (e.g., F above) as the physical quantities that they represent. Yet, as in most matters of notation, it is not difficult to adapt to these conventions; the context will always establish the meaning of the symbols unambiguously.

Table 3.1 presents some of the most important physical quantities or dynamical variables, and the operators that represent them.

Table 3.1

Physical Quantity		*Operator in Coordinate Representation*	*Operator in Momentum Representation*
Position	$\mathbf{r}$	$\mathbf{r}$	$i\hbar\nabla_{\mathbf{p}}$
Momentum	$\mathbf{p}$	$\frac{\hbar}{i}\nabla$	$\mathbf{p}$
Angular momentum	$\mathbf{L}$	$\mathbf{r} \times \frac{\hbar}{i}\nabla$	$i\hbar\nabla_{\mathbf{p}} \times \mathbf{p}$
Kinetic energy	T	$-\frac{\hbar^2}{2m}\nabla^2$	$\frac{p^2}{2m}$
Potential energy	V	$V(\mathbf{r})$	$V(i\hbar\nabla_{\mathbf{p}})$
Total energy	E	$-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})$	$\frac{p^2}{2m} + V(i\hbar\nabla_{\mathbf{p}})$

As the discussion in the last section makes clear, the expectation values of all of these operators are real numbers, as they should be if they are averages of physical quantities that are analogous to the familiar classical quantities of the same name.

The reasonableness of the operator assignments in Table 3.1 is confirmed by showing that their expectation values generally satisfy the laws of classical me-

chanics. For example, by using the continuity equation, the time derivative of the mean x coordinate can be written as

$$\frac{d}{dx}\langle x\rangle = \int x\frac{\partial\rho}{\partial t}\,d^3r = -\int x\nabla\cdot\mathbf{j}\,d^3r = -\int \nabla\cdot(x\,\mathbf{j})d^3r + \int j_x\,d^3r = \int j_x\,d^3r$$

or

$$\boxed{\frac{d}{dt}\langle\mathbf{r}\rangle = \int \mathbf{j}\,d^3r} \tag{3.39}$$

where the divergence term has been removed under the assumption that ψ vanishes sufficiently fast at infinity. Using (3.6a) and integration by parts, we finally obtain

$$m\frac{d}{dt}\langle\mathbf{r}\rangle = \int \psi^*\frac{\hbar}{i}\nabla\psi\,d^3r = \langle\mathbf{p}\rangle \tag{3.40}$$

which shows that for the expectation values of wave packets the usual relation between velocity and momentum holds. We note here that $\mathbf{p}$ is the (kinetic) momentum $m\mathbf{v}$, which is the canonical momentum only if there is no electromagnetic vector potential, $\mathbf{A} = 0$. The appropriate generalization in the case $\mathbf{A} \neq 0$ will be dealt with in Section 4.6.

The time rate of change of $\langle\mathbf{p}\rangle$ is, by use of (3.1) and its complex conjugate, given by

$$\begin{aligned}\frac{d}{dt}\langle p_x\rangle &= -i\hbar\int\frac{\partial\psi^*}{\partial t}\frac{\partial\psi}{\partial x}\,d^3r - i\hbar\int\psi^*\frac{\partial}{\partial x}\left(\frac{\partial\psi}{\partial t}\right)d^3r\\ &= -\frac{\hbar^2}{2m}\int\left[(\nabla^2\psi^*)\frac{\partial\psi}{\partial x} - \psi^*\nabla^2\frac{\partial\psi}{\partial x}\right]d^3r + \int\left[V\psi^*\frac{\partial\psi}{\partial x} - \psi^*\frac{\partial(V\psi)}{\partial x}\right]d^3r\end{aligned}$$

The integral containing Laplacians can be transformed into a surface integral by Green's theorem:[3]

$$\int_V\left[(\nabla^2\psi^*)\frac{\partial\psi}{\partial x} - \psi^*\nabla^2\left(\frac{\partial\psi}{\partial x}\right)\right]d^3r = \int_S\left[\frac{\partial\psi}{\partial x}\nabla\psi^* - \psi^*\nabla\left(\frac{\partial\psi}{\partial x}\right)\right]\cdot d\mathbf{S}$$

It is assumed that the last integral vanishes when taken over a very large surface S. What remains is

$$\frac{d}{dt}\langle p_x\rangle = -\int\psi^*\frac{\partial V}{\partial x}\psi\,d^3r$$

or, more generally,

$$\boxed{\frac{d}{dt}\langle p\rangle = -\int\psi^*(\nabla V)\,\psi\,d^3r = -\langle\nabla V\rangle = \langle\mathbf{F}\rangle} \tag{3.41}$$

[3]Green's theorem for continuously differentiable functions u and v:

$$\int_S(u\nabla\omega - v\nabla u)\cdot d\mathbf{S} = \int_V(u\nabla^2 v - v\nabla^2 u)d^3r$$

where the surface S encloses the volume V.

Equation (3.41), known as *Ehrenfest's theorem*, looks like Newton's second law, written for expectation values, and conforms to the correspondence principle. If the state ψ is such that $\langle \mathbf{F}(\mathbf{r}) \rangle \approx \mathbf{F}(\langle \mathbf{r} \rangle)$, it follows from (3.40) and (3.41) that $\langle \mathbf{r} \rangle$ moves in time very nearly like a classical particle. In particular, this condition holds for a sharply peaked wave packet.

We can establish a general formula for the time derivative of the expectation value $\langle F \rangle$ of any operator F. To this end we write the wave equation (3.1) in the form

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi \tag{3.42}$$

by making use of the notation

$$H = -\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \tag{3.43}$$

for the *Hamiltonian* (or total energy) operator, H. The time derivative of the expectation value of any operator F, which may be explicitly time dependent, can be written as

$$\begin{aligned} i\hbar \frac{d}{dt} \langle F \rangle &= i\hbar \int \psi^* F \frac{\partial \psi}{\partial t} d^3r + i\hbar \int \frac{\partial \psi^*}{\partial t} F\psi \, d^3r + i\hbar \int \psi^* \frac{\partial F}{\partial t} \psi \, d^3r \\ &= \int \psi^* FH\psi d^3r - \int (H\psi)^* F\psi d^3r + i\hbar \left\langle \frac{\partial F}{\partial t} \right\rangle \\ &= \int \psi^* F\left(-\frac{\hbar^2}{2m} \nabla^2 + V \right)\psi \, d^3r - \int \psi^* \left(-\frac{\hbar^2}{2m} \nabla^2 + V \right) F\psi \, d^3r + i\hbar \left\langle \frac{\partial F}{\partial t} \right\rangle \end{aligned}$$

In the last step, use has again been made of Green's theorem, and vanishing boundary surface terms have been omitted. The potential energy is, of course, assumed to be real. Compactly, we may write the last result as

$$\boxed{i\hbar \frac{d}{dt} \langle F \rangle = \langle FH - HF \rangle + i\hbar \left\langle \frac{\partial F}{\partial t} \right\rangle} \tag{3.44}$$

This formula is of the utmost importance in all facets of quantum mechanics.

Exercise 3.10. Show that Eq. (3.44) is also true in the momentum representation.

The order in which two operators F and H, both acting to their right, are applied to a state ψ (or ϕ) is generally important, and their *commutator*,

$$\boxed{[F, H] = FH - HF} \tag{3.45}$$

is a measure of the difference between the two products. If

$$[F, H] = 0 \tag{3.46}$$

the two operators are said to *commute*.

According to Eq. (3.44), if F commutes with H and does not explicitly depend on t, $\langle F \rangle$ is constant in time for any state. The physical quantity F is then termed a *constant of the motion* and is said to be *conserved*.

Since H commutes with itself, $\langle H \rangle$ is constant in time, expressing the quantum version of the law of conservation of energy for a system with a potential that does not explicitly depend on time (conservative forces).

Linear momentum $\mathbf{p} = m\mathbf{v}$ is a constant of the motion if $\mathbf{p}$ commutes with H. Since $\mathbf{p}$ commutes with the kinetic energy operator $p^2/2m$, it is only necessary to examine whether it also commutes with $V(x, y, z)$. All three components of momentum commute with V only if V is constant. Hence, in quantum mechanics as in classical mechanics, linear momentum is conserved only if no external forces act upon the system.

Exercise 3.11. Prove that the operator p_x commutes with a function $V(x)$ of x only if V is constant, and generalize this result to three dimensions.

4. ***Commutators and Operator Algebra.*** The commutator $[A, B] = AB - BA$ of two operators plays a prominent role in quantum mechanics, as exemplified in the last section. The most important commutators are those involving canonically conjugate variables, such as the pairs x and p_x, y and p_y, z and p_z. The manipulation of commutation relations for various *dynamical variables*, as Dirac called the operators that represent physical quantities, is most safely accomplished by allowing the operators to act on an arbitrary function, f, which is removed at the end of the calculation. Thus, in the coordinate representation,

$$(xp_x - p_x x)f = \left(x\frac{\hbar}{i}\frac{\partial}{\partial x} - \frac{\hbar}{i}\frac{\partial}{\partial x}x\right)f = -i\hbar\left[x\frac{\partial f}{\partial x} - \frac{\partial (xf)}{\partial x}\right] = i\hbar f$$

from which we infer the fundamental commutation relation

$$[x, p_x] = xp_x - p_x x = i\hbar\, 1 \tag{3.47}$$

The identity operator 1 may be omitted and the fundamental commutation relations for this and the other canonically conjugate variables may be written as

$$xp_x - p_x x = yp_y - p_y y = zp_z - p_z z = i\hbar \tag{3.48}$$

or, in commutator bracket notation,

$$\boxed{[x, p_x] = [y, p_y] = [z, p_z] = i\hbar\, 1} \tag{3.49}$$

All other products formed from Cartesian coordinates and their conjugate momenta are commutative, i.e., $xy = yz$, $xp_y = p_y x$, $xp_z = p_z x$, and so on.

In commutator bracket notation, for any operators the following elementary rules are easy to verify:

$$\boxed{\begin{aligned}
&[A, B] + [B, A] = 0 \\
&[A, A] = 0 \\
&[A, B + C] = [A, B] + [A, C] \\
&[A + B, C] = [A, C] + [B, C] \\
&[AB, C] = A[B, C] + [A, C]B \\
&[A, BC] = [A, B]C + B[A, C] \\
&[A, [B, C]] + [C, [A, B]] + [B, [C, A]] = 0
\end{aligned}} \tag{3.50}$$

Exercise 3.12. If A and B are two operators that both commute with their commutator, $[A, B]$, prove that, for a positive integer n,

$$\begin{aligned}[A, B^n] &= nB^{n-1}[A, B] \\ [A^n, B] &= nA^{n-1}[A, B]\end{aligned} \tag{3.51}$$

Note the similarity of the process with differentiation. Apply to the special case $A = x$, $B = p_x$. Prove that

$$[p_x, f(x)] = \frac{\hbar}{i}\frac{df}{dx}\,1 \tag{3.52}$$

if $f(x)$ can be expanded in a power series of the operator x.

Exercise 3.13. Show that each component of orbital angular momentum commutes with the kinetic energy operator, e.g.,

$$[L_x, T] = \left[yp_z - zp_y, \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2)\right] = 0$$

Exercise 3.14. Prove that

$$\frac{d}{dt}\langle \mathbf{L} \rangle = -\langle \mathbf{r} \times \nabla V \rangle = \langle \mathbf{r} \times \mathbf{F} \rangle \tag{3.53}$$

as expected from correspondence with classical mechanics.

Another operator relation that is frequently used involves the function e^A, defined by the power series

$$e^A = 1 + \frac{A}{1!} + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$$

Consider the function

$$f(\lambda) = e^{\lambda A} B e^{-\lambda A}$$

of a real number λ, and make a Taylor series expansion of $f(\lambda)$, observing that

$$\begin{aligned}\frac{df}{d\lambda} &= Af(\lambda) - f(\lambda)A = [A, f(\lambda)] \\ \frac{d^2 f}{d\lambda^2} &= \left[A, \frac{df}{d\lambda}\right] = [A, [A, f(\lambda)]]\end{aligned}$$

and so on. Since $f(0) = B$, we get the identify

$$\boxed{e^{\lambda A} B e^{-\lambda A} = B + \frac{\lambda}{1!}[A, B] + \frac{\lambda^2}{2!}[A, [A, B]] + \frac{\lambda^3}{3!}[A, [A, [A, B]]] + \cdots} \tag{3.54}$$

If the operators satisfy the relation (β: constant)

$$[A,[A, B]] = \beta B \tag{3.55}$$

the right-hand side of (3.54) can be summed to give

$$\boxed{e^{\lambda A}Be^{-\lambda A} = B\cosh\lambda\sqrt{\beta} + \frac{[A, B]}{\sqrt{\beta}}\sinh\lambda\sqrt{\beta}} \tag{3.56}$$

A special case arises if (μ: constant)

$$[A, B] = \mu\, 1 \tag{3.57}$$

leading from (3.56) to

$$e^{\lambda A}Be^{-\lambda A} = B + \lambda\mu\, 1 \tag{3.58}$$

Exercise 3.15. If $[A, B] = \gamma B$ (γ: constant) show that

$$e^{\lambda A}Be^{-\lambda A} = e^{\lambda\gamma}B \tag{3.59}$$

Equations like (3.54) caution us against the indiscriminate use of well-known mathematical rules, when the variables are noncommuting operators. In particular, the simple law $e^Ae^B = e^{A+B}$ is generally not valid for operators. Its place is taken by the *Campbell-Baker-Hausdorff* formula, according to which in the equation

$$e^Ae^B = e^{A+B+F(A,B)}$$

the additional term $F(A, B)$ is expressible as a generally infinite sum of multiple commutators of A and B.

Exercise 3.16. Consider the product $G(\lambda) = e^{\lambda A}e^{\lambda B}$ and prove that

$$\begin{aligned}\frac{dG}{d\lambda} &= \left\{A + B + \frac{\lambda}{1!}[A, B] + \frac{\lambda^2}{2!}[A, [A, B]] + \cdots\right\}G \\ &= G\left\{A + B + \frac{\lambda}{1!}[A, B] + \frac{\lambda^2}{2!}[[A, B], B] + \cdots\right\}\end{aligned} \tag{3.60}$$

Show that if A and B are two operators that both commute with their commutator $[A, B]$, then

$$e^Ae^B = e^{A+B+[A,B]/2} \tag{3.61}$$

Another operator identity with many useful applications concerns differentiation of an operator that depends on a parameter t. The parameter could be the time t, but the formula is general and no commitment as to the nature of the parameter is required. If we suppose that $A(t)$ is an operator that depends on t, the derivative of $e^{A(t)}$ may be expressed as

$$\begin{aligned}\frac{d}{dt}e^{A(t)} &= \frac{d}{dt}\lim_{N\to\infty}\left\{1 + \frac{A(t)}{N}\right\}^N = \lim_{N\to\infty}\sum_{n=1}^{N}\left\{1 + \frac{A}{N}\right\}^{N-n}\frac{A'(t)}{N}\left\{1 + \frac{A}{N}\right\}^{n-1} \\ &= \lim_{N\to\infty}\sum_{k=1}^{N}\left\{1 + \frac{A}{N}\right\}^{k-1}\frac{A'(t)}{N}\left\{1 + \frac{A}{N}\right\}^{N-k}\end{aligned} \tag{3.62}$$

where we use the notation

$$A'(t) = \frac{dA}{dt} \tag{3.63}$$

The sum in (3.62) can be turned into an integral over the real variable x, defined as

$$x = \frac{n}{N}, \qquad \Delta x = \frac{\Delta n}{N} = \frac{1}{N} \quad \text{or} \quad y = \frac{k}{N}, \qquad \Delta y = \frac{\Delta k}{N} = \frac{1}{N}$$

so that

$$\frac{d}{dt} e^{A(t)} = \lim_{N\to\infty} \left(1 + \frac{A}{N}\right)^N \int_0^1 e^{-xA(t)} A'(t) e^{xA(t)}\, dx$$
$$= \int_0^1 e^{yA(t)} A'(t) e^{-yA(t)}\, dy \lim_{N\to\infty} \left(1 + \frac{\dot{A}}{N}\right)^N$$

and finally,

$$\frac{d}{dt} e^{A(t)} = e^{A(t)} \int_0^1 e^{-xA(t)} A'(t) e^{xA(t)}\, dx \tag{3.64}$$
$$= \int_0^1 e^{yA(t)} A'(t) e^{-yA(t)}\, dy\, e^{A(t)}$$

which is the desired result.[4]

Combining (3.64) with (3.54), and performing the x integrations term by term, we obtain

$$\frac{d}{dt} e^{A(t)} = e^{A(t)} \left\{ \frac{1}{1!} A'(t) - \frac{1}{2!} [A(t), A'(t)] + \frac{1}{3!} [A(t), [A(t), A'(t)]] - \cdots \right\} \tag{3.65}$$
$$= \left\{ \frac{1}{1!} A'(t) + \frac{1}{2!} [A(t), A'(t)] + \frac{1}{3!} [A(t), [A(t), A'(t)]] + \cdots \right\} e^{A(t)}$$

Exercise 3.17. Prove the identity (3.64) by a second method, starting from the function $F(\lambda) = \dfrac{d}{dt} e^{\lambda A(t)}$. Differentiate with respect to the variable λ and derive a first-order differential equation for the function F. Then solve this equation.

5. ***Stationary States and General Solutions of the Wave Equation.*** The operator formalism provides us with the tools to construct formal solutions of the time-dependent Schrödinger equation simply and efficiently. The fundamental mathematical problem is to obtain a solution of Eq. (3.42),

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi \tag{3.42}$$

such that at $t = 0$ the solution agrees with a given initial state $\psi(\mathbf{r}, 0)$.

If the system is conservative and the Hamiltonian operator H does not depend on t, the solution of (3.42) may be written as

$$\psi(\mathbf{r}, t) = \exp\left(-\frac{i}{\hbar} Ht\right) \psi(\mathbf{r}, 0) \tag{3.66}$$

The simplicity of this formula is deceptive. The Hamiltonian is usually the sum of a differential operator (kinetic energy T) and a function of position (potential energy V), and $e^{-iHt/\hbar}$ cannot be replaced by a product of operators, $e^{-iTt/\hbar}\, e^{-iVt/\hbar}$. It is fair

[4]For a compilation of many applications of (3.64), see Wilcox (1967).

to say that much of the rest of this book will be devoted to turning Eq. (3.66) into a more explicit and calculable form for different dynamical systems.

Here we make the more general observation that any initial state represented by an *eigenfunction* $\psi_E(\mathbf{r})$ of the Hamiltonian operator, corresponding to an *eigenvalue* E, gives rise to a particularly simple and important class of solutions (3.66). Eigenfunctions $\psi_E(\mathbf{r})$ and eigenvalues E of H are defined by the *eigenvalue equation*

$$H\psi_E(\mathbf{r}) = E\psi_E(\mathbf{r}) \tag{3.67}$$

or, more explicitly,

$$\boxed{-\frac{\hbar^2}{2m}\nabla^2\psi_E(\mathbf{r}) + V(\mathbf{r})\psi_E(\mathbf{r}) = E\psi_E(\mathbf{r})} \tag{3.68}$$

known plainly as the *Schrödinger equation*. This equation is also called the *time-independent Schrödinger equation* in contrast to the *time-dependent Schrödinger equation* (3.1).

Use of the symbol $\psi_E(\mathbf{r})$, or merely $\psi(\mathbf{r})$, in the (time-independent) Schrödinger equation for a function that is in general entirely different from $\psi(\mathbf{r}, t)$ is sanctioned by usage. Using a different symbol might spare the student a little confusion now but would make the reading of the literature more difficult later. Whenever necessary, care will be taken to distinguish the two functions by including the variables on which they depend.

If the initial state is represented by an eigenfunction $\psi_E(\mathbf{r})$ of H, the time-dependent wave function becomes, according to (3.66),

$$\boxed{\psi(\mathbf{r}, t) = \exp\left(-\frac{i}{\hbar}Et\right)\psi_E(\mathbf{r})} \tag{3.69}$$

because for any analytic function f of H,

$$f(H)\psi_E(\mathbf{r}) = f(E)\psi_E(\mathbf{r}) \tag{3.70}$$

The particular solution (3.69) of the time-dependent Schrödinger wave equation derives its great significance from the fact that more general solutions of (3.1) can be constructed by superposition of these *separable solutions*.

If the states represented by (3.69) are to be physically acceptable, they must conserve probability. Hence, the probability density

$$\rho = \psi^*(\mathbf{r}, t)\,\psi(\mathbf{r}, t) = \exp\left[-\frac{i}{\hbar}(E - E^*)t\right]\psi_E{}^*(\mathbf{r})\psi_E(\mathbf{r})$$

must satisfy condition (3.10), and this is possible only if the eigenvalues E are real numbers.

If E is real, the time-dependent factor in (3.69) is purely oscillatory. Hence, in the separable state (3.69) the functions $\psi(\mathbf{r}, t)$ and $\psi_E(\mathbf{r})$ differ only by a time-dependent phase factor of constant magnitude. This justifies the occasional designation of the spatial factor $\psi_E(\mathbf{r})$ as a time-independent wave function. (Often, for economy of notation, the subscript E will be dropped.) For such a state, the expectation value of any physical quantity that does not depend on time explicitly is constant and may be calculated from the time-independent wave function alone:

$$\langle A\rangle = \int \psi^*(\mathbf{r}, t)A\psi(\mathbf{r}, t)d^3r = \int \psi_E{}^*(\mathbf{r})A\psi_E(\mathbf{r})d^3r \tag{3.71}$$

From (3.67) we obtain

$$\langle H \rangle = \int \psi_E^*(\mathbf{r}) H \psi_E(\mathbf{r}) d^3r = E \tag{3.72}$$

showing that the eigenvalue E is also the expectation value of the energy in this state.

We thus see that the particular solutions

$$\psi(\mathbf{r}, t) = \exp\left(-\frac{i}{\hbar} Et\right)\psi_E(\mathbf{r}) \tag{3.69}$$

have intriguing properties. The probability density $\rho = |\psi(\mathbf{r}, t)|^2$ and all other probabilities are *constant in time*. For this reason, wave functions of the form (3.69) are said to represent *stationary states*. Note that although the state is stationary, the particle so described is in motion and not stationary at all. Also, $\nabla \cdot \mathbf{j} = 0$ for such a state.

Exercise 3.18. Using Eq. (3.7), prove the orthogonality relation

$$\int \psi_1^*(\mathbf{r})\psi_2(\mathbf{r}) d^3r = 0 \tag{3.73}$$

for any two stationary states with energies $E_1 \neq E_2$ and suitable boundary conditions.

We have already encountered one special example of a stationary state. This is the state of free particle of momentum $\mathbf{p}$ and consequently of energy $E = p^2/2m$, i.e., the plane wave

$$\psi(\mathbf{r}, t) = \exp\left(-\frac{i}{\hbar} Et\right) \exp\left(\frac{i}{\hbar} \mathbf{p} \cdot \mathbf{r}\right)$$

This wave function represents a state that has the *definite* value E for its energy. There is no uncertainty in the momentum and in the energy of this plane wave state; but there is, correspondingly, a complete ignorance of the position of the particle and of the transit time of the particle at a chosen position.

When conservative forces are introduced, the properties of momentum are, of course, drastically altered. But conservation of energy remains valid, as does the relation according to which the energy equals h times the frequency. Since a stationary state (3.69) has the definite (circular) frequency $\omega = E/\hbar$, it follows that a stationary state is a state of well-defined energy, E being the definite value of its energy and not only its expectation value. This is to be understood in the sense that any determination of the energy of a particle that is in a stationary state always yields a particular value E and only that value. Such an interpretation conforms with the uncertainty relation, $\Delta E \Delta t \geq \hbar$, which implies that a quantum state with a precise energy ($\Delta E = 0$) is possible only if unlimited time is available to determine that energy. Stationary states are of just such a nature in view of the constancy of $|\psi|$ in time.

The observed discrete energy levels of atoms and nuclei are approximated to a high degree of accuracy by stationary states. The simple regularity conditions that physical considerations lead us to impose on the wave function give rise in many cases to solutions of (3.68), which, instead of depending continuously on the parameter E, exist only for certain *discrete* values of E. The values of E for which (3.68)

has solutions consistent with the boundary conditions are said to constitute the *energy spectrum* of the Schrödinger equation. This spectrum may consist of isolated points (discrete energy levels) and continuous portions. If there is a lowest energy level, the corresponding stationary state is called the *ground state* of the system. All discrete higher energy levels are classified as *excited states*.

The careful phrasing concerning the approximate nature of the connection between energy levels and stationary states is intentional. The point at issue here is that no excited state of an atom is truly discrete, stationary, or permanent. If it were, we would not see the spectral lines that result from its decay to some other level by photon emission. The excited states are nevertheless stable to a high degree of approximation, for the period associated with a typical atomic state is of the order of the period of an electron moving in a Bohr orbit, i.e., about 10^{-15} sec, whereas the lifetime of the state is roughly 10^{-8} sec, as will be shown in Chapter 19. Hence, during the lifetime of the state some 10^7 oscillations take place. The situation is even more pronounced in the case of the so-called *unstable* particles, such as the kaon or the lambda particle. Their natural periods are $\cong 10^{-21}$ sec, and their lifetimes $\cong 10^{-10}$ sec. Thus, they live for about 10^{11} periods. Comparing this with the fact that during its history the earth has completed some 10^9 revolutions around the sun—a motion that has always been considered one of the stablest—the striking, if paradoxical, feature of the unstable particles is their stability!

The stationary state wave functions are particular solutions of the wave equation. More general solutions of (3.1) can be constructed by superposition of such particular solutions. Summing over various allowed values of E, we get the solution.

$$\psi(\mathbf{r}, t) = \sum_E c_E \exp\left(-\frac{i}{\hbar} Et\right) \psi_E(\mathbf{r}) \tag{3.74}$$

Here as elsewhere in this book the summation may imply an integration if part or all of the energy spectrum of E is continuous. Furthermore, the summation in (3.74) may include the same energy value more than once. This occurs if the Schrödinger equation, (3.67) or (3.68), has more than one linearly independent *physically acceptable* solution for the same eigenvalue E. It is then not sufficient to designate the eigenfunction by $\psi_E(\mathbf{r})$, and one or more additional labels must be used to identify a particular eigenfunction. The energy eigenvalue E is said to be *degenerate* in this case.

The example of the free particle in one dimension illustrates these features. In this case (3.74) becomes

$$\psi(\mathbf{r}, t) = \sum_E c_E \exp\left(-\frac{i}{\hbar} Et\right) \exp\left(\frac{i}{\hbar} px\right) \tag{3.75}$$

There is degeneracy of *degree two* here, because for any positive value of E, p can assume two values, $p = \pm\sqrt{2mE}$. The corresponding eigenfunctions are linearly independent. Since the energy spectrum is continuous and E can be any positive number, (3.75) must be regarded as an integral. Negative values of E, which correspond to imaginary values of p, are excluded, because ψ would become singular for very large values of $|x|$. The exclusion of negative values of E on this ground is an example of the operation of the physically required boundary conditions.

For a free particle in *three dimensions*, each positive energy E has an infinite degree of degeneracy, since plane waves propagating with the same wavelength in different directions are linearly independent. The eigenfunctions are $\exp(i\,\mathbf{p}\cdot\mathbf{r}/\hbar)$ and the eigenvalues are $E = (p_x^2 + p_y^2 + p_z^2)/2m$.

As a second-order linear differential equation, the Schrödinger equation in one dimension for an arbitrary potential V

$$-\frac{\hbar^2}{2m}\frac{d^2\psi_E(x)}{dx^2} + V(x)\psi_E(x) = E\psi_E(x) \tag{3.76}$$

has two linearly independent solutions for each value of E. The Wronskian, $\psi_E^{(1)}\psi_E^{(2)\prime} - \psi_E^{(1)\prime}\psi_E^{(2)}$, of two such solutions of (3.76), $\psi_E^{(1)}$ and $\psi_E^{(2)}$, is a nonzero constant for all x.

Exercise 3.19. Prove that the Wronskian for two solutions of the one-dimensional Schrödinger equation is constant.

Four possibilities can be distinguished:

(a) Both solutions and their derivatives are finite at all x, as is the case for the free particle ($V = 0$) and for a particle in a periodic potential, if the energy is in an allowed band. The degeneracy is *twofold*.

(b) One solution and its derivative approach zero as x approaches either $+\infty$ or $-\infty$, or both. The constancy of the Wronskian requires that the other solution diverges at large distance, so there is *no degeneracy*. If a solution approaches zero for *both* $x \to \pm\infty$ sufficiently rapidly, it represents a *bound* state.

(c) Both solutions *diverge* at large distance, for either $x \to +\infty$ or $-\infty$, or both. Neither solution represents a physically acceptable state. This is the case if E lies between discrete energy levels or, for a periodic potential, in a forbidden band.

(d) Finally, for a periodic potential if the energy E is at a band edge, only one solution remains finite, while the other one diverges.

In the preceding discussion, it was assumed that the coordinate x is the usual Cartesian coordinate ranging from $-\infty$ to $+\infty$. Sometimes we encounter systems that require the imposition of *periodic* boundary conditions on the solutions of the Schrödinger equation. For instance, this is the case for a particle that is constrained to move on a closed loop, such as a ring, and the boundary condition to be applied is the *single-valuedness* of the wave function.[5] Generally, the allowed stationary states will have twofold degeneracy.

The potential function $V(x)$ in the Schrödinger equation is usually smooth and well-behaved for most values of x. In many applications, however, $V(x)$ is nonanalytic at isolated points. In particular, we commonly encounter potential functions that are discontinuous or that have discontinuous slopes at some isolated points x_0. If such a singular potential function can be approximated as the limiting case of a regular one, the wave function ψ and its first derivative ψ' must be continuous everywhere, since the Schrödinger equation is a second-order differential equation.

More specifically, if $x = x_0$ is the point of singularity, we get by integrating the Schrödinger equation from $x = x_0 - \varepsilon$ to $x = x_0 + \varepsilon$:

$$\psi'(x_0 + \varepsilon) - \psi'(x_0 - \varepsilon) = \int_{x_0-\varepsilon}^{x_0+\varepsilon} \frac{2m}{\hbar^2}[V(x) - E]\psi(x)\,dx$$

As long as $V(x)$ is finite, whether analytic or not, this equation implies that ψ' is continuous across the singularity. Hence, ψ itself must also be continuous.

[5]For an analysis of the single-valuedness condition for wave functions, see Merzbacher (1962).

Thus, the *joining* condition to be assumed at a singular point of the potential, if V has a finite value, is that the *wave function and its slope must be matched continuously*. The probability current density, which is made up of ψ and ψ', must also be continuous—an evident physical requirement, if there are no sources and sinks of probability, i.e., of particles.

Fourier analysis tells us that for any given initial wave packet, $\psi(x, 0)$, the expansion coefficients in (3.75) are uniquely defined. Hence, $\psi(x, t)$ in the form (3.75) is not only *a* solution of the wave equation

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} \tag{3.77}$$

but it is the *most general* solution that can be of physical utility when describing the motion of a free particle. The generalization to three dimensions is straightforward.

Similarly, if the Schrödinger equation for an arbitrary potential yields solutions $\psi_E(\mathbf{r})$ that are *complete* in the sense that any initial state $\psi(\mathbf{r}, 0)$ of physical interest may be expanded in terms of them, i.e.,

$$\psi(\mathbf{r}, 0) = \sum_E c_E \psi_E(\mathbf{r}) \tag{3.78}$$

then we automatically know the solution of (3.1) for all t:

$$\boxed{\psi(\mathbf{r}, t) = \sum_E c_E \exp\left(-\frac{i}{\hbar} Et\right) \psi_E(\mathbf{r})} \tag{3.79}$$

Thus, $\psi(\mathbf{r}, t)$ is a superposition of stationary states, if the initial wave function can be expanded in terms of solutions of the Schrödinger equation.

One more important simple property of the wave equation (3.1) can be discussed without specifying V in further detail. It is based on the observation that, if the complex conjugate of (3.1) is taken and t is replaced by $-t$, we get

$$i\hbar \frac{\partial \psi^*(\mathbf{r}, -t)}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi^*(\mathbf{r}, -t) + V(\mathbf{r})\psi^*(\mathbf{r}, -t) \tag{3.80}$$

provided only that V is real. This equation has the same form as (3.1). Hence, if $\psi(\mathbf{r}, t)$ is a solution of (3.1), $\psi^*(\mathbf{r}, -t)$ is also a solution. The latter is often referred to as the *time-reversed* solution of (3.1) with respect to $\psi(\mathbf{r}, t)$. The behavior of the wave equation exhibited by (3.80) is called its *invariance under time reversal*. For a stationary state, invariance under time reversal implies that, if $\psi_E(\mathbf{r})$ is a stationary wave function, $\psi_E^*(\mathbf{r})$ is also one. Hence, it follows that if $\psi_E(\mathbf{r})$ is a nondegenerate solution of (3.68), it must be real, except for an arbitrary constant complex factor.

Exercise 3.20. Prove the statements of the last paragraph, and show that if degeneracy is present, the solutions of the Schrödinger equation with real V may be chosen to be real.

Most of the analysis in this section can be applied directly to the wave equation (3.21) in *momentum space*. The Fourier transform of the time-independent Schrödinger equation (3.68) for stationary states of energy E is the integral equation

$$\boxed{\begin{aligned}\frac{p^2}{2m}\phi_E(\mathbf{p}) + \frac{1}{(2\pi\hbar)^{3/2}} \int \exp\left(-\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}\right) V(\mathbf{r}) \\ \exp\left(\frac{i}{\hbar}\mathbf{p}'\cdot\mathbf{r}\right)\phi_E(\mathbf{p}')d^3r\, d^3p' = E\phi_E(\mathbf{p})\end{aligned}} \tag{3.81}$$

The corresponding time-dependent momentum wave function that solves the wave equation (3.21) is

$$\phi(\mathbf{p}, t) = e^{-iEt/\hbar}\, \phi_E(\mathbf{p}) \tag{3.82}$$

For a free particle, (3.81) is simply an algebraic equation, which may be written as

$$\left(\frac{p^2}{2m} - E\right)\phi_E(\mathbf{p}) = 0$$

This equation implies that $\phi_E(\mathbf{p}) = 0$ for all $\mathbf{p}$, except if $p^2 = 2mE$. In three dimensions, if $\mathbf{p}_0$ is a fixed momentum with energy $E = p_0^2/2m$, the momentum space energy eigenfunctions are then

$$\phi(\mathbf{p}) = \delta(\mathbf{p} - \mathbf{p}_0) \tag{3.83}$$

For a linear harmonic oscillator, with $V(x) = m\omega^2 x^2/2$, the Schrödinger equation in momentum space (3.80) reduces to (3.23), or

$$\frac{1}{2m}\, p_x^2 \phi_E(x) - \frac{1}{2}\, m\omega^2\, \frac{d^2\phi_E(x)}{dp_x^2} = E\phi_E(x) \tag{3.84}$$

which is similar to the Schrödinger equation in coordinate space. Furthermore, the boundary conditions are analogous: ϕ must approach zero for large values of $|p_x|$, just as ψ is required to tend to zero as $|x| \to \infty$. Since in one dimension there is no degeneracy for bound states, it follows that for the energy eigenstates of the harmonic oscillator the Fourier transforms $\psi_E(x)$ and $\phi_E(p_x)$ must be simply related to each other.

Exercise 3.21. Show that for the stationary states of the harmonic oscillator,

$$\phi_E(p_x) = \frac{e^{i\alpha_E}}{\sqrt{m\omega}}\, \psi_E\left(\frac{p_x}{m\omega}\right) \tag{3.85}$$

which is a special case of the result in Exercise 3.8.

Exercise 3.22. For real V, prove that the wave equation in momentum space is invariant under time reversal and that $\phi^*(-\mathbf{p}, -t)$ is the time-reversed solution with respect to $\phi(\mathbf{p}, t)$.

6. ***The Virial Theorem.*** A simple example will illustrate the principles of the last two sections. Consider the equation of motion for the operator $\mathbf{r} \cdot \mathbf{p}$. According to (3.44),

$$i\hbar\, \frac{d}{dt} \langle \mathbf{r} \cdot \mathbf{p} \rangle = \langle [\mathbf{r} \cdot \mathbf{p}, H] \rangle$$

By applying several of the rules (3.50) and the fundamental commutation relation (3.47), $xp_x - p_x x = i\hbar$, we obtain

$$\begin{aligned}
[xp_x, H] &= \left[xp_x, \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{p_z^2}{2m} + V\right] = \left[xp_x, \frac{p_x^2}{2m}\right] + [xp_x, V] \\
&= [x, p_x^2]\, \frac{p_x}{2m} + x[p_x, V] \\
&= p_x[x, p_x]\, \frac{p_x}{2m} + [x, p_x]\, \frac{p_x^2}{2m} + x[p_x, V] = i\hbar\, \frac{p_x^2}{m} + x[p_x, V]
\end{aligned}$$

Similar relations hold for the y and z components. Combining these results, we get

$$\begin{aligned}\frac{d}{dt}\langle \mathbf{r}\cdot\mathbf{p}\rangle &= \left\langle\frac{p^2}{m}\right\rangle - \int \psi^*\mathbf{r}\cdot\{\nabla(V\psi) - V\nabla\psi\}d^3r \\ &= \left\langle\frac{p^2}{m}\right\rangle - \int \psi^*(\mathbf{r}\cdot\nabla V)\psi d^3r\end{aligned} \tag{3.86}$$

Exercise 3.23. Show that $\mathbf{r}\cdot\mathbf{p} - \mathbf{p}\cdot\mathbf{r} = 3i\hbar\, 1$, and prove that in spite of the noncommutivity of the operators $\mathbf{r}$ and $\mathbf{p}$, their order on the left-hand side of (3.86) does not matter; the same equation holes for $\langle\mathbf{p}\cdot\mathbf{r}\rangle$.

As in classical mechanics,[6] we obtain from (3.86) for the time average over a time τ

$$\frac{1}{\tau}[\langle\mathbf{r}\cdot\mathbf{p}\rangle_{t=\tau} - \langle\mathbf{r}\cdot\mathbf{p}\rangle_{t=0}] = 2\overline{\langle T\rangle} - \overline{\langle\mathbf{r}\cdot\nabla V\rangle} = 2\overline{\langle T\rangle} + \overline{\langle\mathbf{r}\cdot\mathbf{F}\rangle}$$

If the expectation values in this expression remain finite as $\tau\to\infty$, the left-hand side tends to zero, and

$$2\overline{\langle T\rangle} = \overline{\langle\mathbf{r}\cdot\nabla V\rangle} = -\overline{\langle\mathbf{r}\cdot\mathbf{F}\rangle} \tag{3.87}$$

For a *stationary state* all expectation values in (3.87) are constant in time, and it follows that

$$\boxed{2\langle T\rangle = \langle\mathbf{r}\cdot\nabla V\rangle = -\langle\mathbf{r}\cdot\mathbf{F}\rangle} \tag{3.88}$$

The results (3.87) and (3.88) are known as the *virial theorem*. In a gas of weakly interacting molecules, departures from the ideal gas law can be related to the *virial*, $\sum_i \mathbf{r}_i\cdot\mathbf{F}_i$, which explains the terminology.

Exercise 3.24. Apply the virial theorem for a stationary state to the three-dimensional anisotropic harmonic oscillator for which

$$V(\mathbf{r}) = \frac{1}{2}m(\omega_1^2x^2 + \omega_2^2y^2 + \omega_3^2z^2)$$

Show that the same result could have been derived from the equality given in Exercise 3.21.

Exercise 3.25. If the potential has the property $V(\lambda x, \lambda y, \lambda z) = \lambda^n V(x, y, z)$, using Euler's theorem for homogeneous functions, show that, for a stationary state

$$\langle T\rangle = \frac{n}{2}\langle V\rangle \tag{3.89}$$

Exercise 3.26. Apply the virial theorem to a central potential of the form $V = Ar^n$, and express the expectation value of V for bound states in terms of the energy.

[6]Goldstein (1980), Section 3–4.

Problems

1. If the state $\psi(\mathbf{r})$ is a superposition,

$$\psi(\mathbf{r}) = c_1\psi_1(\mathbf{r}) + c_2\psi_2(\mathbf{r})$$

where $\psi_1(\mathbf{r})$ and $\psi_2(\mathbf{r})$ are related to one another by time reversal, show that the probability current density can be expressed without an interference term involving ψ_1 and ψ_2.

2. For a free particle in one dimension, calculate the variance at time t, $(\Delta x)_t^2 \equiv \langle (x - \langle x\rangle_t)^2\rangle_t = \langle x^2\rangle_t - \langle x\rangle_t^2$ without explicit use of the wave function by applying (3.44) repeatedly. Show that

$$(\Delta x)_t^2 = (\Delta x)_0^2 + \frac{2}{m}\left[\frac{1}{2}\langle xp_x + p_x x\rangle_0 - \langle x\rangle_0\langle p_x\rangle\right]t + \frac{(\Delta p_x)^2}{m^2}t^2$$

and

$$(\Delta p_x)_t^2 = (\Delta p_x)_0^2 = (\Delta p_x)^2$$

3. Consider a linear harmonic oscillator with Hamiltonian

$$H = T + V = \frac{p_x^2}{2m} + \frac{1}{2}m\omega^2 x^2$$

(a) Derive the equation of motion for the expectation value $\langle x\rangle_t$ and show that it oscillates, similarly to the classical oscillator, as

$$\langle x\rangle_t = \langle x\rangle_0 \cos\,\omega t + \frac{\langle p_x\rangle_0}{m\omega}\sin\,\omega t$$

(b) Derive a second-order differential equation of motion for the expectation value $\langle T - V\rangle_t$ by repeated application of (3.44) and use of the virial theorem. Integrate this equation and, remembering conservation of energy, calculate $\langle x^2\rangle_t$.

(c) Show that

$$(\Delta x)_t^2 \equiv \langle x^2\rangle_t - \langle x\rangle_t^2 = (\Delta x)_0^2\cos^2\,\omega t + \frac{(\Delta p_x)_0^2}{m^2\omega^2}\sin^2\,\omega t + \left[\frac{1}{2}\langle xp_x + p_x x\rangle_0 - \langle x\rangle_0\langle p_x\rangle_0\right]\frac{\sin\,2\omega t}{m\omega}$$

Verify that this reduces to the result of Problem 1 in the limit $\omega \to 0$.

(d) Work out the corresponding formula for the variance $(\Delta p_x)_t^2$.

4. Prove that the probability density and the probability current density at position $\mathbf{r}_0$ can be expressed in terms of the *operators* $\mathbf{r}$ and $\mathbf{p}$ as expectation values of the operators

$$\rho(\mathbf{r}_0) \to \delta(\mathbf{r} - \mathbf{r}_0) \qquad \mathbf{j}(\mathbf{r}_0) \to \frac{1}{2m}[\mathbf{p}\delta(\mathbf{r} - \mathbf{r}_0) + \delta(\mathbf{r} - \mathbf{r}_0)\mathbf{p}]$$

Derive expressions for these densities in the momentum representation.

5. For a system described by the wave function $\psi(\mathbf{r}')$, the *Wigner distribution* function is defined as

$$W(\mathbf{r}', \mathbf{p}') = \frac{1}{(2\pi\hbar)^3}\int e^{-i\mathbf{p}'\cdot\mathbf{r}''/\hbar}\,\psi^*\left(\mathbf{r}' - \frac{\mathbf{r}''}{2}\right)\psi\left(\mathbf{r}' + \frac{\mathbf{r}''}{2}\right)d^3r''$$

(In formulas involving the Wigner distribution, it is essential to make a notational distinction between the unprimed *operators*, **r** and **p**, and the real number variables, which carry primes.)

(a) Show that $W(\mathbf{r}', \mathbf{p}')$ is a real-valued function, defined over the six-dimensional "phase space" $(\mathbf{r}', \mathbf{p}')$.[7]

(b) Prove that

$$\int W(\mathbf{r}', \mathbf{p}')\, d^3p' = |\psi(\mathbf{r}')|^2$$

and that the expectation value of a function of the operator **r** in a normalized state is

$$\langle f(\mathbf{r})\rangle = \int\int f(\mathbf{r}')W(\mathbf{r}', \mathbf{p}')\, d^3r'\, d^2p'$$

(c) Show that the Wigner distribution is normalized as

$$\int W(\mathbf{r}', \mathbf{p}')\, d^3r'\, d^3p' = 1$$

(d) Show that the probability density $\rho(\mathbf{r}_0)$ at position $\mathbf{r}_0$ is obtained from the Wigner distribution with[8]

$$\rho\,(\mathbf{r}_0) \rightarrow f(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_0)$$

6. (a) Show that if $\phi(\mathbf{p}')$ is the momentum wave function representing the state, the Wigner distribution is

$$W(\mathbf{r}', \mathbf{p}') = \frac{1}{(2\pi\hbar)^3}\int \phi^*(\mathbf{p}' - \mathbf{p}'')\phi(\mathbf{p}' + \mathbf{p}'')e^{2i\mathbf{p}''\cdot\mathbf{r}'/\hbar}\, d^3p''$$

(b) Verify that

$$\int W(\mathbf{r}', \mathbf{p}')\, d^3r' = |\phi(\mathbf{p}')|^2$$

and that the expectation value of a function of the operator **p** is

$$\langle g(\mathbf{p})\rangle = \int\int g(\mathbf{p}')W(\mathbf{r}', \mathbf{p}')\, d^3r'\, d^3p'$$

[7]Although the integrals of $W(\mathbf{r}', \mathbf{p}')$ over coordinate and momentum space are probability distributions in momentum and coordinate space, respectively, the function $W(\mathbf{r}', \mathbf{p}')$, which can take negative values, is not a probability distribution (in phase space).

[8]Because it involves both **r** and **p** in the same expression, the analogous question for the probability current density **j** is deferred until Problem 2, Chapter 15.

CHAPTER 4

The Principles of Wave Mechanics

This chapter concludes our account of the general principles on which wave mechanics rests. The central issue is the existence of complete sets of orthonormal eigenfunctions of self-adjoint (Hermitian) operators, generalized to include the continuous spectrum. The momentum operator is related to coordinate displacements, and we encounter the first examples of symmetry operations. The formalism of nonrelativistic quantum mechanics is joined with the gauge-symmetric electromagnetic field. At the end, we should be ready for the application of quantum mechanics to spinless particles in one or more dimensions.

1. ***Hermitian Operators, their Eigenfunctions and Eigenvalues.*** We have learned that every physical quantity F can be represented by a linear operator, which for convenience is denoted by the same letter, F. The expectation value of F is given by the formula

$$\boxed{\langle F \rangle = \int \psi^* F \psi \, d^3r = \int \phi^* F \phi \, d^3p} \tag{4.1}$$

expressed either in the coordinate or the momentum representation. We now ask what general properties a linear operator must possess in order to be admitted as a candidate for representing a physical quantity.

If $\langle F \rangle$ is the expectation (or average) value of a linear operator representing a physical quantity whose measured values are real, $\langle F \rangle$ must also be *real*. By (4.1) this implies that for all ψ which may represent possible states we must demand that

$$\boxed{\int (F\psi)^* \psi \, d^3r = \int \psi^* F \psi \, d^3r} \tag{4.2}$$

Operators that have this property are called *Hermitian*. The property of being Hermitian is independent of the choice of representation.

Exercise 4.1. Prove that if F is a Hermitian operator, (4.2) can be generalized to

$$\int (F\psi_1)^* \psi_2 \, d^3r = \int \psi_1^* F \psi_2 \, d^3r \tag{4.3}$$

for any two possible states ψ_1 and ψ_2.

The momentum $\mathbf{p}$ is an example of a Hermitian operator. It was shown in Section 3.3 that for the calculation of the expectation value $\langle \mathbf{p} \rangle$ the momentum may be

represented by $\frac{\hbar}{i}\nabla$ if ψ vanishes sufficiently fast at large distances. But under the very same boundary conditions, integration by parts gives

$$\int \left(\frac{\hbar}{i}\nabla\psi\right)^*\psi\, d^3r = -\frac{\hbar}{i}\int (\nabla\psi^*)\psi\, d^3r = \int \psi^* \frac{\hbar}{i}\nabla\psi\, d^3r$$

Hence, condition (4.3) is satisfied. [The momentum **p** is also represented by the differential operator $(\hbar/i)\nabla$ if ψ satisfies periodic boundary conditions (Section 4.4).]

Given an arbitrary, not necessarily Hermitian, operator F, it is useful to define its (Hermitian) adjoint $F^\dagger$ by the requirement

$$\int f^* F^\dagger g\, d^3r = \int (Ff)^* g\, d^3r \tag{4.4}$$

where f and g are arbitrary functions (but subject to the condition that the integrals exist). The existence of the linear operator $F^\dagger$ for operators of the type $F(\mathbf{r}, \mathbf{p})$ can be verified by integration by parts. Comparing (4.4) with (4.3), we note that F is Hermitian if it is self-adjoint,

$$F^\dagger = F \tag{4.5}$$

Conversely, if an operator F is Hermitian, and if its adjoint exists and has the same domain as F, we have

$$\int (F\psi_1)^*\psi_2\, d^3r = \int \psi_1^* F\psi_2\, d^3r = \int \psi_1^* F^\dagger\psi_2\, d^3r$$

and hence F is self-adjoint. Since the physical interpretation requires that operators representing measurable physical quantities must be self-adjoint, it has become customary in quantum mechanics to use the terms *Hermitian* and *self-adjoint* synonymously. We shall follow this usage, although it glosses over some mathematical distinctions that are addressed in more thorough presentations.[1]

A number of simple theorems about operators can now be given. Their proofs, if not trivial, are indicated briefly:

1. The adjoint of the sum of two operators equals the sum of their adjoints: $(F + G)^\dagger = F^\dagger + G^\dagger$. The sum of two Hermitian operators is Hermitian.

2. The identity operator 1, which takes every function into itself, is Hermitian. If λ is a real number, $\lambda 1$ is Hermitian.

3. If F is non-Hermitian, $F + F^\dagger$ and $i(F - F^\dagger)$ are Hermitian. Hence, F can be written as a linear combination of two Hermitian operators:

$$\boxed{F = \frac{F + F^\dagger}{2} + i\frac{F - F^\dagger}{2i}} \tag{4.6}$$

[1]Many mathematically unimpeachable treatments have been published since the first appearance of von Neumann (1932). For a compact account, see Jordan (1969) and the references cited here. For a more recent discussion, see Ballentine (1990).

4. If F and G are two arbitrary operators, the adjoint of their product is given by

$$\boxed{(FG)^\dagger = G^\dagger F^\dagger} \tag{4.7}$$

with an important reversal of the order of the factors.

Proof

$$\int \psi^*(FG)^\dagger\psi\, d^3r = \int (FG\psi)^*\psi\, d^3r = \int (G\psi)^*F^\dagger\psi\, d^3r = \int \psi^*G^\dagger F^\dagger\psi\, d^3r$$

But ψ is arbitrary, hence (4.7) follows. If F and G are Hermitian,

$$(FG)^\dagger = GF \tag{4.8}$$

Corollary. The product of two Hermitian operators is Hermitian if and only if they commute.

5. The adjoint of a *complex* number λ is its complex conjugate λ^*.

6. All the operators listed in Table 3.1 representing physical quantities are Hermitian.

Since the weak requirement of being Hermitian is sufficient to establish the most important properties common to the operators that represent physical quantities, the generic term *dynamical variable* is often used for such operators. Thus, x and p_x are dynamical variables, but their product xp_x is not, because the two operators fail to commute (see property 4).

Exercise 4.2. Show that the fundamental commutation relation (3.47) for x and p_x is consistent with the Hermitian nature of the operators.

Exercise 4.3. From the time-dependent Schrödinger equation (3.42), prove the equation of motion (3.44) for the expectation value of an operator F, making use only of the Hermitian property of H but not of its explicit form.

It is not farfetched to suppose that the failure of x and p_x to commute is connected with the uncertainty relation for x and p_x and with the incompatibility of precise simultaneous values for the x coordinate and the x-component of momentum. Given any dynamical variable A, how do we find and characterize those particular states of the system in which A may be said to have a *definite* value? If we make the convention that specific numerical values of dynamical variables (physical quantities) will be denoted by primed symbols (e.g., A') to distinguish them from the operators representing physical quantities (e.g., A), the same question can be phrased in physical terms: If we measure A in a large number of separate systems, all replicas of each other and each represented by the same wave function ψ, under what conditions will all these systems yield the same sharp and accurately predictable value A'? Phrased still differently, what kind of ψ corresponds to a probability distribution of the value of A that is peaked at A' and has no spread?

In the particular state ψ, in which the outcome of every measurement of A is the same, A', the expectation value, $\langle f(A)\rangle$, of any function of A must be equal to $f(A')$. Hence, we must demand that, in such a state,

$$\langle f(A)\rangle = f(A') \tag{4.9}$$

for any function of A. In particular, for $f(A) = A$ we demand

$$\langle A\rangle = A' \tag{4.10}$$

and for $f(A) = A^2$ we require that

$$\langle A^2\rangle = A'^2 = \langle A\rangle^2 \tag{4.11}$$

According to Eq. (A.35) in the Appendix, (4.11) expresses the vanishing of the variance $(\Delta A)^2$ of A:

$$(\Delta A)^2 = \langle (A - \langle A\rangle)^2\rangle = \langle A^2\rangle - \langle A\rangle^2$$

If ψ is assumed to be quadratically integrable, this condition implies for the special state in which A has a sharp value A' that

$$\int \psi^*(A - A')^2\psi \, d^3r = 0$$

The operator A, being a dynamical variable, is Hermitian, and A' is real. Hence, the last equation becomes

$$\int [(A - A')\psi]^*(A - A')\psi \, d^3r = \int |(A - A')\psi|^2 \, d^3r = 0$$

from which it follows that

$$\boxed{A\psi = A'\psi} \tag{4.12}$$

A quadratically integrable function ψ that satisfies (4.12) is an *eigenfunction* of A, and the real number A' is the corresponding *eigenvalue*. All eigenvalues of A belong to its *spectrum*. An eigenfunction of A is characterized by the fact that operating with A on it has no effect other than multiplication of the function by the eigenvalue A'. It may be noted that (4.10) and (4.11) are sufficient to satisfy condition (4.9) for any function f of a dynamical variable.

The eigenvalues are the same in the coordinate or momentum representations; the corresponding eigenfunctions are Fourier transforms of each other. This follows from the linearity of Fourier transforms.

The Hamiltonian H representing the energy is an example of a dynamical variable. If the system has a classical analogue, the operator is obtained from the classical Hamiltonian function of q's and p's by replacing these physical variables by operators. Its eigenfunctions $\psi_E(\mathbf{r})$ or $\phi_E(\mathbf{p})$ represent the states in which measurement of the energy gives the certain and sharp values E. (We follow the customary usage of denoting the eigenvalues of H by E, rather than H', which consistency in notation would suggest.)

Our conclusion is that a system will reveal a definite value A', and no other value, when the dynamical variable A is measured if and only if it is in a state represented by the eigenfunction $\psi_{A'}$. We often say that the system is *in* an eigenstate of A. The only definite values of A which a system can thus have are the eigenvalues A'.

The reality of the eigenvalues of Hermitian operators needs no further proof, but a most important property of the eigenfunctions, their orthogonality, does. Actually, the two properties follow from the same simple argument. Consider two eigenfunctions of A, ψ_1 and ψ_2, corresponding to the eigenvalues A_1' and A_2':

$$A\psi_1 = A_1'\psi_1 \tag{4.13}$$

$$A\psi_2 = A_2'\psi_2 \tag{4.14}$$

Multiply (4.13) on the left by ψ_2^*; take the complex conjugate of (4.14), and multiply it on the right by ψ_1. Then integrate both equations over all space and subtract one from the other. Owing to the Hermitian property of A,

$$(A_1' - A_2'^*) \int \psi_2^* \psi_1 \, d^3r = 0 \tag{4.15}$$

If $\psi_2 = \psi_1$, then $A_2' = A_1'$, and hence $A_1'^* = A_1'$, which demonstrates again that all eigenvalues of A are real. Using this result, we then see from (4.15) that if $A_2' \neq A_1'$ the two eigenfunctions are *orthogonal* in the sense that

$$\int \psi_2^* \psi_1 \, d^3r = 0 \tag{4.16}$$

We conclude that eigenfunctions belonging to different eigenvalues are orthogonal.

Since the eigenvalues A' are real, Eq. (4.15) is trivially satisfied if two different eigenfunctions belong to a particular eigenvalue of A. An eigenfunction that is obtained from another one merely by multiplication with a constant is, of course, not considered "different." Rather, to be at all interesting, all the n eigenfunctions belonging to an eigenvalue A' must be *linearly independent*. This means that none of them may be expressed as a linear combination of the others.

Although any two eigenfunctions belonging to the same eigenvalue may or may not be orthogonal, it is always possible to express all of them as linear combinations of n eigenfunctions that are orthogonal and belong to that same eigenvalue (the *Schmidt orthogonalization method*).

Proof. Suppose that

$$A\psi_1 = A'\psi_1 \quad \text{and} \quad A\psi_2 = A'\psi_2$$

where ψ_2 is *not* a multiple of ψ_1, and where the overlap integral of ψ_1 and ψ_2 is defined as

$$\int \psi_1^* \psi_2 \, d^3r = K \neq 0$$

By interpreting ψ_1 and ψ_2 as vectors in a two-dimensional space, and $\int \psi_1^* \psi_2 \, d^3r$ as their scalar (Hermitian inner) product, we can depict the Schmidt procedure in Figure 4.1. We construct a new vector ψ_2' as a linear combination of ψ_1 and ψ_2 and demand that

$$\int \psi_1^* \psi_2' \, d^3r = 0$$

If ψ_1 and ψ_2 are assumed to be normalized to unity, ψ_2', also normalized, is

$$\psi_2' = \frac{K\psi_1 - \psi_2}{\sqrt{1 - |K|^2}} \tag{4.17}$$

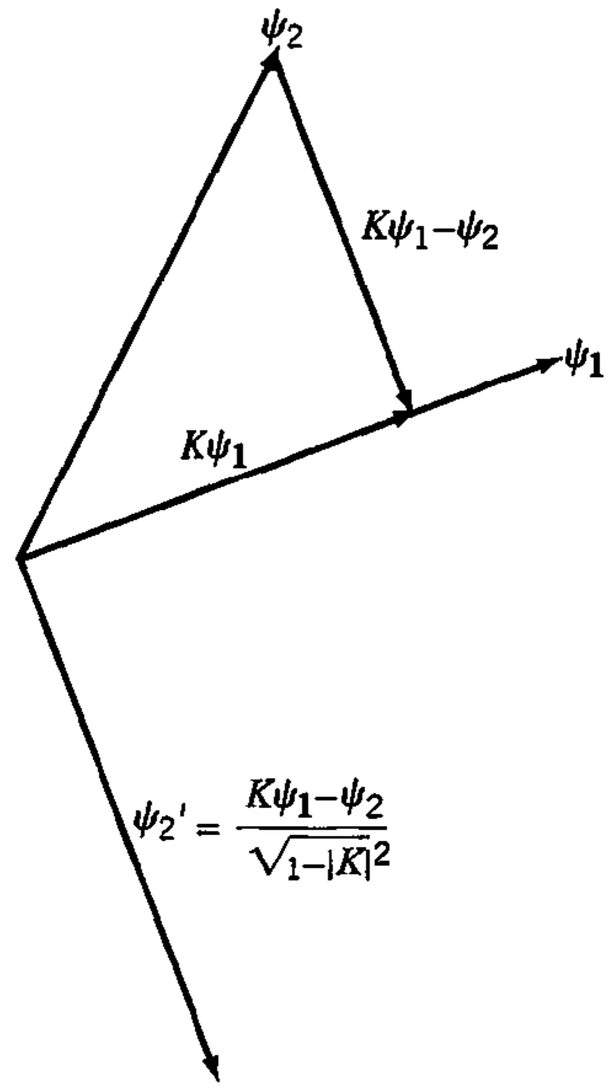


Figure 4.1. Two-dimensional (real-valued) vector analogue of Schmidt orthogonalization. The unit vectors represent the (generally complex-valued) functions ψ_1, ψ_2, and ψ_2'. The scalar product of the vectors ψ_1 and ψ_2 equals K, the length of the projection.

This new eigenfunction of A is orthogonal to ψ_1. If there are other eigenfunctions of A which correspond to the same eigenvalue A' and are linearly independent of ψ_1 and ψ_2, this process of orthogonalization can be continued systematically by demanding that ψ_3 be replaced by ψ_3', a linear combination of ψ_1, ψ_2', and ψ_3, so that ψ_3' is orthogonal to ψ_1 and ψ_2', and so on. When the process is completed, any eigenfunction of A with eigenvalue A' will be expressible as a linear combination of $\psi_1, \psi_2', \psi_3', \ldots, \psi_n'$. This orthogonal set of functions is by no means unique. Its existence shows, however, that there is no loss of generality if we assume that all eigenfunctions of the operator A are mutually orthogonal. Any other eigenfunction of A can be written as a *linear combination* of orthogonal ones. Henceforth we will therefore always assume that all the eigenfunctions of A are *orthogonal* and that the eigenfunctions, which are quadratically integrable, have been normalized to unity.

Since

$$\boxed{\int \psi_i^* \psi_j \, d^3r = \delta_{ij}} \tag{4.18}$$

the solutions of the equation

$$A\psi_i = A_i' \psi_i \tag{4.19}$$

are said to form an *orthonormal* set. In writing (4.18) and (4.19) we are, therefore, allowing for the possibility that two or more eigenvalues of A may be the same. If this happened, we speak of a *repeated* eigenvalue; otherwise we say that the eigenvalue is *simple*. The possible occurrence of repeated eigenvalues is not always made explicit in the formalism but should be taken into account when it arises, as it frequently does. In the case of the energy operator, energy levels that correspond to repeated eigenvalues of the Hamiltonian H are said to be *degenerate*.

2. ***The Superposition and Completeness of Eigenstates.*** We have seen the physical significance of the eigenstates of A. What happens if the physical system is not in such an eigenstate? To find out, let us first look at a state that is a *superposition* of eigenstates of A, such that the wave function can be written as

$$\boxed{\psi = \sum_i c_i \psi_i} \tag{4.20}$$

Owing to the orthonormality of the eigenfunctions, (4.18), the coefficients c_i are related to ψ by

$$\boxed{c_i = \int \psi_i^* \psi \, d^3r} \tag{4.21}$$

and the normalization integral is

$$\boxed{\int \psi^* \psi \, d^3r = \sum_I |c_i|^2 = 1} \tag{4.22}$$

The expectation value of A in the state ψ is given by

$$\langle A \rangle = \int \psi^* A \psi \, d^3r = \sum_i A_i' \, |c_i|^2 \tag{4.23}$$

It is important to remember that certain eigenvalues of A appearing in this sum may be repeated. The sums of all the $|c_i|^2$ which belong to the same eigenvalue of A will be symbolized by $\sum_i' |c_i|^2$.

More generally, for any function of A,

$$\boxed{\langle f(A) \rangle = \sum_i f(A_i') |c_i|^2} \tag{4.24}$$

of which (4.22) and (4.23) are special cases.

Exercise 4.4. Prove (4.21), (4.22), and (4.23).

From these equations we may infer that the eigenvalues of A_i' which are characteristic of the operator A, are in fact the only values of A that can be found in the measurement of this physical quantity, even if ψ is not an eigenstate of A. If this interpretation is made, then it follows that for simple eigenvalues, $|c_i|^2$, which depends on the state, is the *probability of finding the value A_i' when A is measured.* This conclusion comes about because two probability distributions of the random variable A_i' which produce the same expectation values for any function of A_i' must be identical; hence, the probabilities $|c_i|^2$ are uniquely determined. If ψ happens to be an eigenstate, ψ_i, of A, then $c_i = 1$, and all other $c_j = 0$ if $j \neq i$. In this particular case, (4.24) agrees with (4.9), showing the consistency of the present interpretation with our earlier conclusions. If an eigenvalue of A is repeated, the probability of finding it is the restricted sum $\sum_i' |c_i|^2$.

Although this interpretation is natural, it is still an *assumption* that a sharp line can be drawn between the *dynamical variables* of the system, which determine the

possible results of a measurement, and the state of the system, which determines the *actual* outcome of the measurement, at least in a probability sense.

Equation (4.20) shows how this division of the theory into physical quantities (operators) and particular states (wave functions) with its attendant consequences for the interpretation of A_i' and $|c_i|^2$ is related to the possibility of superposing states by constructing linear combinations of eigenfunctions. The state ψ is in a certain sense intermediate between its component states ψ_i. It bears the mark of each of them because a measurement of A in the state ψ may yield any of the eigenvalues A_i' which correspond to eigenfunctions ψ_i represented with nonvanishing contributions in the expansion (4.20) of ψ.

The interpretation of the formalism is incomplete, however, until we are assured that the wave function ψ of an arbitrary state can always be written as a linear combination of the eigenfunction of A and that there are no physical states which cannot be so represented.

Cogitation alone cannot provide this assurance, but experience has shown that it is legitimate to assume that any state can be represented to a sufficient approximation by a superposition of eigenstates of those dynamical variables that are actually observed and measured in experiment. Mathematically, the superposition appears as a linear combination of the appropriate eigenfunctions, so that we may generalize (4.20) and write for any state ψ:

$$\psi = \sum_{i=1}^{\infty} c_i \psi_i \tag{4.25}$$

where the sum on the right side includes all physically acceptable eigenfunctions of A. Since this infinite sum does not generally converge pointwise, the sense in which it represents the wave function ψ must be clarified. If (4.25) is to be treated as if the sum were finite, so that integrations, summations, and differentiations may be liberally interchanged, it is sufficient to require that the sum $\sum_i c_i \psi_i$ converge to a state ψ in accordance with the condition

$$\lim_{n\to\infty} \int \left| \psi - \sum_{i=1}^{n} c_i \psi_i \right|^2 d^3r = 0 \tag{4.26}$$

This kind of convergence, which is far more permissive than point-by-point uniform convergence, is known as *convergence in the mean*. The sum $\sum c_i \psi_i$ may be said to approximate ψ in the sense of a *least-square fit* at all coordinate points.

Exercise 4.5. If a set of n orthonormal functions ψ_i is given, show that the superposition $\sum_{i=1}^{n} c_i \psi_i$ which best approximates a state ψ in the sense of a least-square fit and which minimizes the integral

$$I_n = \int \left| \psi - \sum_{i=1}^{n} c_i \psi_i \right|^2 d^3r \tag{4.27}$$

corresponds to the choice (4.21) for the coefficients c_i. Prove that for this optimal condition Bessel's inequality

$$\int |\psi|^2 \, d^3r \geq \sum_{i=1}^{n} |c_i|^2 \tag{4.28}$$

holds. As n is increased and a new $n + 1$-st function is included in the superposition of orthonormal functions, note that the first n coefficients c_i retain their values and that $I_n \geq I_{n+1}$.

If in the limit $n \to \infty$ the sequence I_n approaches zero, the set of orthonormal functions represents ψ in the sense of (4.26) and we have for all ψ *Parseval's formula*,

$$\boxed{\int |\psi|^2 \, d^3r = \sum_{i=1}^{\infty} |c_i|^2} \tag{4.29}$$

This equation is called the *completeness relation* and expresses the condition that an operator A must satisfy if its eigenfunctions ψ_i are to be sufficient in number to represent an arbitrary state. Such a set of eigenfunctions is said to be *complete*. For a large class of simple operators—notably those known to be of the Sturm-Liouville type—the completeness of the eigenfunctions can be proved rigorously, but quantum mechanics requires the use of many far more general operators for which the completeness proof is hard to come by, and one has to proceed heuristically, assuming completeness unless proven otherwise.

The assumption that every physical operator possesses a complete set of orthogonal eigenfunctions is spelled out in the *expansion postulate* of quantum mechanics:

> *Every physical quantity can be represented by a Hermitian operator with eigenfunctions* $\psi_1, \psi_2, \ldots, \psi_n$, *and every physical state by a sum* $\sum_i c_i \psi_i$, *where the coefficients are defined by* (4.21).

Following Dirac, we sometimes call an Hermitian operator that possesses a complete set of eigenfunctions an *observable*.

According to the discussion at the beginning of this section, if A_i' is a simple eigenvalue, $|c_i|^2$ is the probability of finding the value A_i' for the physical quantity A. If A_i' is repeated, the probability is the restricted sum $\sum_i' |c_i|^2$. The coefficients c_i are called *probability amplitudes*. They are determined by the state according to (4.21) and conversely determine the state fully [Eq. 4.20)].

In the last section, we interpreted functions as vectors in order to give a geometric interpretation of the Schmidt orthogonalization procedure. More generally, the expansion of an arbitrary wave function in terms of an orthonormal set of eigenfunctions is reminiscent of the expansion of a vector in terms of an orthonormal set of basis vectors. The expansion coefficients are analogous to vector components, but the dimensionality of the vector space is no longer finite. The integrals in Eqs. (4.18), (4.21), and (4.22) can be thought of as *inner products*, akin to the scalar products for ordinary vectors, and (4.22) and (4.29) represent the formula for the square of the "length," or *norm*, of the "vector" ψ. Linear operators are like tensors of rank two, which act on vectors to produce other vectors.

This geometrically inspired approach of working in a generalized vector (or Hilbert) space, suitably extended to complex-valued vectors, underlies the general theory of Chapters 9 and 10. It owes its power and aesthetic appeal to its independence of any particular representation.

3. ***The Continuous Spectrum and Closure.*** In the last section, we have assumed that the eigenfunctions of the Hermitian operator A representing a physical quantity are *countable*, although they may be infinite in number. It is essential to remove this limitation, since many physical quantities, such as the linear momentum of a particle, have a *continuum* of values. The chief obstacle that restricts the eigenvalue spectrum to discrete eigenvalues is the requirement (4.18) of quadratic integrability of eigenfunctions.

If quadratic integrability is too narrow a restriction to impose on the eigenfunctions of A, what should be the criterion for selecting from the solutions of the equation

$$A\psi_{A'} = A'\psi_{A'} \tag{4.30}$$

those eigenfunctions that make up a complete set for *all* physical states?

The subscript A' has been attached to the eigenfunctions in the continuum to label them. As before, the discrete eigenfunctions are labeled as ψ_i corresponding to the eigenvalue A_i'.

Boundedness of the eigenfunctions everywhere is usually, but not always, a useful condition that draws the line between those solutions of (4.30) that must be admitted to the complete set of eigenfunctions and those that must be rejected. Generally, it is best to assume that the expansion postulate holds and to let nature tell us how wide we must make the set of eigenfunctions so that any physically occurring ψ can be expanded in terms of it. We thus tailor the mathematics to the requirements of the physics, the only assumption being that there is some fabric that will suit these requirements.[2]

If the probability interpretation of quantum mechanics is to be maintained, all eigenvalues of A must be real, and eigenfunctions that are not quadratically integrable can appear in the expansion of a quadratically integrable wave function ψ only with infinitesimal amplitude. Hence, these functions are part of the complete set only if they belong to a continuum of real eigenvalues, depend on the eigenvalue continuously, and can be integrated over a finite range of eigenvalues. Thus, (4.25) must be generalized to read

$$\boxed{\psi = \sum_i c_i\psi_i + \int c_{A'}\psi_{A'}\, dA'} \tag{4.31}$$

In the notation it is tacitly assumed that for a continuous eigenvalue A' there is only one eigenfunction. If there are several (possibly infinitely many) linearly independent eigenfunctions corresponding to an eigenvalue A', more indices must be used and summations (or integrations) carried out over these.

We now extend to the continuum the fundamental assumptions formulated in the expansion postulate and generalize (4.24) by requiring that

$$\langle f(A)\rangle = \int \psi^* f(A)\psi\, d^3r = \sum_i |c_i|^2\, f(A_i') + \int |c_{A'}|^2\, f(A')\, dA' \tag{4.32}$$

Thus, $|c_{A'}|^2$ in the continuum case is the *probability density* of finding A to have the measured value A'. More precisely, $|c_{A'}|^2\, dA'$ is the probability that a measurement of A yields values between A' and $A'+dA'$. Substitution of (4.31) into

[2]See references in footnote 1.

$\int \psi^* f(A)\psi \, d^3r$ shows that the right-hand side of (4.32) is obtained only if, for the continuous spectrum, we have the orthogonality condition

$$\boxed{\int \psi^*_{A'} \psi_{A''} \, d^3r = \delta(A' - A'')} \tag{4.33}$$

and if the continuous and discrete eigenfunctions are orthogonal to each other. We say that the eigenfunctions $\psi_{A'}$ are subject to *A-normalization* if (4.33) holds. This equation expresses orthogonality and *delta-function normalization* of the continuum eigenfunctions. With the requirement that the eigenvalues must be real, it can be merged into the single condition

$$\int \psi^*_{A'} A\psi_{A''} \, d^3r - \int (A\psi_{A'})^* \psi_{A''} \, d^3r = (A'' - A')\delta(A' - A'') \tag{4.34}$$

But the right-hand side of this equation vanishes, since $x\delta(x) = 0$. Hence, whether or not the eigenfunctions are quadratically integrable, we must demand that, the Hermitian property of A,

$$\int \psi^*_{A'} A\psi_{A''} \, d^3r = \int (A\psi_{A'})^* \psi_{A''} \, d^3r \tag{4.35}$$

should hold for the physically admissible eigenfunctions, in the usual sense of equations involving delta functions. The orthonormality condition (4.33) permits evaluation of the *expansion coefficients* in (4.31), giving

$$c_{A'} = \int \psi^*_{A'} \psi \, d^3r \tag{4.36}$$

in close analogy with (4.21).

We remark that the derivations in the last paragraph succeed only if the order of the integrations over space and over the range of eigenvalues A' can be interchanged. The license for such interchanges is implicit in the extended expansion postulate and is assumed without further notice.

To keep the notation compact, we often write the equations of quantum mechanics as if all eigenvalues were discrete, the summations over the eigenvalues implying integrations over any continuous portions of the spectrum. This convention is suggested by the formal similarity of the expansion equations

$$\psi = \sum_i c_i \psi_i, \qquad c_i = \int \psi^*_i \psi \, d^3r \tag{4.37}$$

for the discrete, and

$$\psi = \int c_{A'} \psi_{A'} \, dA', \qquad c_{A'} = \int \psi^*_{A'} \psi \, d^3r \tag{4.38}$$

for the continuous spectrum.

Moreover, it is always possible to make all eigenfunctions quadratically integrable, and the entire spectrum discrete, by applying boundary conditions that confine the particle to a limited portion of space. For instance, impenetrable walls at great distances can be erected, or periodicity in a large unit cell can be required. The imposition of such boundary conditions ''quantizes'' the continuous spectrum

of A. The spacing between the discrete eigenvalues A_i' decreases as the confining volume grows, and in the limit of infinitely receding boundaries, portions of the discrete spectrum go over into a continuous one. The transition is made by introducing a *density of (discrete) states*, $\rho(A')$. This is the number of eigenstates ψ_i per unit eigenvalue interval:

$$\rho(A') = \frac{\Delta N}{\Delta A'} \tag{4.39}$$

In the limit of an infinitely large confining region, the density ρ becomes infinite, since then the eigenvalues are truly continuous. But usually ρ is proportional to the volume of the enclosure, and it is then possible to speak of a finite *density of states per unit volume*.

It is easy to see that consistency between the expansion formulas (4.37) and (4.38) is achieved if the relations

$$\psi_{A'} = \psi_i \sqrt{\rho(A')}, \qquad c_{A'} = c_i \sqrt{\rho(A')} \tag{4.40}$$

are adopted. Similarly,

$$\delta(A' - A'') = \rho(A')\delta_{ij} \tag{4.41}$$

relates the discrete Kronecker delta to the ''continuous'' delta function. Hence, if the equations are written with the notation of the discrete spectrum, it is a simple matter to take proper cognizance of the equally important continuous spectrum, without which the set of eigenfunctions of an operator A would not be complete. Many of the features of the continuous spectrum will be illustrated in Section 4.4.

A useful condition that a given set of orthonormal functions must satisfy, if it is to be complete can be derived from the identity

$$\psi(\mathbf{r}) = \sum_i c_i \psi_i(\mathbf{r}) = \sum_i \psi_i(\mathbf{r}) \int \psi_i^*(\mathbf{r}')\psi(\mathbf{r}')\, d^3r'$$

$$= \int \left\{ \sum_i \psi_i^*(\mathbf{r}')\psi_i(\mathbf{r}) \right\} \psi(\mathbf{r}')\, d^3r'$$

Since this must be true for *any* function $\psi(\mathbf{r})$, we infer that

$$\boxed{\sum_i \psi_i^*(\mathbf{r}')\psi_i(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}')} \tag{4.42}$$

which is known as the *closure relation*. If the set of orthonormal functions included in the sum is incomplete, the sum will fail to represent the delta function.

Exercise 4.6. We know that $|\psi_i(\mathbf{r})|^2$ is the probability density in coordinate space if the system is known to be in the i-th eigenstate of A. Use the closure relation to deduce, conversely, that the same quantity measures the (relative) probability of finding the system to yield A_i' in a measurement of A, if the system is known to be at position $\mathbf{r}$.

4. ***A Familiar Example: The Momentum Eigenfunctions and the Free Particle.*** In the coordinate representation, the linear momentum $\mathbf{p}$ is represented by the operator $\mathbf{p} \to \hbar\nabla/i$. Its eigenvalue equation

$$\frac{\hbar}{i} \nabla \psi_{\mathbf{p}}(\mathbf{r}) = \mathbf{p}\psi_{\mathbf{p}}(\mathbf{r}) \tag{4.43}$$

has the orthonormal eigenfunctions

$$\boxed{\psi_{\mathbf{p}}(\mathbf{r}) = \frac{1}{(2\pi\hbar)^{3/2}} e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}} \tag{4.44}$$

and eigenvalues $-\infty < p_x, p_y, p_z < +\infty$. Each of the three momentum components supports a spectrum of repeated eigenvalues, but the plane waves (4.44) are simultaneous eigenfunctions of all three components, and the three eigenvalues of p_x, p_y, and p_z together completely specify the eigenfunction. The occurrence of simultaneous eigenfunctions of commuting operators, like the three momentum components, to specify a complete set of eigenfunctions will be discussed in detail later, but the explicit example of (4.44) already illustrates the issues involved.

The expansion postulate implies that an arbitrary wave function can be written in terms of momentum probability amplitudes $\phi(\mathbf{p})$ as

$$\psi(\mathbf{r}) = \frac{1}{(2\pi\hbar)^{3/2}} \int \phi(\mathbf{p}) e^{(i/\hbar)\mathbf{p}\cdot\mathbf{r}}\, d^3p \tag{4.45}$$

and that the expansion coefficient is given by

$$\phi(\mathbf{p}) = \frac{1}{(2\pi\hbar)^{3/2}} \int \psi(\mathbf{r}) e^{-(i/\hbar)\mathbf{p}\cdot\mathbf{r}}\, d^3r \tag{4.46}$$

These are, of course, just the standard formulas of Fourier analysis (Section 3.3). The orthogonality of the eigenfunctions is expressed by

$$\frac{1}{(2\pi\hbar)^3} \int e^{(i/\hbar)(\mathbf{p}-\mathbf{p}')\cdot\mathbf{r}}\, d^3r = \delta(\mathbf{p} - \mathbf{p}') \tag{4.47}$$

which is an integral representation for the delta function (see Appendix, Section 1). This equation shows that (4.44) is given in the **p**-normalization.

If, in the above formulas, we set $\mathbf{p} = \hbar\mathbf{k}$ and then choose units so that $\hbar = 1$, the eigenfunctions are said to be given in the **k**-normalization, and Fourier transforms like (2.16) and (2.17) are obtained.

Exercise 4.7. Show that the eigenfunctions of momentum satisfy the closure relation (4.42).

In one dimension it is sometimes convenient to use a closely related set of eigenfunctions. These are the eigenfunctions of the free particle Hamiltonian,

$$H = \frac{p^2}{2m} = -\frac{\hbar^2 d^2}{2m\, dx^2} \tag{4.48}$$

They are doubly degenerate for each energy $E > 0$:

$$\psi_E(x) = A(E)e^{ikx} \quad \text{and} \quad B(E)e^{-ikx}$$

corresponding to the two directions in which the particle can move. The nonnegative quantum number k is related to the energy by

$$\hbar k = \sqrt{2mE}$$

The *energy normalization* is determined by the conditions

$$\int_{-\infty}^{+\infty} A^*(E)A(E')e^{i(k'-k)x}\, dx = \delta(E - E') \tag{4.49}$$

and

$$\int_{-\infty}^{+\infty} B^*(E)B(E')e^{-i(k'-k)x}\,dx = \delta(E - E') \tag{4.50}$$

Using the properties of the delta function (compiled in the Appendix, Section 1), we can write (4.49) as

$$2\pi|A(E)|^2 \frac{\hbar}{\sqrt{2m}} \delta(\sqrt{E} - \sqrt{E'}) = \delta(E - E')$$

If $E \neq 0$, the identity (A.32) in the Appendix,

$$\delta(\sqrt{E} - \sqrt{E'}) = 2\sqrt{E}\delta(E - E')$$

gives the normalization constant

$$|A(E)|^2 = \sqrt{\frac{m}{2E}} \frac{1}{2\pi\hbar} = \frac{m}{hp} = \frac{1}{hv} \tag{4.51}$$

where $v = p/m$ is the particle speed. A similar calculation gives $|B(E)|^2$. Hence, if the phase constants are chosen equal to zero, the energy-normalized eigenfunctions for the free particle in one dimension are ($k > 0$)

$$\psi_E(x) = \left(\frac{m}{8\pi^2\hbar^2 E}\right)^{1/4} e^{\pm ikx} \tag{4.52}$$

Any two eigenfunctions (4.52) with different signs in the exponent (representing waves propagating in opposite directions) are orthogonal, whether or not they belong to the same energy.

The theory of Fourier series teaches us that there are alternate methods of constructing complete sets of discrete quantized energy eigenfunctions for the free particle Hamiltonians. In *three dimensions*, the eigenvalues of the Hamiltonian

$$H = -\frac{\hbar^2}{2m}\nabla^2$$

are defined by imposing certain boundary conditions on the eigenfunctions

$$\psi_{\mathbf{k}}(\mathbf{r}) = C(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}}$$

A particularly useful set is obtained if we require the eigenfunctions to satisfy *periodic boundary conditions*, such that upon translation by a fixed length L in the direction of any one of the three coordinate axes all eigenfunctions of H shall assume their previous values. This periodicity condition restricts the allowed momenta or wave vectors to

$$\boxed{k_x = \frac{2\pi}{L} n_x, \qquad k_y = \frac{2\pi}{L} n_y, \qquad k_z = \frac{2\pi}{L} n_z} \tag{4.53}$$

where n_x, n_y, and n_z are integers (positive, negative, or zero). The wave functions may be normalized in the basic cube of side length L; thus:

$$\int_0^L \int_0^L \int_0^L |\psi_{\mathbf{k}}(\mathbf{r})|^2\,dx\,dy\,dz = 1$$

Hence,

$$\psi_{\mathbf{k}}(\mathbf{r}) = L^{-3/2}\, e^{i\mathbf{k}\cdot\mathbf{r}} \tag{4.54}$$

and the energy eigenvalues are the *discrete* numbers

$$E = \frac{\hbar^2 k^2}{2m} = \frac{2\pi^2\hbar^2}{mL^2}\,(n_x^2 + n_y^2 + n_z^2) \tag{4.55}$$

Since each allowed energy value can be obtained by a number of different choices of the quantum numbers n_x, n_y, n_z, there is, except for $E = 0$, considerable degeneracy here. As the energy E increases, the number of ways in which three perfect squares can be added to yield

$$n_x^2 + n_y^2 + n_z^2 = \frac{mEL^2}{2\pi^2\hbar^2} \tag{4.56}$$

rises rapidly. The number of states (4.54) in the volume $L^3\,\Delta p_x\,\Delta p_y\,\Delta p_z$ of phase space is $\Delta n_x\,\Delta n_y\,\Delta n_z$, and according to (4.53) we have

$$L^3\,\frac{\Delta p_x\,\Delta p_y\,\Delta p_z}{\Delta n_x\,\Delta n_y\,\Delta n_z} = h^3 \tag{4.57}$$

In statistical mechanics, this last equation is somewhat loosely interpreted by saying: "Each state occupies a volume h^3 in phase space." The number, $\Delta N/\Delta E$, of eigenstates per unit energy interval for any but the lowest energies is to good approximation

$$\boxed{\frac{\Delta N}{\Delta E} = \frac{L^3 4\pi p^2 dp}{h^3\, dE} = \frac{m^{3/2}\sqrt{E}L^3}{\sqrt{2}\pi^2\hbar^3}} \tag{4.58}$$

Exercise 4.8. Use the Euclidean space in which lattice points are labeled by the integer coordinates n_x, n_y, n_z to derive the (asymptotic) formula (4.58) for the density of states[3]. Obtain a formula for the cumulative number Z of states whose energies are less than or equal to E. Compute the number of ways in which one can add the squares of three integers (positive, negative, or zero) such that their sum, $n_x^2 + n_y^2 + n_z^2$, successively assumes values no greater than 0, 1, 2, . . . , 100. Compare the numerical results with the asymptotic formula.

Exercise 4.9. Prove that the free particle energy eigenfunctions, which are subject to periodic boundary conditions, satisfy the orthogonality conditions appropriate to a discrete eigenvalue spectrum, provided that the integrations are extended over the periodicity volume.

In *one dimension*, free particle energy eigenfunctions that are subject to *periodic* boundary conditions give rise to an energy spectrum

$$\boxed{E = \frac{2\pi^2\hbar^2}{mL^2}\,n^2 \qquad (n = 0, \pm 1, \pm 2, \ldots)} \tag{4.59}$$

[3]Brehm and Mullin (1989), pp. 82–84.

and each eigenvalue, except $E = 0$, is doubly degenerate. The density of states is approximately

$$\frac{\Delta N}{\Delta E} = \frac{2|\Delta n|}{\Delta E} \approx \frac{|n|}{E} = \sqrt{\frac{2m}{E}}\frac{L}{h} \tag{4.60}$$

The main difference between this density of states and its three-dimensional analogue (4.58) is that it decreases with energy as $1/\sqrt{E}$, whereas the latter is proportional to $\sqrt{E}$. This different behavior is due to the fact that the degree of degeneracy of each energy value is independent of E, and equal to 2, for one dimension, whereas in three-dimensional space it increases linearly with E.

Closely related to the foregoing formulation of the free particle problem is the problem of the "particle in a box," familiar from every introductory course in quantum physics. In Section 6.4 we will return to this problem as a special case of a one-dimensional potential square well of infinite depth, but here we simply assume zero potential energy inside a finite volume at whose boundaries V jumps to an infinitely high value. Since the expectation value of the potential energy must be finite, the wave function must vanish within and on the walls of the box. As a convenient example, we choose the box to be a cube of length L, situated in the first coordinate octant with one corner at the origin. Inside the box, the Schrödinger equation is

$$-\frac{\hbar^2}{2m}\nabla^2\psi(x, y, z) = E\psi(x, y, z) \tag{4.61}$$

In a box geometry, the solutions are obtained by separation of variables. Thus, we set

$$\psi(x, y, z) = X(x)Y(y)Z(z) \tag{4.62}$$

and find that the Schrödinger equation separates into three equations:

$$\frac{d^2X}{dx^2} + k_x^2X = 0, \qquad \frac{d^2Y}{dy^2} + k_y^2Y = 0, \qquad \frac{d^2Z}{dz^2} + k_z^2Z = 0 \tag{4.63}$$

and

$$E = \frac{\hbar k^2}{2m} = \frac{\hbar^2}{2m}(k_x^2 + k_y^2 + k_z^2)$$

The eigenfunctions that satisfy the boundary conditions

$$\psi(0, y, z) = \psi(L, y\, z) = \psi(x, 0, z) = \psi(x, L, z) = \psi(x, y, 0) = \psi(x, y, L) = 0$$

are

$$\psi(x, y, z) = C \sin k_x x \sin k_y y \sin k_z z = C \sin \frac{\ell_x \pi}{L} x \sin \frac{\ell_y \pi}{L} y \sin \frac{\ell_z \pi}{L} z \tag{4.64}$$

with ℓ_x, ℓ_y, ℓ_z satisfying the condition

$$\ell_x^2 + \ell_y^2 + \ell_z^2 = \frac{2mEL^2}{\pi^2\hbar^2} \tag{4.65}$$

The values of ℓ_x, ℓ_y, ℓ_z are restricted to be positive integers, since a change of sign in (4.64) produces no new linearly independent eigenfunction. Hence, the degree of degeneracy of any given energy value is less here by a factor of 8 than in the case of periodic boundary conditions. However, by virtue of the different right-hand sides of (4.56) and (4.65), for a given energy E there are now more ways of finding suitable

values to satisfy (4.65). These two effects compensate each other, and as a result the density of states is the same for the case of periodic boundary conditions and the box. In either case, the eigenfunctions constitute a complete set, in terms of which an arbitrary wave packet can be expanded by a Fourier series within the basic periodicity cube.

This comforting result is an example of a very general property of the density of states: It is insensitive to the specific details of the boundary conditions and is dependent only on the size (volume) of the region of integration.[4]

Exercise 4.10. For the energy eigenfunctions in the one-dimensional box (Figure 4.2), compute the first few terms in the closure relation (4.42). Show how better

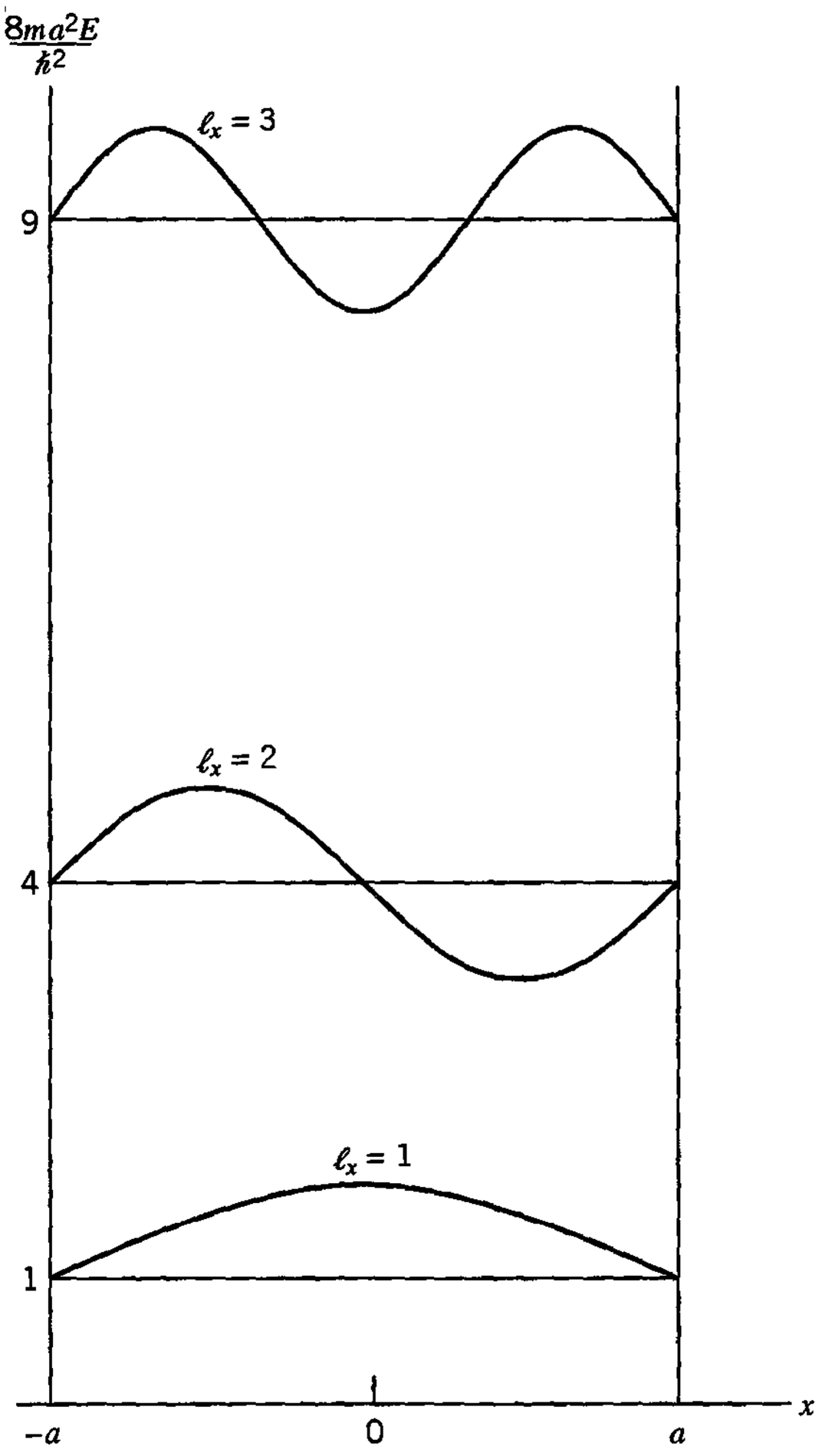


Figure 4.2. The three lowest energy eigenvalues ($\ell_x = 1,2,3$) and the corresponding eigenfunctions for a one-dimensional box of dimension $L = 2a$. The unit of energy on the ordinate is $\hbar^2/8ma^2$. The maximum amplitude of the normalized eigenfunctions is $\sqrt{1/a}$.

[4]Courant and Hilbert (1953).

approximations to the delta function $\delta(x - x')$ are obtained as the number of terms is increased.

5. ***Unitary Operators: The Displacement Operator.*** As representatives of physical observables, Hermitian operators play a paramount role in quantum mechanics, but *unitary operators* are of comparable importance. The time development operator, which links the initial state of a system at $t = 0$ to its state at an arbitrary time t, as formalized in (3.66), is unitary. Unitary operators also serve to transform quantum mechanical states that describe equivalent physical situations into one another. They are, therefore, at the heart of every analysis that seeks to take advantage of the symmetry properties of the physical system under consideration. A systematic discussion of symmetry in quantum mechanics is taken up in Chapter 17, but it is instructive to illustrate the underlying ideas by some simple examples in the realm of one-particle wave mechanics.

A unitary operator U is defined by the condition that it must have an inverse and preserve the norm of all state vectors:

$$\boxed{\int [U\psi(\mathbf{r})]^* U\psi(\mathbf{r})\, d^3r = \int \psi^*(\mathbf{r})\psi(\mathbf{r})\, d^3r} \tag{4.66}$$

From (4.66) it follows that a unitary operator preserves the inner (Hermitian scalar) product of any two states:

$$\int [U\psi_1(\mathbf{r})]^* U\psi_2(\mathbf{r})\, d^3r = \int \psi_1^*(\mathbf{r})\psi_2(\mathbf{r})\, d^3r \tag{4.67}$$

Exercise 4.11. Prove (4.67) from (4.66) by letting $\psi = c_1\psi_1 + c_2\psi_2$, for any choice of the coefficients c_1 and c_2.

By the definition of the adjoint operator $U^\dagger$, Eq. (4.67) can also be written as

$$\int \psi_1^*(\mathbf{r})U^\dagger U\psi_2(\mathbf{r})\, d^3r = \int [U^\dagger U\psi_1(\mathbf{r})]^* \psi_2(\mathbf{r})\, d^3r$$

so that a unitary operator is one that satisfies the relation

$$\boxed{UU^\dagger = U^\dagger U = 1} \tag{4.68}$$

Unitary operators, like self-adjoint (Hermitian) operators, can be assumed to have complete sets of orthonormal eigenfunctions, since for two *eigenfunctions* ψ_1 and ψ_2 of a unitary operator, corresponding to the eigenvalues U_1' and U_2', the equations

$$\int (U\psi_1)^* U\psi_2\, d^3r = U_1'^* U_2' \int \psi_1^*\psi_2\, d^3r \tag{4.69}$$

and

$$(1 - U_1'^* U_2') \int \psi_1^*\psi_2\, d^3r = 0$$

follow. Hence, reasoning analogous to that given for Hermitian operators leads to the conclusion that all eigenvalues of a unitary operator have an absolute value equal to unity:

$$U_j' = e^{i\alpha_j} \qquad (\alpha_j: \text{real}) \tag{4.70}$$

and eigenfunctions belonging to different eigenvalues are orthogonal. Again, as for Hermitian operators, linearly independent eigenfunctions belonging to the same eigenvalue may be chosen to be orthogonal.

An example of a useful unitary operator is the *displacement* or *translation* operator D_ξ, defined by its action on any wave function ψ:

$$\boxed{D_\xi \psi(x) = \psi(x - \xi)} \tag{4.71}$$

where ξ is a real number. Although this definition of the displacement operator is for wave functions in one dimension, the generalization to higher dimensions is immediate. Figure 4.3 illustrates the transformation.

Since $\psi\ (x - \xi)$ is normalized to the same value as $\psi(x)$, the operator D_ξ is unitary. Successive application of two displacement operators D_ξ and D_η leads to

$$D_\eta D_\xi \psi(x) = \psi(x - \xi - \eta) = D_\xi D_\eta \psi(x) = D_{\xi+\eta}\psi(x) \tag{4.72}$$

Hence, displacement operators have the property

$$D_\xi D_\eta = D_\eta D_\xi = D_{\xi+\eta} \tag{4.73}$$

which expresses the fact that the product of two displacements is again a displacement. Such closure under multiplication is the chief characteristic of a *group*. The additional properties defining a group are the existence of an *identity*:

$$D_0 = 1 \tag{4.74}$$

and an *inverse* of every element:

$$D_\xi^{-1} = D_{-\xi} \tag{4.75}$$

such that

$$D_\xi^{-1} D_\xi = D_0 = 1 \tag{4.76}$$

Finally, the multiplication of the elements of a group must be *associative*:

$$D_\zeta (D_\xi D_\eta) = (D_\zeta D_\xi) D_\eta \tag{4.77}$$

which is true for displacement operators, since both sides of this equation are equal to $D_{\zeta+\xi+\eta}$. The translation group is said to be *Abelian* because its elements commute, but it should be noted that this is not a general requirement for groups.

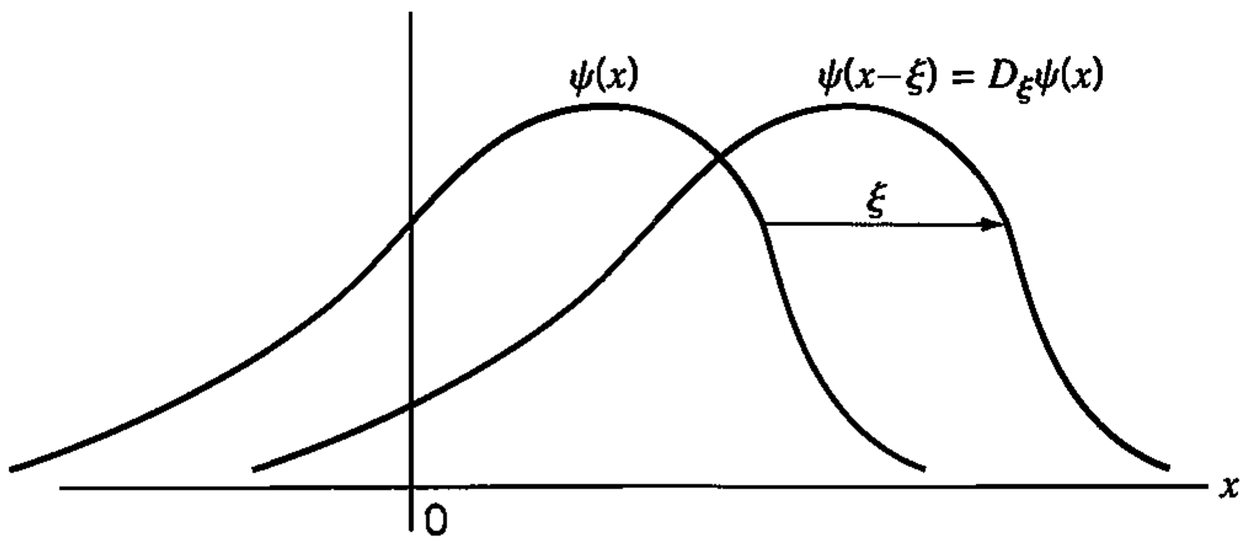


Figure 4.3. Linear displacement or translation of a function $\psi(x) \rightarrow D_\xi\ \psi(x) = \psi(x - \xi)$ by a displacement vector ξ in one dimension. The transformation of the function is *active*. If $\xi = vt$, this figure illustrates an active Galilean transformation (Section 4.7) for a positive value of t.

From the basic property of the displacement operator it is seen that for an arbitrarily small displacement ε, to first order:

$$D_\varepsilon = 1 - i\varepsilon k \tag{4.78}$$

where k must be a *Hermitian* operator to ensure the unitarity of D_ξ.

Exercise 4.12. Prove that if D_ξ is to be unitary, it is necessary that k be Hermitian.

Combining an arbitrary displacement ξ with a small displacement ε, we can write:

$$D_{\xi+\varepsilon} = D_\xi D_\varepsilon = D_\xi(1 - i\varepsilon k) \tag{4.79}$$

from which we infer the differential equation

$$\frac{dD_\xi}{d\xi} = \lim_{\varepsilon\to 0} \frac{D_{\xi+\varepsilon} - D_\xi}{\varepsilon} = -iD_\xi k \tag{4.80}$$

With $D_0 = 1$ for $\xi = 0$, (4.80) has the solution

$$D_\xi = e^{-ik\xi} \tag{4.81}$$

Since

$$D_\xi\psi(x) = \psi(x - \xi) = \psi(x) - \frac{\xi}{1!}\psi'(x) + \frac{\xi^2}{2!}\psi''(x) \pm \ldots = \exp\left(-\xi\frac{\partial}{\partial x}\right)\psi(x) \tag{4.82}$$

it is seen that, in the coordinate representation,

$$k = \frac{1}{i}\frac{\partial}{\partial x} \tag{4.83}$$

which is proportional to the linear momentum operator p_x. The Hermitian operator k is said to be the *generator of infinitesimal translations*.

The eigenvalues and eigenfunctions of the unitary operator D_ξ can be determined from its representation in terms of the Hermitian operator k. For the latter we already know that the eigenvalue equation

$$kf(x) = \frac{1}{i}\frac{\partial f(x)}{\partial x} = k'f(x) \tag{4.84}$$

is satisfied by the eigenfunctions

$$f(x) = e^{ik'x} \tag{4.85}$$

The eigenvalue k' of the operator k may be any real number. It cannot have an imaginary part, because the eigenfunctions would then diverge for $x \to +\infty$ or $x \to -\infty$ and not represent possible physical states.

Since D_ξ is related to the operator k by Eq. (4.81), the eigenvalues of the displacement operator are expressed as

$$D'_\xi = e^{-ik'\xi} \tag{4.86}$$

from which it is seen that the same eigenvalue results if k' is replaced by $k' + 2\pi n/\xi$, where n is any integer. Thus, we can limit ourselves to values of k' in an interval of length $2\pi/\xi$, such as $-\pi/\xi < k' \leq +\pi/\xi$, which is referred to as the *first*

Brillouin zone. The degenerate eigenfunctions of D_ξ corresponding to the eigenvalue $D'_\xi = \exp(-ik'\xi)$ are

$$f_n(x) = e^{i(k'+2\pi n/\xi)x} \qquad (n\text{: any integer})$$

Hence, the linear combination

$$g(x) = \sum_{n=-\infty}^{+\infty} c_n e^{i(k'+2\pi n/\xi)x} = e^{ik'x} \sum_{n=-\infty}^{+\infty} c_n e^{i2\pi n/\xi x} \tag{4.87}$$

is the most general eigenfunction of D_ξ corresponding to the eigenvalue (4.86). The coefficients c_n are arbitrary. The eigenfunctions

$$g(x) = e^{ik'x}u(x) \tag{4.88}$$

are therefore products of plane waves and arbitrary *periodic* functions $u(x)$ of x:

$$u(x - \xi) = u(x) \tag{4.89}$$

with period ξ. Functions that have the structure (4.88) are known as *Bloch functions*.

Exercise 4.13. By requiring that $g(x)$ defined by (4.88) be an eigenfunction of D_ξ, with eigenvalue $D'_\xi = e^{-ik'\xi}$, verify the periodicity property (4.89) directly from the definition of D_ξ as a displacement operator.

Since displacements in three dimensions commute and may be carried out equivalently in any order, it is evident that, in three-dimensional space, Eq. (4.81) is generalized to

$$D_{\boldsymbol{\delta}} = e^{-i\mathbf{k}\cdot\boldsymbol{\delta}} = \exp\{-i(k'_x\xi + k'_y\eta + k'_z\zeta)\} \tag{4.90}$$

where the displacement vector $\boldsymbol{\delta}$ is represented by its three Cartesian components (ξ, η, ζ). The generator of infinitesimal displacements

$$\boxed{\mathbf{k} = \frac{1}{i}\nabla} \tag{4.91}$$

is proportional to the momentum operator $\mathbf{p}$. The generalization of the concepts of Brillouin zones and Bloch functions to three dimensions involves no new ideas. It forms the backbone of every analysis of the physics of periodic lattices, such as crystalline solids.

Exercise 4.14. If $\mathbf{d}_1$, $\mathbf{d}_2$, $\mathbf{d}_3$ are the *primitive* translation vectors of a three-dimensional infinite lattice whose points are at positions $\mathbf{R} = n_1\mathbf{d}_1 + n_2\mathbf{d}_2 + n_3\mathbf{d}_3$ (n_i = *integer*), show that any simultaneous eigenfunction of *all* the translation operators $D_{\mathbf{R}} = e^{-i\mathbf{k}\cdot\mathbf{R}}$ corresponding to eigenvalues $e^{-i\mathbf{k}'\cdot\mathbf{R}}$ is of the form $u(\mathbf{r})e^{i\mathbf{k}'\cdot\mathbf{r}}$, where $u(\mathbf{r})$ is an arbitrary periodic function of $\mathbf{r}$ over the lattice, i.e., $u(\mathbf{r} + \mathbf{R}) = u(\mathbf{r})$, for every $\mathbf{R}$. Note that any two values of $\mathbf{k}'$ which produce the same eigenvalue $e^{-i\mathbf{k}'\cdot\mathbf{R}}$ differ by a *reciprocal lattice vector* $\mathbf{G}$, defined by the condition $\mathbf{G}\cdot\mathbf{R} = 2\pi n$, where n is any integer.

6. ***The Charged Particle in an Electromagnetic Field and Gauge Invariance.*** The Hamiltonian operator considered so far and listed as the total energy in Table 3.1 had the simple form

$$H = \frac{p^2}{2m} + V(\mathbf{r}) \tag{4.92}$$

corresponding to a conservative system with potential energy function V, such as a charged particle in an electrostatic field. But classical Hamiltonian mechanics is not limited to such systems, and neither is quantum mechanics. If a particle of charge q moves in an external *electromagnetic field*, which may be time-dependent, the classical equations of motion may be derived as Hamilton's equations from a Hamiltonian

$$\boxed{H = \frac{1}{2m}\left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)^2 + q\phi} \tag{4.93}$$

Here ϕ and $\mathbf{A}$ are the scalar and vector potentials of the electromagnetic field, respectively, both of which are now assumed to be given functions of $\mathbf{r}$ and t. The validity of this Hamiltonian operator in quantum mechanics is supported by a calculation of the time rate of change of the expectation value of $\mathbf{r}$ for the particle:

$$\frac{d}{dt}\langle \mathbf{r} \rangle = \frac{1}{i\hbar}\langle [\mathbf{r}, H] \rangle = \frac{1}{m}\left\langle \mathbf{p} - \frac{q}{c}\mathbf{A} \right\rangle \tag{4.94}$$

Evidently, this equation defines the *velocity* operator:

$$\boxed{\mathbf{v} = \frac{1}{m}\left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)} \tag{4.95}$$

and this relation between the *canonical* momentum $\mathbf{p}$ and the velocity $\mathbf{v}$, in the presence of a vector potential $\mathbf{A}$, has the same form in quantum mechanics as in classical mechanics.

From (4.95) we infer that the components of the velocity operator generally do not commute and that, since the magnetic field $\mathbf{B} = \nabla \times \mathbf{A}$,

$$\mathbf{v} \times \mathbf{v} = -\frac{q}{m^2 c}(\mathbf{p} \times \mathbf{A} + \mathbf{A} \times \mathbf{p}) = i\hbar \frac{q}{m^2 c}(\nabla \times \mathbf{A}) = i\hbar \frac{q}{m^2 c}\mathbf{B} \tag{4.96}$$

Newton's second law in quantum mechanical form is

$$\frac{d}{dt}\langle \mathbf{v} \rangle = \frac{1}{i\hbar}\langle [\mathbf{v}, H] \rangle + \left\langle \frac{\partial \mathbf{v}}{\partial t} \right\rangle = \frac{1}{i\hbar}\left\langle \left[\mathbf{v}, \frac{m}{2}\mathbf{v} \cdot \mathbf{v}\right] \right\rangle + \frac{1}{i\hbar}\langle [\mathbf{v}, q\phi] \rangle - \frac{q}{mc}\left\langle \frac{\partial \mathbf{A}}{\partial t} \right\rangle$$

The right-hand side can be evaluated by using the operator identity

$$[\mathbf{v}, \mathbf{v} \cdot \mathbf{v}] = \mathbf{v} \times (\mathbf{v} \times \mathbf{v}) - (\mathbf{v} \times \mathbf{v}) \times \mathbf{v} \tag{4.97}$$

Exercise 4.15. Prove the identity (4.97) for *any* operator $\mathbf{v}$ whose components may not commute.

Combining (4.96) and (4.97) we find

$$m\frac{d}{dt}\langle \mathbf{v} \rangle = \frac{q}{2c}\langle \mathbf{v} \times \mathbf{B} - \mathbf{B} \times \mathbf{v} \rangle - q\langle \nabla\phi \rangle - \frac{q}{c}\left\langle \frac{\partial \mathbf{A}}{\partial t} \right\rangle$$

If we recall that the electric field $\mathbf{E}$ is

$$\mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}$$

we arrive at the quantum mechanical version of the equation for the acceleration of the particle in terms of the Lorentz force:

$$\boxed{m\frac{d}{dt}\langle\mathbf{v}\rangle = \frac{q}{2c}\langle\mathbf{v}\times\mathbf{B} - \mathbf{B}\times\mathbf{v}\rangle + q\langle\mathbf{E}\rangle} \tag{4.98}$$

In all of this, the cross products of two vectors require careful treatment because noncommuting factors are involved.

At this stage, it is appropriate to discuss the choice of the electromagnetic potentials in quantum mechanics. Since

$$\nabla\times(\nabla f) = 0 \quad \text{and} \quad \nabla\frac{\partial f}{\partial t} = \frac{\partial}{\partial t}\nabla f$$

the fields **E** and **B** are unchanged by the *gauge* transformation on the potentials:

$$\boxed{\mathbf{A}\to\mathbf{A}' = \mathbf{A} + \nabla f \qquad \phi\to\phi' = \phi - \frac{1}{c}\frac{\partial f}{\partial t}} \tag{4.99}$$

where $f(\mathbf{r}, t)$ is an arbitrary differentiable function of **r** and t. The primed and unprimed potentials are said to represent two different but equally valid *gauges*. In the new primed gauge, the Hamiltonian is

$$H' = \frac{1}{2m}\left(\mathbf{p} - \frac{q}{c}\mathbf{A}'\right)^2 + q\phi' \tag{4.100}$$

The time-dependent Schrödinger equation in the old gauge is

$$i\hbar\frac{\partial\psi(\mathbf{r}, t)}{\partial t} = -\frac{\hbar^2}{2m}\left(\nabla - \frac{iq}{\hbar c}\mathbf{A}\right)^2\psi(\mathbf{r}, t) + q\phi\psi(\mathbf{r}, t) \tag{4.101}$$

and in the new gauge it has the form

$$i\hbar\frac{\partial\psi'(\mathbf{r}, t)}{\partial t} = -\frac{\hbar^2}{2m}\left(\nabla - \frac{iq}{\hbar c}\mathbf{A}'\right)^2\psi'(\mathbf{r}, t) + q\phi'\psi'(\mathbf{r}, t) \tag{4.102}$$

where the new potentials $\mathbf{A}'$ and ϕ' are related to the old potentials by (4.99). The new wave function ψ' is related to the old wave function ψ by the unitary phase transformation,

$$\psi'(\mathbf{r}, t) = U\psi(\mathbf{r}, t) \tag{4.103}$$

with

$$\boxed{U = e^{iqf(\mathbf{r}, t)/\hbar c}} \tag{4.104}$$

If the phase is a constant, we speak of a *global* gauge transformation. Since multiplication by a *constant* phase factor $e^{i\alpha}$ is unimportant in ordinary quantum mechanics where ψ is a probability amplitude, global gauge transformations are not interesting here. In quantum field theory, however, where particles can be created and destroyed, these transformations have significance.

To prove that (4.104) constitutes a local gauge transformation, we merely need to substitute ψ' as given by (4.104) in (4.102) and deduce that this equation is

identical to (4.101) if the Hamiltonians (4.93) and (4.100) in the two gauges are connected by

$$H' = UHU^\dagger + i\hbar \frac{\partial U}{\partial t} U^\dagger \tag{4.105}$$

To check the validity of this condition, we use the definition (4.99) and the commutation relations to verify that the gauge transformation (4.104) leaves the operators **r** and **v** (but *not* **p**) invariant:

$$\begin{aligned} U\mathbf{r}U^\dagger &= \mathbf{r} \\ U\left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)U^\dagger &= \mathbf{p} - \frac{q}{c}\nabla f - \frac{q}{c}\mathbf{A} = \mathbf{p} - \frac{q}{c}\mathbf{A}' \end{aligned} \tag{4.106}$$

and that, further,

$$i\hbar \frac{\partial U}{\partial t} U^\dagger = -\frac{q}{c}\frac{\partial f}{\partial t} \tag{4.107}$$

Exercise 4.16. Carry out the details of the gauge transformation explicitly to show that (4.105) holds with any unitary transformation of the form (4.104).

The representation of the canonical momentum operator **p**, as distinct from the kinetic momentum $m\mathbf{v} = \mathbf{p} - \frac{q}{c}\mathbf{A}$, has been kept the same, $\mathbf{p} = \frac{\hbar}{i}\nabla$, for all choices of the gauge. The fundamental commutation relations are gauge-invariant, and the gauge invariance of the theory finds its expression in the relation

$$\boxed{\psi^*\left(\frac{\hbar}{i}\nabla - \frac{q}{c}\mathbf{A}\right)\psi = \psi^* U^\dagger\left(\frac{\hbar}{i}\nabla - \frac{q}{c}\mathbf{A}'\right)U\psi = \psi'^*\left(\frac{\hbar}{i}\nabla - \frac{q}{c}\mathbf{A}'\right)\psi'} \tag{4.108}$$

This connection ensures the gauge invariance of the expectation value of the kinetic momentum.

Exercise 4.17. Show that the substitution $\nabla \to \nabla - \frac{iq}{\hbar c}\mathbf{A}$ in (3.6) produces a gauge-invariant current density and that this new **j** satisfies the continuity equation

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

for the Schrödinger equation (4.101) in the presence of an electromagnetic field.

Although we have assumed in the foregoing discussion that an electromagnetic field is actually present, gauge invariance considerations apply equally if the particle is free (zero **E** and **B** fields) and the potentials satisfy the conditions

$$\nabla \times \mathbf{A} = 0 \quad \text{and} \quad \nabla\phi + \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} = 0$$

Under these conditions, the trivial gauge $\phi = 0$, $\mathbf{A} = 0$ can be chosen. Other gauges are perfectly acceptable, and the canonical and kinetic momentum do not necessarily

have to be equal. But it is generally assumed that in the absence of electric and magnetic fields, or if the particle is neutral, the choice $\phi = 0$, $\mathbf{A} = 0$ has been made, fixing the most *natural* gauge.

Exercise 4.18. For a charged particle in a uniform static electric field $\mathbf{E}$, show the connection between the two gauges: $\phi = -\mathbf{E} \cdot \mathbf{r}$, $\mathbf{A} = 0$, and $\phi' = 0$, $\mathbf{A}' = -c\mathbf{E}t$. Derive the gauge transformation function $f(\mathbf{r}, t)$ and the two Hamiltonians. Obtain the equations of motion for the expectation values of $\mathbf{r}$ and $\mathbf{p}$ in the two gauges and compare their solutions.

Exercise 4.19. Show that the Schrödinger equation in the presence of a static magnetic field is *not* invariant under time reversal.

The main lesson to be learned from this analysis of gauge invariance is that the wave function for a particle is dependent on the explicit or implied choice of a gauge and that the potentials $\mathbf{A}$ and ϕ must be known if the meaning of ψ is to be properly communicated. In particle physics, the invariance of the quantum mechanical description under multiplication of all wave functions by a spacetime-dependent phase factor $\exp[iq\, f(\mathbf{r}, t)/\hbar c]$ is elevated to a fundamental principle of *local gauge symmetry*. Its imposition necessarily requires that, in addition to a wave function ψ, the state of the system be described by a *gauge field* represented by $\mathbf{A}$ and ϕ. In this view, the electromagnetic field is regarded as a manifestation of local gauge symmetry. The extension of these ideas to particles with additional (internal) degrees of freedom will be sketched in Section 17.10.

7. ***Galilean Transformations and Gauge Invariance.*** The translations considered in Section 4.5 were purely spatial displacements and had nothing to do with time. We now turn to transformations that involve uniform motion with constant velocity $\mathbf{v}$. Using the tools developed in Section 4.5, we attempt to relate the state $\psi(\mathbf{r}, t)$ of a system at time t to the state $\psi_{\mathbf{v}}(\mathbf{r}, t)$ of the same system, after it has been *boosted* so that it is moving with velocity $\mathbf{v}$. According to the *principle of relativity*, the absolute uniform motion of a closed physical system is undetectable. In nonrelativistic physics we speak of the *Galilean principle of relativity*. There is every reason to believe that the principle of relativity is valid in quantum mechanics, as it is in classical physics. If the wave function were itself an observable physical quantity (which it is not), we would expect that the boosted state was just $\psi(\mathbf{r} - \mathbf{v}t, t)$ for all times t. For the simple example of a free particle state represented by a plane wave with momentum $\mathbf{p}$,

$$\psi(\mathbf{r}, t) = e^{i(\mathbf{p}\cdot\mathbf{r} - Et)/\hbar} = e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)} \tag{4.109}$$

This would imply that the boosted plane wave state is

$$\psi(\mathbf{r} - \mathbf{v}t, t) = e^{i[\mathbf{p}\cdot(\mathbf{r} - \mathbf{v}t) - Et]/\hbar} = e^{i[\mathbf{k}\cdot\mathbf{r} - \omega t - \mathbf{k}\cdot\mathbf{v}t]} \tag{4.110}$$

This expression is plainly wrong, since it implies that the momentum and the corresponding de Broglie wavelength are the same for the boosted system as for the system at rest, while the associated frequency has been (Doppler) shifted from ω to $\omega + \mathbf{k} \cdot \mathbf{v}$. Instead, we expect that in the boosted state the momentum is $\mathbf{p} + m\mathbf{v}$ and the energy $E_{\mathbf{v}} = E + \mathbf{p} \cdot \mathbf{v} + mv^2/2$, corresponding to a frequency $\omega + \mathbf{k} \cdot \mathbf{v} + mv^2/2\hbar$, where $|\mathbf{v}| = v$.

To be in compliance with the Galilean principle of relativity and obtain a correct wave function for the boosted state $\psi_{\mathbf{v}}(\mathbf{r}, t)$, we must multiply $\psi(\mathbf{r} - \mathbf{v}t)$ by an appropriate spacetime-dependent phase factor:

$$\boxed{\psi_{\mathbf{v}}(\mathbf{r}, t) = e^{i(m\mathbf{v}\cdot\mathbf{r} - mv^2t/2)/\hbar}\psi(\mathbf{r} - \mathbf{v}t, t)} \tag{4.111}$$

The phase factor in (4.112) depends on **v**, but *not* on **p** and E.

Exercise 4.20. Check that (4.111) gives the correct plane wave function for the boosted state.

More generally, we find that if $\psi(\mathbf{r}, t)$ satisfies the wave equation (4.101):

$$i\hbar \frac{\partial\psi(\mathbf{r}, t)}{\partial t} = -\frac{\hbar^2}{2m}\left[\nabla - \frac{iq}{\hbar c}\mathbf{A}(\mathbf{r}, t)\right]^2\psi(\mathbf{r}, t) + q\phi(\mathbf{r}, t)\psi(\mathbf{r}, t) \tag{4.101}$$

the wave function $\psi(\mathbf{r} - \mathbf{v}t, t)$ satisfies

$$\begin{aligned} i\hbar \frac{\partial\psi(\mathbf{r} - \mathbf{v}t, t)}{\partial t} = &-\frac{\hbar^2}{2m}\left[\nabla - \frac{iq}{\hbar c}\mathbf{A}(\mathbf{r} - \mathbf{v}t, t)\right]^2\psi(\mathbf{r} - \mathbf{v}t, t) \\ &+ q\phi(\mathbf{r} - \mathbf{v}t, t)\psi(\mathbf{r} - \mathbf{v}t, t) - i\hbar\mathbf{v}\cdot\nabla\psi(\mathbf{r} - \mathbf{v}t, t) \end{aligned}$$

where $\mathbf{A}(\mathbf{r} - \mathbf{v}t, t)$ and $\phi(r - \mathbf{v}t, t)$ are the boosted potentials acting on the boosted quantum system. The extra term on the right-hand side shows that this is not the correct wave equation and that $\psi(\mathbf{r} - \mathbf{v}t, t)$ is not the correct wave function for the boosted state. We may rewrite the last equation as

$$\begin{aligned} i\hbar \frac{\partial\psi(\mathbf{r} - \mathbf{v}t, t)}{\partial t} = &-\frac{\hbar^2}{2m}\left[\nabla - \frac{iq}{\hbar c}\mathbf{A}(\mathbf{r} - \mathbf{v}t, t) + \frac{i}{\hbar}m\mathbf{v}\right]^2\psi(\mathbf{r} - \mathbf{v}t, t) \\ &+ \left[q\phi(\mathbf{r} - \mathbf{v}t, t) - \frac{mv^2}{2}\right]\psi(\mathbf{r} - \mathbf{v}t, t) \end{aligned} \tag{4.112}$$

If we invert (4.111):

$$\psi(\mathbf{r} - \mathbf{v}t, t) = e^{-i(m\mathbf{v}\cdot\mathbf{r} - mv^2t/2)/\hbar}\psi_{\mathbf{v}}(\mathbf{r}, t)$$

and substitute this into (4.112), we obtain the correct wave equation for the boosted state:

$$i\hbar \frac{\partial\psi_{\mathbf{v}}(\mathbf{r}, t)}{\partial t} = -\frac{\hbar^2}{2m}\left[\nabla - \frac{iq}{\hbar c}\mathbf{A}(\mathbf{r} - \mathbf{v}t, t)\right]^2\psi_{\mathbf{v}}(\mathbf{r}, t) + q\phi(\mathbf{r} - \mathbf{v}t, t)\psi_{\mathbf{v}}(\mathbf{r}, t) \tag{4.113}$$

The transformation (4.111) describes the relation between *two different* but equivalent states of a system or two replicas of a system. Such a boost is called an *active* Galilean transformation (Figure 4.3). It is to be distinguished from a transformation between two descriptions of the same state of *one* system, viewed from two different frames of reference, which are in relative motion. This transformation is called a *passive* transformation.

To show how a passive transformation works, we start again with the wave function $\psi(\mathbf{r}, t)$ characterizing the given state in frame-of-reference S with coordinates x, y, z, and position vector **r**. We assume that another frame of reference

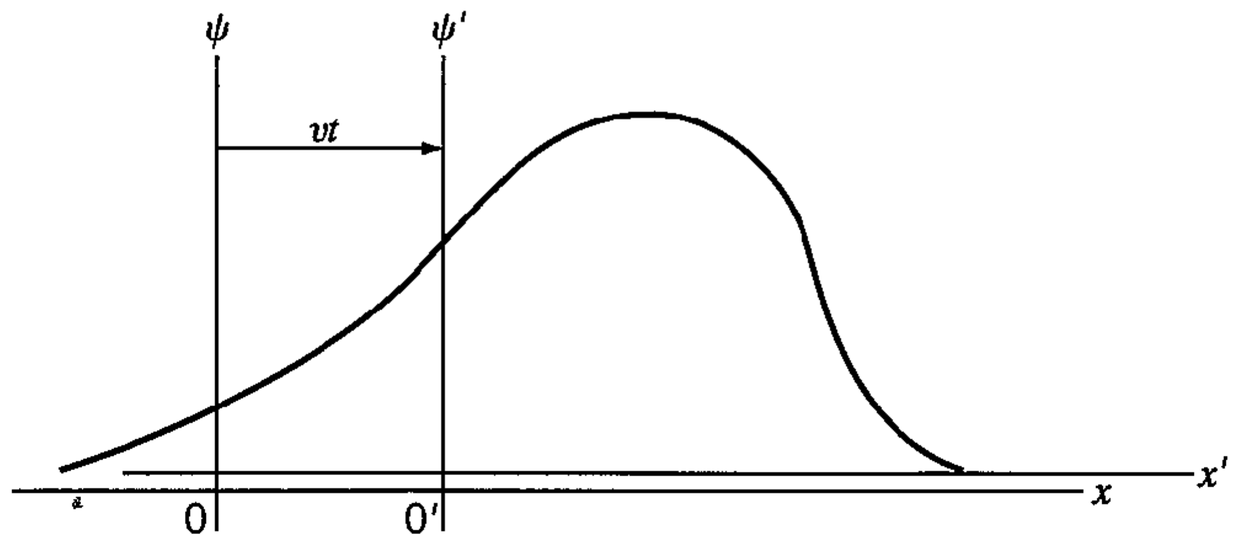


Figure 4.4. Passive Galilean transformation, $x' = x - vt$, of a wave function, effected by the mapping $\psi(x) \rightarrow \psi'(x') = \psi(x' + vt)$. Compare with the active transformation shown in Figure 4.3.

$S'(x', y', z')$ is in uniform translational motion with velocity $\mathbf{v}$ relative to S. A Galilean coordinate transformation (Figure 4.4) is given by

$$\mathbf{r}' = \mathbf{r} - \mathbf{v}t \quad \text{and} \quad t' = t \tag{4.114}$$

Here the second, trivial, equation has been introduced to reduce the likelihood of mathematical errors in carrying out the transformation. It also prepares the way for use of Lorentz rather than Galilean transformations, in which case the time transforms nontrivially. From (4.114) we deduce that

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial t'} - \mathbf{v} \cdot \nabla' \quad \text{and} \quad \nabla = \nabla' \tag{4.115}$$

Hence, if we define the potentials in S' as

$$\mathbf{A}'(\mathbf{r}', t') = \mathbf{A}(\mathbf{r}, t) = \mathbf{A}(\mathbf{r}' + \mathbf{v}t', t') \quad \text{and} \quad \phi'(\mathbf{r}', t') = \phi(\mathbf{r}, t) = \phi(\mathbf{r}' + \mathbf{v}t', t')$$

the wave equation (4.101) in S is transformed into

$$i\hbar \frac{\partial \psi(\mathbf{r}' + \mathbf{v}t', t')}{\partial t'} = -\frac{\hbar^2}{2m}\left[\nabla' - \frac{iq}{\hbar c}\mathbf{A}'(\mathbf{r}', t') - \frac{i}{\hbar}m\mathbf{v}\right]^2 \psi(\mathbf{r}' + \mathbf{v}t', t') + \left[q\phi'(\mathbf{r}', t') - \frac{mv^2}{2}\right]\psi(\mathbf{r}' + \mathbf{v}t', t')$$

Since this equation in the primed frame does not have the same form as (4.101) in the unprimed frame, the Galilean principle of relativity appears to be jeopardized. However, the simple gauge transformation that defines the wave function $\psi'(\mathbf{r}', t')$ in S',

$$\boxed{\psi'(\mathbf{r}', t') = e^{-i(m\mathbf{v}\cdot\mathbf{r} + mv^2t/2)/\hbar}\psi(\mathbf{r}' + \mathbf{v}t', t') = e^{-i(m\mathbf{v}\cdot\mathbf{r} - mv^2t/2)/\hbar}\psi(\mathbf{r}, t)} \tag{4.116}$$

restores the expected manifest Galilean form invariance:

$$i\hbar \frac{\partial \psi'(\mathbf{r}', t')}{\partial t'} = -\frac{\hbar^2}{2m}\left[\nabla' - \frac{iq}{\hbar c}\mathbf{A}'(\mathbf{r}', t')\right]^2 \psi'(\mathbf{r}', t') + q\phi'(\mathbf{r}', t')\psi'(\mathbf{r}', t') \tag{4.117}$$

Nonrelativistic quantum mechanics is thus seen to be consistent with the Galilean principle of relativity owing to the gauge dependence of the wave function.

It is remarkable that if the momentum and energy in the free particle wave function are regarded as components of a relativistic four-vector, the plane wave is

invariant under Lorentz transformations, and there is no additional phase or gauge transformation. In Chapters 23 and 24 we will return to these issues in the context of relativistic quantum mechanics.

Exercise 4.21. If the plane wave (4.109) describes a free particle state in the frame of reference S, show that the passively transformed wave function $\psi'(\mathbf{r}', t')$ correctly accounts for the momentum of the same particle in frame of reference S'. This result implies that the de Broglie wavelength is different as viewed from two frames, which are in relative motion. Compare and contrast the Galilean transformation of ψ waves with the analogous transformation of classical waves. (See also Section 1.2.)

Exercise 4.22. In frame S, consider a standing-wave state ψ that is a superposition of a plane wave (4.109) and a plane wave of equal magnitude and arbitrary phase but opposite momentum, as might be generated by reflection from an impenetrable barrier. Show that the nodes in the wave function ψ are separated by half a wavelength. If the same state is viewed from a moving frame S', show that the two counter-propagating waves have different wavelengths, but that the spacing of the nodes nevertheless is the same as in S.

Problems

1. Show that the addition of an imaginary part to the potential in the quantal wave equation describes the presence of sources or sinks of probability. (Work out the appropriate continuity equation.)

Solve the wave equation for a potential of the form $V = -V_0(1 + i\zeta)$, where V_0 and ζ are positive constants. If $\zeta \ll 1$, show that there are stationary state solutions that represent plane waves with exponentially attenuated amplitude, describing absorption of the waves. Calculate the absorption coefficient.

2. If a particle of mass m is constrained to move in the xy plane on a circular orbit of radius ρ around the origin 0, but is otherwise free, determine the energy eigenvalues and the eigenfunctions.

3. A particle of mass m and charge q is constrained to move in the xy plane on a circular orbit of radius ρ around the origin 0, as in Problem 2, and a magnetic field, represented by the vector potential $\mathbf{A} = \Phi\hat{\mathbf{k}} \times \mathbf{r}/[2\pi(\hat{\mathbf{k}} \times \mathbf{r})^2]$, is imposed.

(a) Show that the magnetic field approximates that of a long thin solenoid with flux Φ placed on the z axis.

(b) Determine the energy spectrum in the presence of the field and show that it coincides with the spectrum for $\Phi = 0$ if the flux assumes certain quantized values. Note that the energy levels depend on the strength of a magnetic field $\mathbf{B}$ which differs from zero only in a region into which the particle cannot penetrate (the *Aharonov-Bohm effect*).

CHAPTER 5

The Linear Harmonic Oscillator

In solving the Schrödinger equation for simple problems, we first assume that ψ and V depend only on the x coordinate. The Hamiltonian for a conservative system in one dimension is

$$H = \frac{p_x^2}{2m} + V(x) \tag{5.1}$$

The eigenvalue equation for this Hamiltonian, in the coordinate representation, is the ordinary Sturm-Liouville differential equation

$$-\frac{\hbar^2}{2m}\frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x) \tag{5.2}$$

We will solve this Schrödinger equation for several special forms of the potential $V(x)$. The harmonic oscillator stands out because in many systems it applies to motion near a stable equilibrium. It governs the dynamics of most continuous physical systems and fields, has equally spaced energy levels, and provides the basis for the description of many-particle states. In this chapter, we discuss the oscillator in the coordinate representation, leaving for later (Chapters 10 and 14) the more general representation-independent treatment.

1. ***Preliminary Remarks.*** For the *linear harmonic oscillator*, the potential energy term in the Schrödinger equation is

$$\boxed{V(x) = \frac{1}{2}m\omega^2x^2} \tag{5.3}$$

We call ω loosely the (classical) frequency of the harmonic oscillator. In the neighborhood of $x = a$, an arbitrary potential $V(x)$ may be expanded in a Taylor series:

$$V(x) = V(a) + V'(a)(x - a) + \frac{1}{2}V''(a)(x - a)^2 + \cdots$$

If $x = a$ is a stable equilibrium position, $V(x)$ has a minimum at $x = a$, i.e., $V'(a) = 0$, and $V''(a) > 0$. If a is chosen as the coordinate origin and $V(a)$ as the zero of the energy scale, then (5.3) is the first approximation to $V(x)$. A familiar example is provided by the oscillations of the atoms in a diatomic molecule, but harmonic oscillations of various generalized coordinates occur in many different systems.

The linear oscillator is important for another reason. The behavior of most continuous physical systems, such as the vibrations of an elastic medium or the electromagnetic field in a cavity, can be described by the superposition of an infinite number of spatial modes, and each expansion coefficient oscillates like a simple harmonic oscillator. In any such physical system, we are then confronted by the quantum mechanics of many linear harmonic oscillators of various frequencies. For

this reason, all quantum field theories make use of the results that we are about to obtain.

Two parameters, the mass m and the frequency ω, appear in the Hamiltonian (5.1) with potential (5.3). They may be used to fix units of mass (m) and time ($1/\omega$). In quantum mechanics, Planck's constant provides a third dimensional number and thus allows us to construct a characteristic energy ($\hbar\omega$) and length ($\sqrt{\hbar/m\omega}$). In the interest of accentuating the great generality of the linear harmonic oscillator, we employ these scaling parameters to replace the canonical coordinate x and momentum p_x by the dimensionless variables

$$\xi = \sqrt{\frac{m\omega}{\hbar}}\, x \tag{5.4}$$

and

$$p_\xi = \sqrt{\frac{1}{\hbar m\omega}}\, p_x \tag{5.5}$$

In terms of the new variables, the Hamiltonian is

$$H = \hbar\omega\left(\frac{1}{2}\, p_\xi^2 + \frac{1}{2}\, \xi^2\right) \tag{5.6}$$

If the Hamiltonian is similarly replaced by the dimensionless quantity

$$H_\xi = \frac{H}{\hbar\omega} = \frac{1}{2}\, p_\xi^2 + \frac{1}{2}\, \xi^2 \tag{5.7}$$

and the time by the dimensionless variable

$$\tau = \omega t \tag{5.8}$$

Hamilton's equations for the system reduce to the usual form

$$\frac{d\xi}{d\tau} = \frac{\partial H_\xi}{\partial p_\xi}; \qquad \frac{dp_\xi}{d\tau} = -\frac{\partial H_\xi}{\partial \xi} \tag{5.9}$$

Exercise 5.1. Show that Eqs. (5.9) are the correct equations of motion for the linear harmonic oscillator defined by (5.1) and (5.3).

The fundamental commutation relation

$$xp_x - p_x x = i\hbar 1 \tag{5.10}$$

is transformed by (5.4) and (5.5) into

$$\xi p_\xi - p_\xi \xi = i1 \tag{5.11}$$

The eigenvalue problem for H_ξ is expressed as

$$H_\xi \psi = \varepsilon\psi \tag{5.12}$$

and the energy E is related to ε by

$$E = \varepsilon\hbar\omega \tag{5.13}$$

The Schrödinger equation for the linear harmonic oscillator is transformed into

$$\frac{d^2\psi}{d\xi^2} + (2\varepsilon - \xi^2)\psi = 0 \tag{5.14}$$

By using the same symbol ψ, in both (5.2) and the transformed equation (5.14) for two different functions for the variables x and ξ, we indulge in an inconsistency that, though perhaps deplorable, is sanctioned by custom.

If a power series solution of this equation is attempted, a *three-term* recursion formula is obtained. To get a differential equation whose power series solution admits a *two-term* recursion relation, which is simpler to analyze, we make the substitution

$$\psi(\xi) = e^{-\xi^2/2}v(\xi) \tag{5.15}$$

This yields the equation

$$\boxed{\frac{d^2v}{d\xi^2} - 2\xi\frac{dv}{d\xi} + 2nv = 0} \tag{5.16}$$

where n is defined by the relation

$$\varepsilon = n + \frac{1}{2} \tag{5.17}$$

or

$$E = \left(n + \frac{1}{2}\right)\hbar\omega \tag{5.18}$$

Exercise 5.2. Substituting a power series with undetermined coefficients for ψ and v into (5.14) and (5.16), obtain the recursion relations and compare these.

2. ***Eigenvalues and Eigenfunctions.*** One simple and important property of the harmonic oscillator derives from the fact that V is an even function of x, if the coordinate origin is chosen suitably. Generally, if the potential energy satisfies the condition $V(-x) = V(x)$, and if $\psi(x)$ is a solution of the Schrödinger equation (5.2), then it follows that $\psi(-x)$ is also a solution of this equation. The Schrödinger equation with even $V(x)$ is said to be *invariant under reflection*, for if x is changed into $-x$, the equation is unaltered except for the replacement of $\psi(x)$ by $\psi(-x)$. Any linear combination of solutions (5.2) also solves (5.2). Hence, if $\psi(x)$ is a solution of the Schrödinger equation, the following two functions must also be solutions:

$$\psi_e(x) = \frac{1}{2}[\psi(x) + \psi(-x)] \qquad \psi_o(x) = \frac{1}{2}[\psi(x) - \psi(-x)]$$

These are the even (e) and odd (o) parts of $\psi(x)$, respectively. Thus, in constructing the solutions of (5.1) for even $V(x)$, we may confine ourselves to all even and all odd solutions. A state which is represented by a wave function that is even, $\psi_e(-x) = \psi_e(x)$, is said to have *even parity*. Similarly, we speak of *odd parity* if $\psi_o(-x) = -\psi_o(x)$.

Exercise 5.3. Extend the notion of invariance under reflection (of all three Cartesian coordinates) to the Schrödinger equation in three dimensions. Show that

if V depends merely on the radial distance r, only solutions of definite parity need be considered.

Since ξ is proportional to x, Eqs. (5.14) and (5.16) are invariant under reflection $\xi \to -\xi$, and we need to look only for solutions of *definite*, even or odd, parity. For a linear second-order differential equation like (5.16), it is easy to construct the even and odd solutions, v_e and v_o, by imposing the boundary conditions:

$$\begin{aligned} &\text{Even case: } v_e(0) = 1, \qquad v_e'(0) = 0 \\ &\text{Odd case: } v_o(0) = 0, \qquad v_o'(0) = 1 \end{aligned}$$

By substituting

$$v_e(\xi) = 1 + a_2\xi^2 + a_4\xi^4 + \cdot$$

or

$$v_o(\xi) = \xi + a_3\xi^3 + a_5\xi^5 + \cdots$$

into (5.16) and equating the coefficient of each power of ξ to zero (see Exercise 5.2), we obtain the power series expansions

$$\begin{aligned} v_e(\xi) &= 1 - \frac{2n}{2!}\xi^2 + \frac{2^2n(n-2)}{4!}\xi^4 - \frac{2^3n(n-2)(n-4)}{6!}\xi^6 + \cdots \\ v_o(\xi) &= \xi - \frac{2(n-1)}{3!}\xi^3 + \frac{2^2(n-1)(n-3)}{5!}\xi^5 - \cdot\cdot \end{aligned} \tag{5.19}$$

The rule that governs these expansions is evident.

In order to determine if such a solution is quadratically integrable and can describe a physical state, we must consider the asymptotic behavior of v_e and v_o. How do these functions behave for large values of $|\xi|$? For the purpose at hand, we need not become involved in complicated estimates. Unless n is an integer both series are infinite, and the ratio of the coefficient of ξ^k to that of ξ^{k-2} is

$$\frac{a_k}{a_{k-2}} = \frac{2(k-n-2)}{k(k-1)} \tag{5.20}$$

Here $k = 2, 4, 6, \ldots$ for v_e and $k = 1, 3, 5, \ldots$ for v_o. For a fixed value of n, other than $n =$ even for v_e and $n =$ odd for v_o, all terms with $k > n + 2$ have the same sign. Furthermore, it is easy to verify that if $k > C(n + 2)$, where C is an arbitrary constant $C > 1$, the ratio (5.20) is always greater than the corresponding ratio of coefficients in the power series expansion of $\exp[(1 - 1/C)\xi^2]$ or $\xi \exp[(1 - 1/C)\xi^2]$, depending on whether k is even or odd. If C is chosen to be greater than 2, it follows from (5.15) that for large values of $|\xi|$ the eigenfunction $\psi(\xi)$ *diverges* faster than $\exp[(1/2 - 1/C)\xi^2]$ or $\xi \exp[(1/2 - 1/C)\xi^2]$, respectively.

Hence, unless n is an integer, $\psi(\xi)$ diverges as $|\xi| \to \infty$. Such wave functions are physically not useful, because they describe a situation in which it is overwhelmingly probable that the particle is not "here" but at infinity. This behavior can be avoided only if n is an integer, in which case one of the two series (5.19) terminates and becomes a polynomial of degree n.

If n is even, we get

$$\psi(\xi) = e^{-\xi^2/2}\, v_e(\xi) \tag{5.21}$$

and the state has even parity. If n is odd, we get

$$\psi(\xi) = e^{-\xi^2/2}\, v_o(\xi) \tag{5.22}$$

and the state has odd parity. Both (5.21) and (5.22) are now finite everywhere and quadratically integrable.

We have come to a very important conclusion: The number n must be a nonnegative integer ($n = 0, 1, 2, \ldots$); hence, E can assume only the discrete values (5.18)

$$\boxed{E_n = \hbar\omega\left(n + \frac{1}{2}\right) \quad \text{or} \quad E = \frac{1}{2}\hbar\omega, \frac{3}{2}\hbar\omega, \frac{5}{2}\hbar\omega, \ldots} \tag{5.23}$$

Classically, all nonnegative numbers are allowed for the energy of a harmonic oscillator. In quantum mechanics, a stationary state of the harmonic oscillator can have only one of a discrete set of allowed energies! The energies are thus indeed quantized, and we may speak of a spectrum of *discrete energy levels* for the harmonic oscillator. The equally spaced energy levels of the harmonic oscillator, familiar to all students of physics since the earliest days of quantum theory, are sometimes referred to as the steps or rungs of a ladder.

The numbers E_n are the *eigenvalues* of the Schrödinger equation

$$\boxed{-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + \frac{1}{2}m\omega^2x^2\psi = E\psi} \tag{5.24}$$

and of the Hamiltonian operator

$$\boxed{H = \frac{p_x^2}{2m} + \frac{1}{2}m\omega^2x^2} \tag{5.25}$$

The corresponding solutions are the *eigenfunctions*. The eigenfunctions (5.21) and (5.22) decay rapidly at large distances, so that the particle is confined to the neighborhood of the center of force. The states described by such eigenfunctions are said to be *bound*. Note that the eigenvalues of the linear harmonic oscillator are not degenerate, since for each eigenvalue there exists only one eigenfunction, apart from an arbitrary constant factor. This property of the one-dimensional Schrödinger equation follows from the discussion of Section 3.5 [case (b)]. That the eigenvalues are equidistant on the energy scale is one of the important peculiar features of the x^2 dependence of the oscillator potential.

The nonnegative integers n are the eigenvalues of the operator

$$H_\xi - \frac{1}{2} = \frac{1}{2}p_\xi^2 + \frac{1}{2}\xi^2 - \frac{1}{2} \tag{5.26}$$

as defined in (5.7). Because its eigenvalues are $n = 0, 1, 2, 3, \ldots$, the operator $H_\xi - 1/2$ is called the *number operator*. It measures how many steps of the energy level ladder of the harmonic oscillator have been excited. Depending on the physical context to which the harmonic oscillator formalism is applied, the state of excitation n is said to represent the presence of n quanta or particles. An alternative and very

general derivation of the eigenvalues and eigenfunctions (or better, eigenvectors or eigenstates) of the number operator, which makes use only of the algebraic properties of the operators ξ and p_ξ and their commutator (5.11), will be found in Section 10.6.

As a label of the eigenvalues and eigenfunctions of the ordinary linear harmonic oscillator, the number n is traditionally and blandly called a *quantum number*. Its minimum value, $n = 0$, corresponds to the *ground state*, but the energy of the oscillator is still $\hbar\omega/2$ and does not vanish as the lowest possible classical energy would. Since in a stationary state, $\langle H \rangle = E$, a zero energy eigenvalue would imply that both $\langle p_x^2 \rangle = 0$ and $\langle x^2 \rangle = 0$. As in Section 4.1, it would follow that both $p_x\psi = 0$ and $x\psi = 0$. But no ψ exists that satisfies these two equations simultaneously. The energy $\hbar\omega/2$ is called the *zero-point energy* of the harmonic oscillator. Being proportional to $\hbar$, it is obviously a quantum phenomenon; it can be understood on the basis of the uncertainty principle (Chapter 10).

3. ***Study of the Eigenfunctions.*** In this section, a few important mathematical properties of the harmonic oscillator eigenfunctions will be derived.

The finite polynomial solutions of (5.16) which can be constructed if n is an integer are known as *Hermite polynomials* of degree n. The complete eigenfunctions are of the form

$$\psi_n(x) = C_n H_n\left(\sqrt{\frac{m\omega}{\hbar}}\, x\right) \exp\left(-\frac{m\omega}{2\hbar}\, x^2\right) \tag{5.27}$$

where H_n denotes a Hermite polynomial of degree n, and C_n is an as yet undetermined normalization constant. But first the Hermite polynomials themselves must be normalized. It is traditional to define them so that the highest power of ξ appears with the coefficient 2^n. Hence, by comparing with (5.19), we see that for even n,

$$H_n(\xi) = (-1)^{n/2} \frac{n!}{(n/2)!}\, v_e(\xi) \tag{5.28}$$

and for odd n,

$$H_n(\xi) = (-1)^{(n-1)/2} \frac{2(n!)}{[(n-1)/2]!}\, v_o(\xi) \tag{5.29}$$

Here is a list of the first few Hermite polynomials:

$$\begin{array}{ll} H_0(\xi) = 1 & H_3(\xi) = -12\xi + 8\xi^3 \\ H_1(\xi) = 2\xi & H_4(\xi) = 12 - 48\xi^2 + 16\xi^4 \\ H_2(\xi) = -2 + 4\xi^2 & H_5(\xi) = 120\xi - 160\xi^3 + 32\xi^5 \end{array} \tag{5.30}$$

They satisfy the differential equation

$$\frac{d^2H_n(\xi)}{d\xi^2} - 2\xi \frac{dH_n(\xi)}{d\xi} + 2nH_n(\xi) = 0 \tag{5.31}$$

The first few harmonic oscillator eigenfunctions are plotted in Figure 5.1 and 5.2.

Exercise 5.4. Prove that

$$\frac{dH_n(\xi)}{d\xi} = 2nH_{n-1}(\xi) \tag{5.32}$$

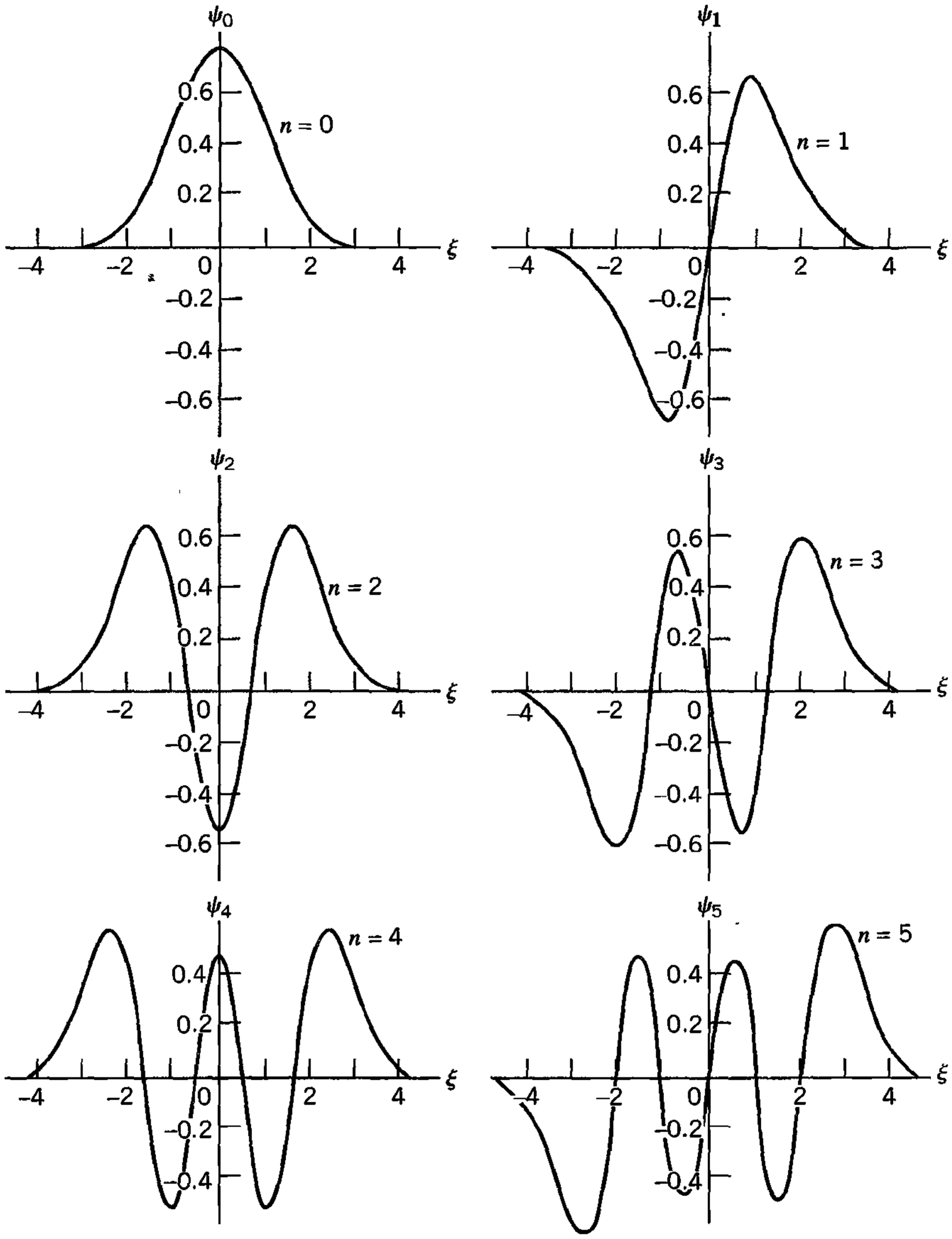


Figure 5.1. The energy eigenfunctions of the linear harmonic oscillator, for the quantum numbers $n = 0$ to 5. The functions ψ_n are plotted versus $\xi = \sqrt{m\omega/\hbar}\, x$ and are normalized as $\int |\psi_n(\xi)|^2 \, d\xi = 1$. The vertical axis is the energy in units of $\hbar\omega$.

A particularly simple representation of the Hermite polynomials is obtained by constructing the *generating function*

$$F(\xi, s) = \sum_{n=0}^{\infty} \frac{H_n(\xi)}{n!} s^n \tag{5.33}$$

As a consequence of the relation (5.32), we see that

$$\frac{\partial F(\xi, s)}{\partial \xi} = 2s\, F(\xi, s)$$

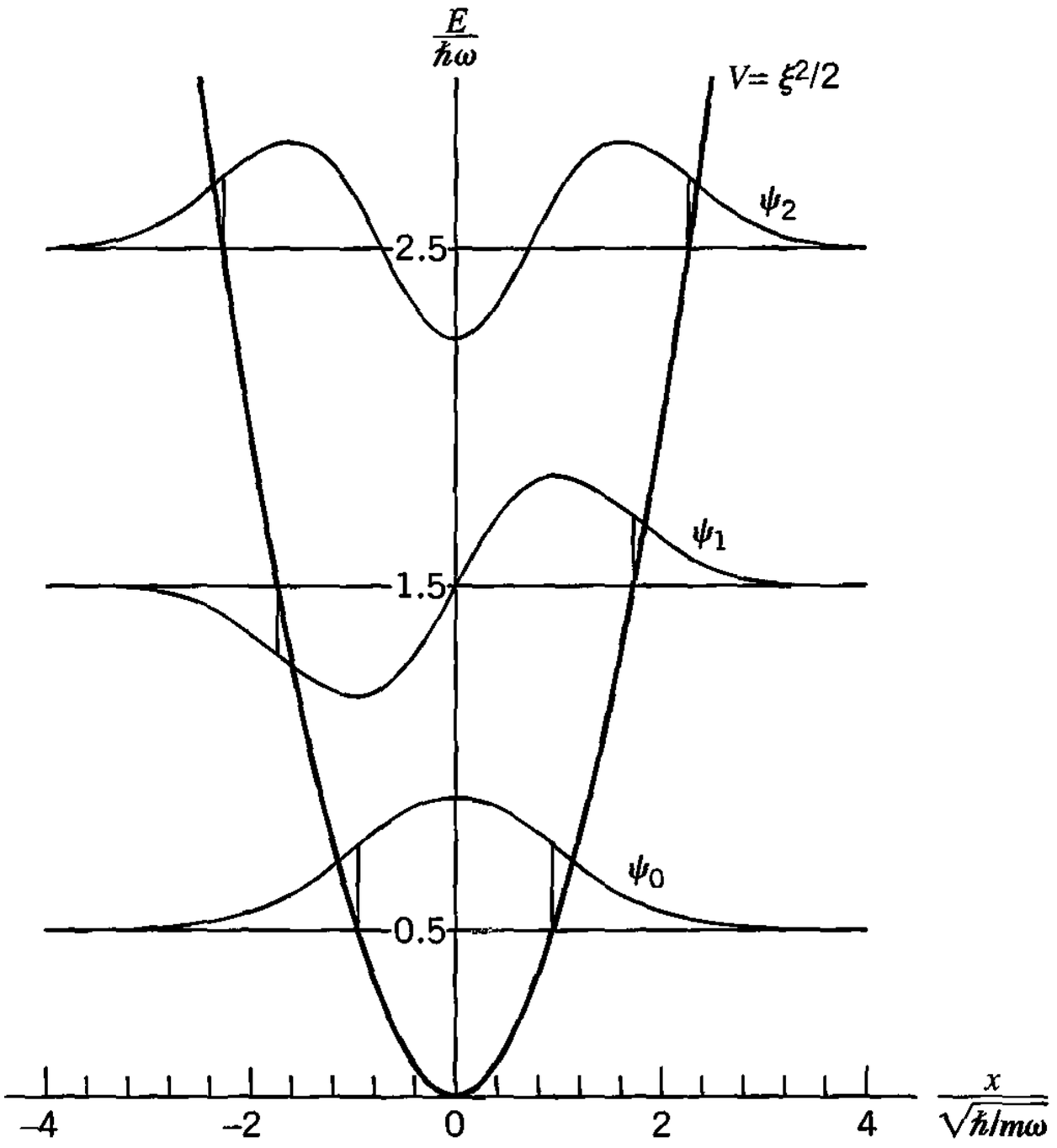


Figure 5.2. The three lowest energy eigenfunctions ($n = 0, 1, 2$) for the linear harmonic oscillator in relation to the potential energy, $V(x) = m\omega^2 x^2/2 = \hbar\omega\xi^2/2$. The intercepts of the parabola with the horizontal lines are the positions of the classical turning points.

This differential equation can be integrated:

$$F(\xi, s) = F(0, s)\, e^{2s\xi}$$

The coefficient $F(0, s)$ can be evaluated from (5.33) and (5.28):

$$F(0, s) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} s^{2n} = e^{-s^2}$$

and therefore the generating function has the form

$$F(\xi, s) = e^{-s^2+2s\xi} = e^{\xi^2-(s-\xi)^2} \tag{5.34}$$

The generating function $F(\xi, s)$ is useful because it allows us to deduce a number of simple properties of the harmonic oscillator wave functions with ease. For example, by Taylor's expansion of (5.34):

$$H_n(\xi) = \left(\frac{d^n}{ds^n} e^{\xi^2-(s-\xi)^2}\right)_{s=0} = (-1)^n e^{\xi^2} \frac{d^n}{d\xi^n} e^{-\xi^2} \tag{5.35}$$

a popular alternative form of definition of the Hermite polynomials.

From this definition it follows that all n roots of $H_n(\xi)$ must be real.

Proof. Assume that $\dfrac{d^{n-1}e^{-\xi^2}}{d\xi^{n-1}}$ has $n-1$ real roots. Since $e^{-\xi^2}$ and all its derivatives tend to zero as $\xi \to \pm\infty$, the derivative $\dfrac{d^n e^{-\xi^2}}{d\xi^n}$ must have at least n real roots. Being

$e^{-\xi^2}$ times a polynomial of degree n, it can have no more than n such roots. The assumption holds for $n = 1$, whence the assertion follows by induction.

The points in coordinate space at which a wave function goes through zero are called its *nodes*. Evidently, the oscillator eigenfunction (5.27) has n nodes.

From the generating function we derive conveniently the value of the useful integral

$$I_{nkp} = \int_{-\infty}^{+\infty} H_n(\xi)H_k(\xi)e^{-\xi^2}\xi^p \, d\xi \tag{5.36}$$

for nonnegative integers n, k, p. To this end we construct the expression

$$\int_{-\infty}^{+\infty} e^{\xi^2-(s-\xi)^2}e^{\xi^2-(t-\xi)^2}e^{2\lambda\xi-\xi^2} \, d\xi = \sum_{n=0}^{\infty}\sum_{k=0}^{\infty}\frac{s^n t^k}{n!k!}\int_{-\infty}^{+\infty} H_n(\xi)H_k(\xi)e^{2\lambda\xi-\xi^2} \, d\xi$$

where (5.33) and (5.34) have been used. The left-hand side can be integrated explicitly; it equals

$$e^{2st+\lambda^2+2\lambda(s+t)}\int_{-\infty}^{+\infty} e^{-(s+t+\lambda-\xi)^2} \, d\xi = \sqrt{\pi}\, e^{\lambda^2+2(st+\lambda s+\lambda t)}$$

Hence,

$$\boxed{\sum_{p=0}^{\infty}\sum_{n=0}^{\infty}\sum_{k=0}^{\infty} I_{nkp}\frac{s^n}{n!}\frac{t^k}{k!}\frac{(2\lambda)^p}{p!} = \sqrt{\pi}\, e^{\lambda^2+2(st+\lambda s+\lambda t)}} \tag{5.37}$$

Comparing the coefficients of equal powers of $s^n t^k \lambda^p$, we obtain the value of I_{nkp}.

In particular, for $p = 0$, we verify that the oscillator energy eigenfunctions for $n \neq k$ are orthogonal, as expected. For $n = k$, we obtain the integral

$$I_{nn0} = \int_{-\infty}^{+\infty} [H_n(\xi)]^2 e^{-\xi^2} \, d\xi = 2^n \, n!\sqrt{\pi}$$

In terms of the variable ξ, the orthonormality of the eigenfunctions is expressed as

$$\frac{1}{2^n \, n!\sqrt{\pi}}\int_{-\infty}^{+\infty} H_n(\xi)H_k(\xi)e^{-\xi^2} \, d\xi = \delta_{nk} \tag{5.38}$$

If we recall that $\xi = \sqrt{\dfrac{m\omega}{\hbar}}\, x$, we have for the normalized eigenfunctions

$$\boxed{\psi_n(x) = 2^{-n/2}(n!)^{-1/2}\left(\frac{m\omega}{\hbar\pi}\right)^{1/4}\exp\left(-\frac{m\omega}{2\hbar}x^2\right)H_n\left(\sqrt{\frac{m\omega}{\hbar}}\, x\right)} \tag{5.39}$$

and the orthonormality relation

$$\int_{-\infty}^{+\infty} \psi_n^*(x)\psi_k(x) \, dx = \delta_{nk} \tag{5.40}$$

For the sake of generality, complex conjugation appears in (5.40), although with the particular choice of the arbitrary phase factor embodied in (5.39) the eigenfunctions are real.

In many calculations, we will need to know the *matrix elements* of the operator x^p:

$$\boxed{\langle n|x^p|k\rangle = \int_{-\infty}^{+\infty} \psi_n^*(x)x^p\psi_k(x)\,dx} \tag{5.41}$$

the "bra-ket" on the left-hand side of this equation is introduced here merely as a notational shorthand. Its name and its significance will be explained in due course. For use in the next section, we record the value of (5.41) for $p = 1$:

$$\boxed{\langle n|x|k\rangle \equiv \int_{-\infty}^{+\infty} \psi_n^*(x)x\psi_k(x)\,dx = \sqrt{\frac{\hbar}{m\omega}}\left[\sqrt{\frac{n}{2}}\,\delta_{k,n-1} + \sqrt{\frac{n+1}{2}}\,\delta_{k,n+1}\right]} \tag{5.42}$$

Exercise 5.5. From (5.36) and (5.37), work out the matrix elements $\langle n|x|k\rangle$ and $\langle n|x^2|k\rangle$ for the harmonic oscillator.

Integral representations of the special functions that we encounter in quantum mechanics are often very useful. For the Hermite polynomials the integral representation

$$H_n(\xi) = \frac{2^n}{\sqrt{\pi}}\int_{-\infty}^{+\infty} (\xi + iu)^n e^{-u^2}\,du \tag{5.43}$$

is valid.

Exercise 5.6. Validate (5.43) by verifying that it satisfies (5.32) and by checking the initial values $H_n(0)$. Alternatively, show that (5.43) can be used to verify the formula (5.33) for the generating function of Hermite polynomials.

If we let $s = \eta + iu$ in the generating function (5.33), multiply the equation by e^{-u^2}, and integrate over u from $-\infty$ to $+\infty$, we obtain from (5.43)

$$\int_{-\infty}^{+\infty} F(\xi, \eta + iu)e^{-u^2}\,du = \sqrt{\pi}\sum_{n=0}^{\infty}\frac{H_n(\xi)}{n!}\frac{H_n(\eta)}{2^n}$$

and by (5.34) this expression equals

$$\int_{-\infty}^{+\infty} e^{-(\eta+iu)^2+2\xi(\eta+iu)-u^2}\,du = \pi e^{-\eta^2+2\xi\eta}\delta(\xi - \eta)$$

Consequently,

$$\sum_{n=0}^{\infty}\frac{1}{2^n\,n!\sqrt{\pi}}H_n(\xi)e^{-\xi^2/2}H_n(\eta)e^{-\eta^2/2} = \delta(\xi - \eta) \tag{5.44}$$

Translated into the x representation, this is the closure relation (4.42):

$$\sum_{n=0}^{\infty}\psi_n^*(x')\psi_n(x) = \delta(x - x') \tag{5.45}$$

This relation shows that the harmonic-oscillator eigenfunctions constitute a complete set of orthonormal functions in terms of which an arbitrary function of x can be expanded. The fundamental expansion postulate of quantum mechanics is thus shown to be valid for the energy eigenfunctions of the harmonic oscillator.

Exercise 5.7. In the generating function (5.33), replace s by $s = t(\eta + iu)$ and prove *Mehler's* formula

$$\sum_{n=0}^{\infty} H_n(\xi)H_n(\eta)\frac{t^n}{2^n\, n!} = \frac{1}{\sqrt{1-t^2}}\exp\left[\frac{2\xi\eta - t^2(\xi^2+\eta^2)}{1-t^2}\right] \tag{5.46}$$

In this section we have derived a number of mathematical results pertaining to the solution of the Schrödinger equation for the linear harmonic oscillator. Although the physical significance of some of these formulas will become apparent only later, it seemed efficient to compile them here in one place.

4. ***The Motion of Wave Packets.*** So far we have considered only the stationary states of the harmonic oscillator. We now turn our attention to the behavior of a general wave $\psi(x, t)$ whose initial form $\psi(x, 0)$ is given. The time-dependent Schrödinger or wave equation

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} + \frac{m\omega^2}{2}x^2\psi(x,t) \tag{5.47}$$

determines the time development of the wave. In Chapters 3 and 4 we saw that the solution of this equation can be obtained automatically, if the initial wave can be expanded in terms of the time-dependent eigenfunctions of the corresponding (time-independent) Schrödinger equation. The completeness of the orthonormal energy eigenfunctions of the harmonic oscillator was proved in the last section. Accordingly, if $\psi(x, 0)$, which we assume to be normalized to unity, is expanded as

$$\psi(x, 0) = \sum_{n=0}^{\infty} c_n\psi_n(x) \tag{5.48}$$

with the expansion coefficients given by

$$c_n = \int_{-\infty}^{+\infty}\psi_n^*(x)\psi(x, 0)\, dx \tag{5.49}$$

then, knowing that for the harmonic oscillator, $E_n = \hbar\omega\left(n + \frac{1}{2}\right)$, we can construct the wave packet at time t by the use of (3.79):

$$\psi(x, t) = \exp\left(-\frac{i\omega t}{2}\right)\sum_{n=0}^{\infty} c_n\psi_n(x)\exp(-in\omega t) \tag{5.50}$$

The center of probability of the normalized wave packet, i.e., the expectation value of the position operator x, is according to (3.15)

$$\langle x\rangle_t = \int_{-\infty}^{+\infty} x|\psi(x, t)|^2\, dx \tag{5.51}$$

Substituting (5.50) into (5.51), we find

$$\begin{aligned}\langle x\rangle_t &= \sum_{n=0}^{\infty}\sum_{k=0}^{\infty} c_n^* c_k e^{i(n-k)\omega t}\int_{-\infty}^{+\infty}\psi_n^*(x)x\psi_k(x)\, dx \\ &= \sum_{n=0}^{\infty}\sum_{k=0}^{\infty} c_n^* c_k e^{i(n-k)\omega t}\langle n|x|k\rangle\end{aligned} \tag{5.52}$$

which shows that the matrix elements of x, obtained in (5.42), enter critically. If (5.42) is substituted in (5.52), the *selection rule* for the matrix elements $\langle n|x|k\rangle$,

$$n - k = \pm 1 \tag{5.53}$$

simplifies the summations and gives the result

$$\langle x\rangle_t = \sqrt{\frac{\hbar}{2m\omega}} \sum_{n=1}^{\infty} \sqrt{n}(c_n^* c_{n-1} e^{i\omega t} + c_{n-1}^* c_n e^{-i\omega t}) \tag{5.54}$$

If we set

$$c_n = |c_n| e^{i\phi_n}$$

we can write

$$\langle x\rangle_t = \sqrt{\frac{2\hbar}{m\omega}} \sum_{n=0}^{\infty} \sqrt{n}|c_n||c_{n-1}| \cos(\omega t + \phi_{n-1} - \phi_n) \tag{5.55}$$

This expression is *exact*. It shows that the expectation value of the coordinate, $\langle x\rangle_t$, oscillates harmonically with frequency ω, just like the classic coordinate $x(t)$.

Exercise 5.8. Use (5.32), (5.40), and (5.42) to calculate the matrix elements of the momentum operator,

$$\langle n|p_x|k\rangle = \int_{-\infty}^{+\infty} \psi_n^*(x) \frac{\hbar}{i} \frac{d\psi_k(x)}{dx}\, dx$$

With this result, evaluate $\langle p_x\rangle_t$ as a function of time for the wave packet (5.50). Verify that $\langle p_x\rangle_t = m \dfrac{d}{dt}\langle x\rangle_t$ for this wave packet.

Exercise 5.9. Verify that (5.55) is expressible as

$$\langle x\rangle_t = \langle x\rangle_0 \cos\,\omega t + \frac{\langle p_x\rangle_0}{m\omega} \sin\,\omega t$$

which can also be derived directly from the equation of motion for $\langle x\rangle_t$ (see Problem 3 in Chapter 3).

Problems

1. Calculate the matrix elements of p_x^2 with respect to the energy eigenfunctions of the harmonic oscillator and write down the first few rows and columns of the matrix. Can the same result be obtained directly by matrix algebra from a knowledge of the matrix elements of p_x?

2. Calculate the expectation values of the potential and kinetic energies in any stationary state of the harmonic oscillator. Compare with the results of the virial theorem.

3. Calculate the expectation value of x^4 for the n-th energy eigenstate of the harmonic oscillator.

4. For the energy eigenstates with $n = 0$, 1, and 2, compute the probability that the coordinate of a linear harmonic oscillator in its ground state has a value greater than the amplitude of a classical oscillator of the same energy.

5. Show that if an ensemble of linear harmonic oscillators is in thermal equilibrium, governed by the Boltzmann distribution, the probability per unit length of finding a particle with displacement of x is a Gaussian distribution. Plot the width of the distribution as a function of temperature. Check the results in the classical and the low-temperature limits. [*Hint:* Equation (5.43) may be used.]

6. Use the generating function for the Hermite polynomials to obtain the energy eigenfunction expansion of an initial wave function that has the same form as the oscillator ground state but that is centered at the coordinate a rather than the coordinate origin:

$$\psi(x, 0) = \left(\frac{m\omega}{\hbar\pi}\right)^{1/4} \exp\left(-\frac{m\omega(x - a)^2}{2\hbar}\right).$$

(a) For this initial wave function, calculate the probability P_n that the system is found to be in the n-th harmonic oscillator eigenstate, and check that the P_n add up to unity.

(b) Plot P_n for three typical values of a, illustrating the case where a is less than, greater than, and equal to $\sqrt{\frac{\hbar}{m\omega}}$.

(c) If the particle moves in the field of the oscillator potential with angular frequency ω centered at the coordinate origin, again using the generating function derive a closed-form expression for $\psi(x, t)$.

(d) Calculate the probability density $|\psi(x, t)|^2$ and interpret the result.

CHAPTER 6

Sectionally Constant Potentials in One Dimension

Potentials like the rectangular barrier or the square well, which are pieced together from constant zero-force sections with sharp discontinuities, do not occur in nature but serve as convenient models. Classically, they are trivial, but here they are useful to exemplify characteristic quantum properties that arise from the smooth joining of the Schrödinger wave function (Section 3.5) at the discontinuities of the potential, such as tunneling and scattering resonances. The mathematics is relatively simple, so that we can concentrate on the physical features, especially the power of symmetry considerations.

1. ***The Potential Step.*** Of all Schrödinger equations, the one for a constant potential is mathematically the simplest. We know from Chapter 2 that the solutions are harmonic plane waves, with wave number

$$k = \sqrt{2m(E - V)}$$

We resume study of the Schrödinger equation with such a potential because the qualitative features of a real physical potential can often be approximated reasonably well by a potential that is made up of a number of constant portions. For instance, unlike the electrostatic forces that hold an atom together, the strong nuclear forces acting between protons and neutrons have a short range; they extend to some distance and then drop to zero very fast. Figure 6.1 shows roughly how a rectangular potential well—commonly called a *square well*—might simulate the properties of such an interaction.

Often, such a schematic potential approximates the real situation and provides a rough orientation with comparatively little mathematical work. As we will see in Section 8.7, a sectionally constant periodic potential exhibits some of the important features of any periodic potential seen by an electron in a crystal lattice.

The case of the free particle, which sees a constant potential $V(x) = \text{const.}$, for all x, in three dimensions as well as one dimension with and without periodic boundary conditions, was already discussed in Section 4.4.

Next in order of increasing complexity is the *potential step* $V(x) = V_0\eta(x)$ as shown in Figure 6.2. There is no physically acceptable solution for $E < 0$ because of the general theorem that E can never be less than the absolute minimum of $V(x)$. Classically, this is obvious. But as the examples of the harmonic oscillator and the free particle have already shown us, it is also true in quantum mechanics despite the possibility of penetration into classically inaccessible regions. We can prove the theorem by considering the real solutions of Schrödinger's equation (see Exercise 3.20):

$$-\frac{\hbar^2}{2m}\psi''(x) + [V(x) - E]\psi(x) = 0$$

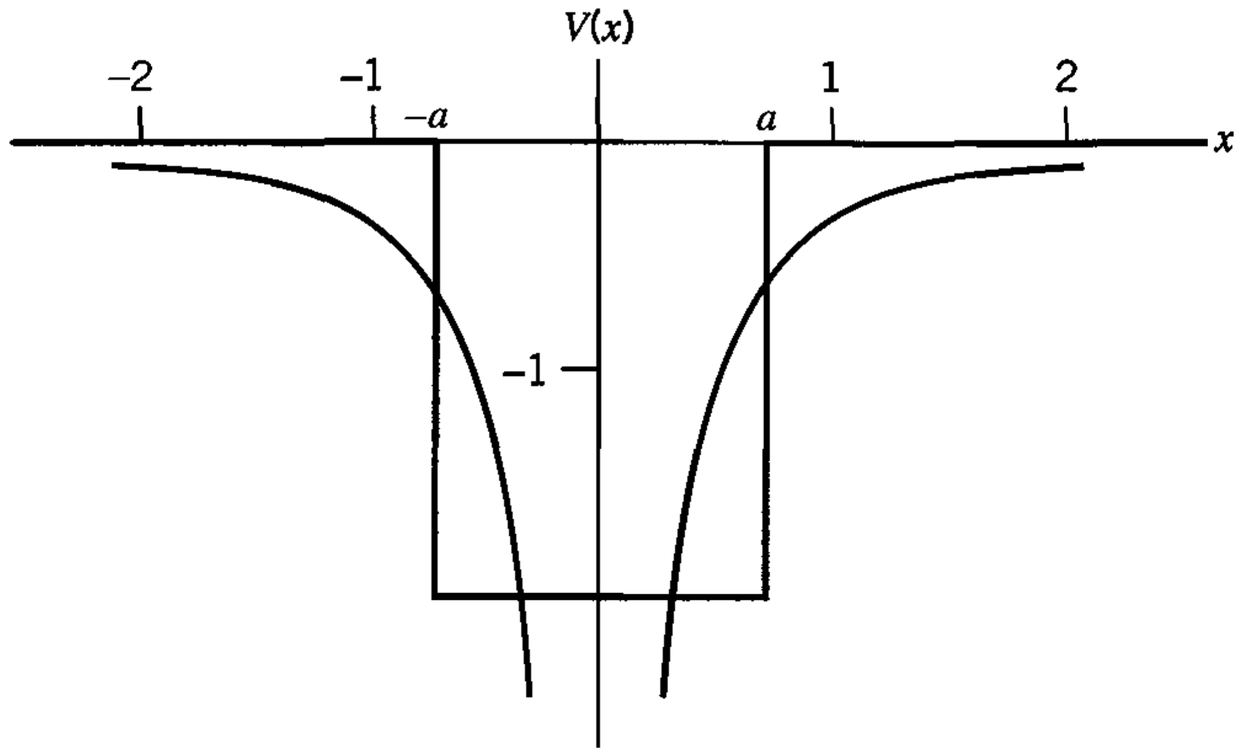


Figure 6.1. Potential approximating the attractive part of nuclear forces ($V = -e^{-|x|}/|x|$) and a one-dimensional square well simulating it.

If $V(x) > E$ for all x, ψ'' has the same sign as ψ, everywhere. Hence, if ψ is positive at some point x, the wave function has one of the two convex shapes shown in Figure 6.3, depending on whether the slope is positive or negative. In Figure 6.3*a*, ψ can never bend down to be finite as $x \to +\infty$. In Figure 6.3*b*, ψ diverges as $x \to -\infty$. To avoid these catastrophes, there must always be some region where $E > V(x)$ and where the particle can be found classically.

Exercise 6.1. Prove that E must exceed the absolute minimum value of the potential $V(x)$ by noting that $E = \langle H \rangle$ in the stationary state $\psi_E(x)$.

Now we consider the potential step with $0 < E < V_0$. Classically, a particle of this energy, if it were incident from the left, would move freely until reflected at the potential step. Conservation of energy requires it to turn around, changing the sign of its momentum.

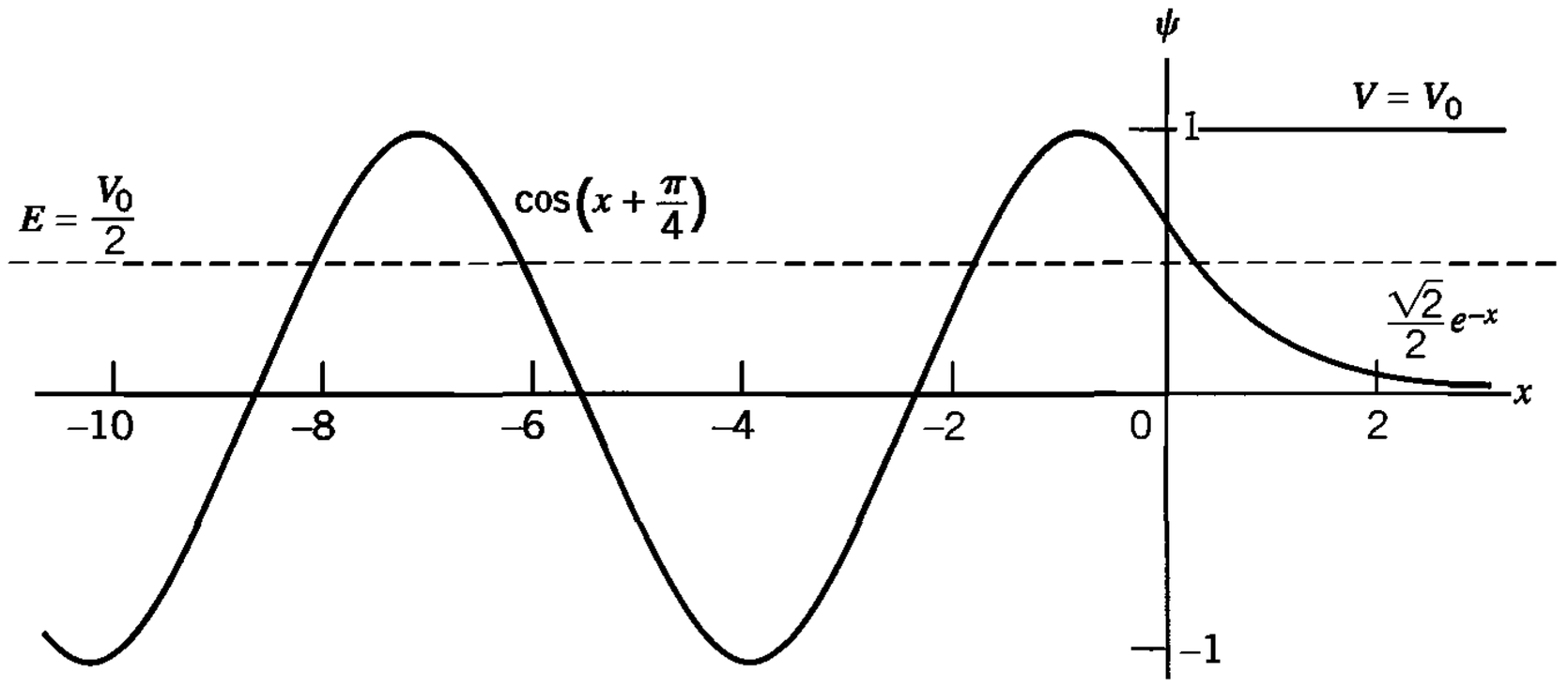


Figure 6.2. Energy eigenfunction for the (Heaviside) step potential function $V(x) = V_0\,\eta(x)$, corresponding to an energy $E = V_0/2$. The step function $\eta(x)$ is defined in Section 1 of the Appendix. The normalization is arbitrary.

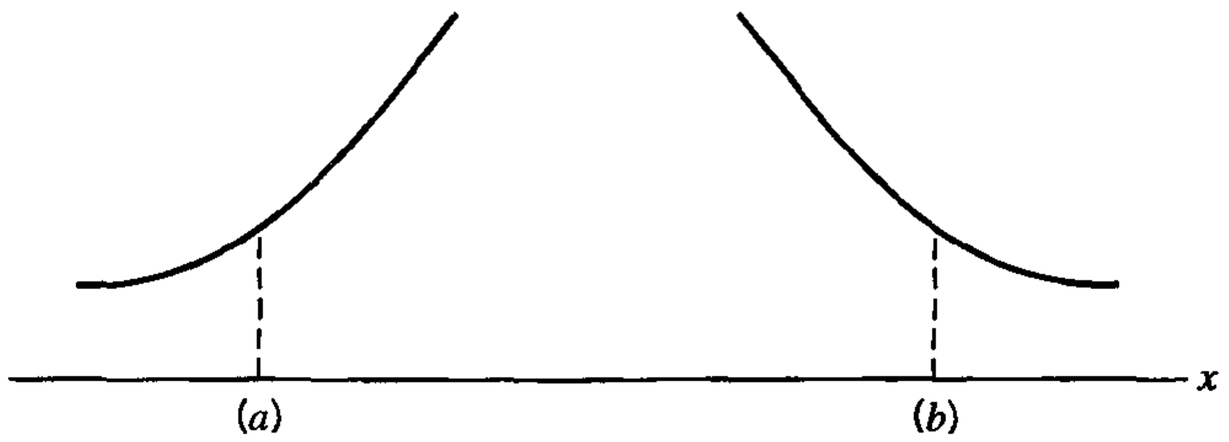


Figure 6.3. Convex shape of the wave function in the nonclassical region ($\psi''/\psi > 0$).

The Schrödinger equation has the solution

$$\psi(x) = \begin{cases} Ae^{ikx} + Be^{-ikx} & (x < 0) \\ Ce^{-\kappa x} & (x > 0) \end{cases} \tag{6.1}$$

Here

$$\hbar k = \sqrt{2mE} \qquad \hbar\kappa = \sqrt{2m(V_0 - E)}$$

Since $\psi(x)$ and its derivative $\psi'(x)$ approach zero as $x \to +\infty$, according to case (b) in Section 3.5 there is no degeneracy for $E < V_0$. The second linearly independent solution for $x > 0$, $e^{\kappa x}$, is in conflict with the boundary condition that $\psi(x)$ remain finite as $x \to +\infty$.

By joining the wave function and its slope smoothly at the discontinuity of the potential, $x = 0$, we have

$$\begin{aligned} A + B &= C \\ ik(A - B) &= -\kappa C \end{aligned}$$

or

$$\begin{aligned} \frac{B}{A} &= \frac{ik + \kappa}{ik - \kappa} = e^{i\alpha} \qquad (\alpha\text{: real}) \\ \frac{C}{A} &= \frac{2ik}{ik - \kappa} = 1 + e^{i\alpha} \end{aligned} \tag{6.2}$$

Substituting these values into (6.1), we obtain

$$\psi(x) = \begin{cases} 2Ae^{i\alpha/2} \cos\left(kx - \dfrac{\alpha}{2}\right) & (x < 0) \\ 2Ae^{i\alpha/2} \cos\dfrac{\alpha}{2}\, e^{-\kappa x} & (x > 0) \end{cases} \tag{6.3}$$

in agreement with the remark made in Section 3.5 that the wave function in the case of no degeneracy is real, except for an arbitrary constant factor. Hence, a graph of such a wave function may be drawn (Figure 6.2). The classical turning point ($x = 0$) is a *point of inflection* of the wave function. The oscillatory and exponential portions can be joined smoothly at $x = 0$ for all values of E between 0 and V_0: *the energy spectrum is continuous.*

The solution (6.1) can be given a straightforward interpretation. It represents a plane wave incident from the left with an amplitude A and a reflected wave that propagates toward the left with an amplitude B. According to (6.2), $|A|^2 = |B|^2$; hence, the reflection is total. A wave packet which is a superposition of eigenfunc-

tions (6.1) could be constructed to represent a particle incident from the left. This packet would move classically, being reflected at the wall and again giving a vanishing probability of finding the particle in the region of positive x after the wave packet has receded; there is no permanent penetration.

Perhaps these remarks can be better understood if we observe that for one-dimensional motion the conservation of probability leads to particularly transparent consequences. For a stationary state, Eq. (3.3) reduces to $dj/dx = 0$. Hence, the current density

$$j = \frac{\hbar}{2m\,i}\left[\psi^* \frac{d\psi}{dx} - \frac{d\psi^*}{dx}\,\psi\right] \tag{6.4}$$

has the same value at all points x. When calculated with the wave functions (6.3), the current density j is seen to vanish, as it does for any essentially real wave function. Hence, there is no net current anywhere at all. To the left of the potential step, the relation $|A|^2 = |B|^2$ ensures that incident and reflected probability currents cancel one another. If there is no current, there is no net momentum in the state (6.1).

Exercise 6.2. Show that for a wave function $\psi(x) = Ae^{ikx} + Be^{-ikx}$, the current density j can be expressed as the sum of an incident and a reflected current density, $j = j_{\text{inc}} + j_{\text{ref}}$, without any interference terms between incident and reflected waves.

The case of an infinitely high potential barrier ($V_0 \to \infty$ or $\kappa \to \infty$) deserves special attention. From (6.1) it follows that in this limiting case $\psi(x) \to 0$ in the region under the barrier, no matter what value the coefficient C may have. According to (6.2), the joining conditions for the wave function at $x = 0$ now reduce formally to

$$\lim_{\kappa\to\infty} \frac{B}{A} = -1 \qquad \lim_{\kappa\to\infty} \frac{C}{A} = 0$$

or $A + B = 0$ and $C = 0$ as $V_0 \to \infty$. These equations show that *at a point where the potential makes an infinite jump the wave function must vanish*, whereas its slope jumps discontinuously from a finite value ($2ikA$) to zero.

We next examine the quantum mechanics of a particle that encounters the potential step in one dimension with an energy $E > V_0$. Classically, this particle passes the potential step with altered velocity but no change of direction. The particle could be incident from the right or from the left. The solutions of the Schrödinger equation are now oscillatory in both regions; hence, to each value of the energy correspond two linearly independent, degenerate eigenfunctions, as discussed in case (a) in Section 3.5. For the physical interpretation, their explicit construction is best accomplished by specializing the general solution:

$$\psi(x) = \begin{cases} Ae^{ikx} + Be^{-ikx} & (x < 0) \\ Ce^{ik_1x} + De^{-ik_1x} & (x > 0) \end{cases} \tag{6.5}$$

where

$$\hbar k = \sqrt{2mE} \quad \text{and} \quad \hbar k_1 = \sqrt{2m(E - V_0)}$$

Two useful particular solutions are obtained by setting $D = 0$, or $A = 0$. The first of these represents a wave incident from the left. Reflection occurs at the potential step, but there is also transmission to the right. The second particular solution rep-

resents incidence from the right, transmission to the left, and reflection toward the right.

Here we consider only the first case ($D = 0$). The remaining constants are related by the condition for smooth joining at $x = 0$,

$$A + B = C$$
$$k(A - B) = k_1 C$$

from which we solve

$$\frac{B}{A} = \frac{k - k_1}{k + k_1} \quad \text{and} \quad \frac{C}{A} = \frac{2k}{k + k_1} \tag{6.6}$$

The current density is again constant, but its value is no longer zero. Instead,

$$j = \begin{cases} \dfrac{\hbar k}{m}(|A|^2 - |B|^2) & (x < 0) \\ \dfrac{\hbar k_1}{m}|C|^2 & (x > 0) \end{cases}$$

in agreement with Exercise 6.2. The equality of these values is assured by (6.6) and leads to the relation

$$\frac{|B|^2}{|A|^2} + \frac{k_1}{k}\frac{|C|^2}{|A|^2} = 1 \tag{6.7}$$

In analogy to optics, the first term in this sum is called the *reflection coefficient*, R, and the second is the *transmission coefficient*, T. We have

$$\boxed{R = \frac{|B|^2}{|A|^2} = \frac{(k - k_1)^2}{(k + k_1)^2}} \tag{6.8}$$

$$\boxed{T = \frac{k_1}{k}\frac{|C|^2}{|A|^2} = \frac{4kk_1}{(k + k_1)^2}} \tag{6.9}$$

Equation (6.7) ensures that $R + T = 1$. The coefficients R and T depend only on the ratio E/V_0.

For a wave packet incident from the left, the presence of reflection means that the wave packet may, when it arrives at the potential step, split into two parts, provided that its average energy is close to V_0. This splitting up of the wave packet is a distinctly nonclassical effect that affords an argument against the early attempts to interpret the wave function as measuring the matter (or charge) density of a particle. For the splitting up of the wave packet would then imply a physical breakup of the particle, and this would be very difficult to reconcile with the facts of observation. After all, electrons and other particles are always found as complete entities with the same distinct properties. On the other hand, there is no contradiction between the splitting up of a wave packet and the probability interpretation of the wave function.

Exercise 6.3. Show that, for a given energy E, the coefficients for reflection and transmission at a potential step are the same for a wave incident from the right as for a wave incident from the left. Note that the relative phase of the reflected to the incident amplitude is zero for reflection from a rising potential step, but π for reflection from a sharp potential drop.

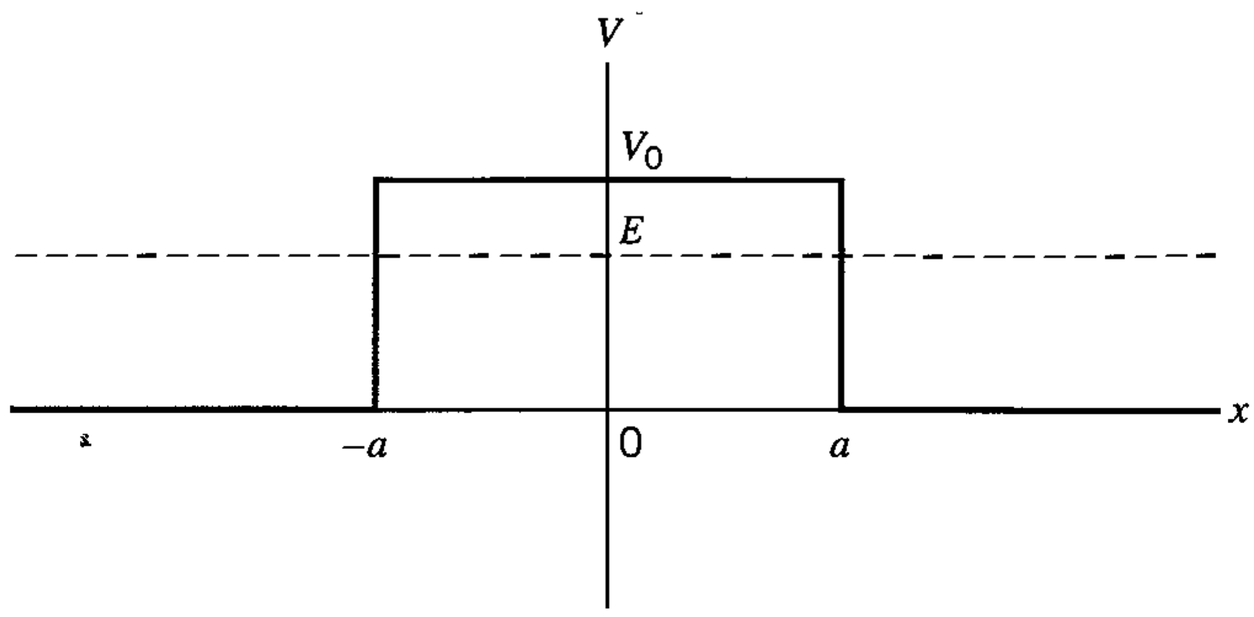


Figure 6.4. Rectangular potential barrier, height V_0, width $2a$.

2. ***The Rectangular Potential Barrier.*** In our study of more and more complicated potential forms, we now reach a very important case, the *rectangular potential barrier* (Figure 6.4). There is an advantage in placing the coordinate origin at the center of the barrier so that $V(x)$ is an even function of x. Owing to the quantum mechanical penetration of a barrier, a case of great interest is that of $E < V_0$. The particle is free for $x < -a$ and $x > a$. For this reason the rectangular potential barrier simulates, albeit schematically, the scattering of a free particle from any potential.

We can immediately write down the general solution of the Schrödinger equation for $E < V_0$:

$$\psi(x) = \begin{cases} Ae^{ikx} + Be^{-ikx} & (x < -a) \\ Ce^{-\kappa x} + De^{\kappa x} & (-a < x < a) \\ Fe^{ikx} + Ge^{-ikx} & (a < x) \end{cases} \tag{6.10}$$

where again $\hbar k = \sqrt{2mE}$, $\hbar\kappa = \sqrt{2m(V_0 - E)}$. The boundary conditions at $x = -a$ require

$$\begin{aligned} Ae^{-ika} + Be^{ika} &= Ce^{\kappa a} + De^{\kappa a} \\ Ae^{-ika} - Be^{ika} &= \frac{i\kappa}{k}(Ce^{\kappa a} - De^{-\kappa a}) \end{aligned} \tag{6.11}$$

These linear homogeneous relations between the coefficients, A, B, C, D are conveniently expressed in terms of matrices:

$$\begin{pmatrix} A \\ B \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \left(1 + \frac{i\kappa}{k}\right)e^{\kappa a + ika} & \left(1 - \frac{i\kappa}{k}\right)e^{-\kappa a + ika} \\ \left(1 - \frac{i\kappa}{k}\right)e^{\kappa a - ika} & \left(1 + \frac{i\kappa}{k}\right)e^{-\kappa a - ika} \end{pmatrix}\begin{pmatrix} C \\ D \end{pmatrix}$$

The joining conditions at $x = a$ are similar. They yield

$$\begin{pmatrix} C \\ D \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \left(1 - \frac{ik}{\kappa}\right)e^{\kappa a + ika} & \left(1 + \frac{ik}{\kappa}\right)e^{\kappa a - ika} \\ \left(1 + \frac{ik}{\kappa}\right)e^{-\kappa a + ika} & \left(1 - \frac{ik}{\kappa}\right)e^{-\kappa a - ika} \end{pmatrix}\begin{pmatrix} F \\ G \end{pmatrix}$$

Combining the last two equations, we obtain the relation between the wave function on both sides of the barrier:

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} \left(\cosh 2\kappa a + \frac{i\varepsilon}{2}\sinh 2\kappa a\right)e^{2ika} & \frac{i\eta}{2}\sinh 2\kappa a \\ -\frac{i\eta}{2}\sinh 2\kappa a & \left(\cosh 2\kappa a - \frac{i\varepsilon}{2}\sinh 2\kappa a\right)e^{-2ika} \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix} \tag{6.12}$$

where the abbreviated notation

$$\varepsilon = \frac{\kappa}{k} - \frac{k}{\kappa}, \qquad \eta = \frac{\kappa}{k} + \frac{k}{\kappa} \tag{6.13}$$

has been used. Note that $\eta^2 - \varepsilon^2 = 4$.

Exercise 6.4. Calculate the determinant of the 2×2 matrix in (6.12).

A particular solution of interest is obtained from (6.12) by letting $G = 0$. This represents a wave incident from the left and transmitted through the barrier to the right. A reflected wave whose amplitude is B is also present. We calculate easily:

$$\frac{F}{A} = \frac{e^{-2ika}}{\cosh 2\kappa a + i(\varepsilon/2)\sinh 2\kappa a} \tag{6.14}$$

The square of the absolute value of this quantity is the transmission coefficient for the barrier. It assumes an especially simple form for a high and wide barrier, which transmits poorly, such that $\kappa a \gg 1$. In first approximation,

$$\cosh 2\kappa a \approx \sinh 2\kappa a \approx e^{2\kappa a}/2$$

Hence,

$$T = \frac{|F|^2}{|A|^2} \approx 16e^{-4\kappa a}\left(\frac{k\kappa}{k^2 + \kappa^2}\right)^2 \tag{6.15}$$

Another limiting case is that of a very narrow but high barrier such that $V_0 \gg E$, $\kappa \gg k$, and $\kappa a \ll 1$, but $V_0 a$ or $\kappa^2 a$ is finite. Under these conditions,

$$T = \left|\frac{F}{A}\right|^2 \approx \frac{k^2}{k^2 + \kappa^4 a^2} = \frac{E}{E + \frac{2m}{\hbar^2}V_0^2 a^2} \tag{6.16}$$

If the "area" under the potential is denoted by

$$g = 2 \lim_{\substack{a\to 0 \\ V_0\to\infty}} V_0 a \tag{6.17}$$

the potential may be represented by a delta function positioned at the origin,

$$V(x) = g\delta(x) \tag{6.18}$$

and the transmission through this potential barrier is

$$T = \frac{E}{E + \frac{mg^2}{2\hbar^2}} \tag{6.19}$$

The matrix that connects A and B with F and G in (6.12) has very simple properties. If we write the linear relations as

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} \alpha_1 + i\beta_1 & \alpha_2 + i\beta_2 \\ \alpha_3 + i\beta_3 & \alpha_4 + i\beta_4 \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix} \tag{6.20}$$

and compare this with (6.12), we observe that the eight *real* numbers α_i and β_i in the matrix satisfy the conditions

$$\alpha_1 = \alpha_4, \qquad \beta_1 = -\beta_4, \qquad \alpha_2 = \alpha_3 = 0, \qquad \beta_2 = -\beta_3 \tag{6.21}$$

These five equations reduce the number of independent variables on which the matrix depends from eight to three. As can be seen from (6.12) and Exercise 6.4, we must add to this an equation expressing the fact that the determinant of the matrix is equal to unity. Using (6.21), this condition reduces to

$$\boxed{\alpha_1^2 + \beta_1^2 - \beta_2^2 = 1} \tag{6.22}$$

Hence, we are left with two parameters, as we must be, since the matrix depends explicitly on the two independent variables ka and κa.

Exercise 6.5. If the matrix elements are constrained by (6.21) and (6.22), show that (6.20) can be written as

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} e^{i\nu} \cosh \lambda & i \sinh \lambda \\ -i \sinh \lambda & e^{-i\nu} \cosh \lambda \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix} \tag{6.23}$$

where λ and ν are two real parameters. For the delta-function barrier (6.18), identify λ and ν in terms of g and k. More generally, verify that (6.12) has the form (6.23).

In the next section it will be shown that the conditions (6.21) and (6.22) imposed on (6.20), rather than pertaining specifically to the rectangular-shaped potential, are consequences of very general symmetry properties of the physical system at hand.

3. ***Symmetries and Invariance Properties.*** Since the rectangular barrier of Figure 6.4 is a real potential and symmetric about the origin, the Schrödinger equation is invariant under time reversal and space reflection. We can exploit these properties to derive the general form of the matrix linking the incident with the transmitted wave.

We recapitulate the form of the general solution of the Schrödinger equation:

$$\psi(x) = \begin{cases} Ae^{ikx} + Be^{-ikx} & (x < -a) \\ Ce^{-\kappa x} + De^{\kappa x} & (-a < x < a) \\ Fe^{ikx} + Ge^{-ikx} & (a < x) \end{cases} \tag{6.10}$$

The smooth joining conditions at $x = -a$ and $x = a$ lead to two linear homogeneous relations between the coefficients A, B, F, and G, but we want to see how far we can proceed without using the joining conditions explicitly. If we regard the wave function on one side of the barrier, say for $x > a$, as given, then the coefficients A and B must be expressible as linear homogeneous functions of F and G. Hence, a matrix M exists such that

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix} \tag{6.24}$$

An equivalent representation expresses the coefficients B and F of the *outgoing* waves in terms of the coefficients A and G of the *incoming* waves by the matrix relation

$$\begin{pmatrix} B \\ F \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} A \\ G \end{pmatrix} \tag{6.25}$$

Whereas the representation in terms of the S matrix is more readily generalized to three-dimensional situations, the M matrix is more appropriate in one-dimensional problems. On the other hand, the symmetry properties are best formulated in terms of the S matrix.

The S and M matrices can be simply related if conservation of probability is invoked. As was shown in Section 6.1, in a one-dimensional stationary state, the probability current density j must be independent of x. Applying expression (6.4) to the wave function (6.10), we obtain the condition

$$|A|^2 - |B|^2 = |F|^2 - |G|^2 \quad \text{or} \quad |B|^2 + |F|^2 = |A|^2 + |G|^2$$

as expected, since $|A|^2$ and $|F|^2$ measures the probability flow to the right, while $|B|^2$ and $|G|^2$ measure the flow in the opposite direction. Using matrix notation, we can write this as

$$(B^* \quad F^*)\begin{pmatrix} B \\ F \end{pmatrix} = (A^* \quad G^*)\tilde{S}^* S \begin{pmatrix} A \\ G \end{pmatrix} = (A^* \quad G^*)\begin{pmatrix} A \\ G \end{pmatrix}$$

where $\tilde{S}$ denotes the transpose matrix of S, and S^* the complex conjugate. It follows that S must obey the condition

$$\boxed{\tilde{S}^* S = 1} \tag{6.26}$$

Since the *Hermitian conjugate* of the matrix S is defined by

$$S^\dagger = \tilde{S}^* = \begin{pmatrix} S_{11}^* & S_{21}^* \\ S_{12}^* & S_{22}^* \end{pmatrix} \tag{6.27}$$

Equation (6.26) implies the statement that the inverse of S must be the same as its Hermitian conjugate. Such a matrix is said to be *unitary*.

For a 2×2 matrix S, the unitarity condition (6.26) implies the following constraints:

$$|S_{11}| = |S_{22}| \quad \text{and} \quad |S_{12}| = |S_{21}| \tag{6.28}$$

$$|S_{11}|^2 + |S_{12}|^2 = 1 \tag{6.29}$$

and

$$S_{11}S_{12}^* + S_{21}S_{22}^* = 0 \tag{6.30}$$

Exercise 6.6. Verify that the conditions (6.28), (6.29), and (6.30) follow from (6.26).

Since the potential is real, the Schrödinger equation has, according to Section 3.5, in addition to (6.10), the *time-reversed* solution,

$$\psi_T(x) = \psi^*(x) = \begin{cases} A^* e^{-ikx} + B^* e^{ikx} & (x < -a) \\ C^* e^{-\kappa x} + D^* e^{\kappa x} & (-a < x < a) \\ F^* e^{-ikx} + G^* e^{ikx} & (a < x) \end{cases} \tag{6.31}$$

Comparison of this solution with (6.10) shows that effectively the directions of motion have been reversed and the coefficient A has been interchanged with B^*, and F with G^*. Hence, in (6.25) we may make the replacements $A \leftrightarrow B^*$ and $F \leftrightarrow G^*$ and obtain an equally valid equation:

$$\begin{pmatrix} A^* \\ G^* \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} B^* \\ F^* \end{pmatrix} \tag{6.32}$$

Equations (6.32) and (6.25) can be combined to yield the condition

$$S^*S = I \tag{6.33}$$

This condition in conjunction with the unitarity relation (6.26) implies that the S matrix must be *symmetric* as a consequence of time reversal symmetry:

$$\tilde{S} = S \tag{6.34}$$

If S is unitary and symmetric, it is easy to verify by comparing Eqs. (6.24) and (6.25) that the M matrix assumes the form:

$$M = \begin{pmatrix} \dfrac{1}{S_{12}} & \dfrac{S_{11}^*}{S_{12}^*} \\ \dfrac{S_{11}}{S_{12}} & \dfrac{1}{S_{12}^*} \end{pmatrix} \tag{6.35}$$

subject to the condition:

$$\det M = \frac{1 - |S_{11}|^2}{|S_{12}|^2} = 1$$

Since the potential is an even function of x, another solution is obtained by replacing x in (6.10) by $-x$. This substitution gives

$$\psi_R(x) = \psi(-x) = \begin{cases} Ae^{-ikx} + Be^{ikx} & (x > a) \\ Ce^{\kappa x} + De^{-\kappa x} & (a > x > -a) \\ Fe^{-ikx} + Ge^{ikx} & (-a > x) \end{cases} \tag{6.36}$$

Now, Ge^{ikx} is a wave incident on the barrier from the left, Be^{ikx} is the corresponding transmitted wave, and Fe^{-ikx} is the reflected wave. The wave Ae^{-ikx} is incident from the right. Hence, in (6.25) we may make the replacements $A \leftrightarrow G$ and $B \leftrightarrow F$ and obtain

$$\begin{pmatrix} F \\ B \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} G \\ A \end{pmatrix}$$

For comparison with (6.25) this relation can also be written as

$$\begin{pmatrix} B \\ F \end{pmatrix} = \begin{pmatrix} S_{22} & S_{21} \\ S_{12} & S_{11} \end{pmatrix} \begin{pmatrix} A \\ G \end{pmatrix}$$

Hence, invariance under reflection implies the relations

$$S_{11} = S_{22} \quad \text{and} \quad S_{12} = S_{21} \tag{6.37}$$

If conservation of probability, time reversal invariance, and invariance under space reflection are *simultaneously* demanded, the matrix M has the structure

$$M_{11} = M_{22}^*, \qquad M_{12} = -M_{12}^* = -M_{21} = M_{21}^*, \qquad \det M = 1 \tag{6.38}$$

Exercise 6.7. Verify the relations (6.38). Check that they are satisfied by the M matrices in Eqs. (6.12) and (6.23) and that they are equivalent to conditions (6.21) and (6.22).

We thus see that the conditions (6.21) and (6.22) can be derived from very general properties without knowledge of the detailed shape of the potential. These general properties are shared by all real potentials that are symmetric with respect to the origin and vanish for large values of $|x|$. For all such potentials the solution of the Schrödinger equation must be asymptotically of the form

$$\psi(x) \to \begin{cases} Ae^{ikx} + Be^{-ikx} & (x \to -\infty) \\ Fe^{ikx} + Ge^{-ikx} & (x \to +\infty) \end{cases}$$

By virtue of the general arguments just advanced, these two portions of the eigenfunctions are related by the equation

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} \alpha_1 + i\beta_1 & i\beta_2 \\ -i\beta_2 & \alpha_1 - i\beta_1 \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix} \tag{6.39}$$

with *real* parameters α_1, β_1, and β_2 subject to the additional constraint

$$\alpha_1^2 + \beta_1^2 - \beta_2^2 = 1 \tag{6.22}$$

The same concepts can be generalized to include long-range forces. All that is needed to define a matrix M with the properties (6.39) and (6.22) is that the Schrödinger equation admit two linearly independent fundamental solutions that have the *asymptotic* property

$$\psi_1(-x) = \psi_2(x) = \psi_1^*(x) \qquad (x \to \pm\infty)$$

For a real even potential function $V(x)$, this can always be accomplished by choosing

$$\psi_1(x) = \psi_{\text{even}}(x) + i\psi_{\text{odd}}(x), \qquad \psi_2(x) = \psi_{\text{even}}(x) - i\psi_{\text{odd}}(x)$$

where ψ_{even} and ψ_{odd} are the real-valued even and odd parity solutions defined in Section 5.2.

Although the restrictions that various symmetries impose on the S or M matrix usually complement each other, they are sometimes redundant. For instance, in the simple one-dimensional problem treated in this section, invariance under reflection, if applicable, guarantees that the S matrix is symmetric [see the second Eq. (6.37)], thus yielding a condition that is equally prescribed by invariance under time reversal together with probability conservation.

It should also be noted that the principle of invariance under time reversal is related to probability conservation and therefore to unitarity. If no velocity-dependent interactions are present, so that V is merely a function of position, the reality of V ensures invariance under time reversal and implies conservation of probability. Velocity-dependent interactions, as they occur for instance in the presence of magnetic fields, can break time reversal symmetry without violating conservation of probability (unitarity).

The matrix method of this section allows a neat separation between the initial conditions for a particular problem and the matrices S and M, which do not depend on the structure of the initial wave packet. The matrices S and M depend only on the nature of the dynamical system, the forces, and the energy. Once either one of

these matrices has been worked out as a function of energy, all problems relating to the potential barrier have essentially been solved. For example, the transmission coefficient T is given by $|F|^2/|A|^2$ if $G = 0$, and therefore

$$T = \left|\frac{F}{A}\right|^2 = \frac{1}{|M_{11}|^2} = |S_{21}|^2 \tag{6.40}$$

Exercise 6.8. If $V = 0$ for all x (free particle), show that

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{6.41}$$

If for real-valued $V \neq 0$, the departure of the S matrix from the form (6.41) is measured by two complex-valued functions of the energy, r and t, which are defined by $S_{11} = 2ir$ and $S_{21} = 1 + 2it$, prove the relation

$$|r|^2 + |t|^2 = \operatorname{Im} t \tag{6.42}$$

The analogue of this identity in three dimensions is known as an *optical theorem* (Section 20.6).

We will encounter other uses of the M and S matrices in the next section and the next chapter. Eventually, in Chapter 20, we will see that similar methods are pertinent in the general theory of collisions, where the S or *scattering matrix* plays a central role. The work of this section is S-matrix theory in its most elementary form.

Exercise 6.9. Noting that the wave number k appears in the Schrödinger equation only quadratically, prove that, as a function of k, the S matrix has the property

$$S(k)\, S(-k) = 1 \tag{6.43}$$

Derive the corresponding properties of the matrix M, and verify them for the example of Eq. (6.12).

Exercise 6.10. Using conservation of probability and invariance under time reversal only, prove that at a fixed energy the value of the transmission coefficient is independent of the direction of incidence. (See also Exercise 6.3.)

4. ***The Square Well.*** Finally, we must discuss the so-called *square* (or rectangular) *well* (Figure 6.5). It is convenient to place the origin of the x axis in the center of the potential well so that $V(x)$ is again an even function of x:

$$V(x) = \begin{cases} -V_0 & \text{for} \quad -a < x < a \qquad (V_0 > 0) \\ 0 & \text{for} \quad |x| \geq a \end{cases}$$

Depending on whether the energy is positive or negative, we distinguish two separate cases. If $E > 0$, the particle is unconfined and is scattered by the potential; if $E < 0$, it is confined and in a bound state. Assuming $-V_0 \leq E < 0$, we treat this last case first and set

$$\hbar k' = \sqrt{2m(E + V_0)}, \qquad \hbar\kappa = \sqrt{-2mE} \tag{6.44}$$

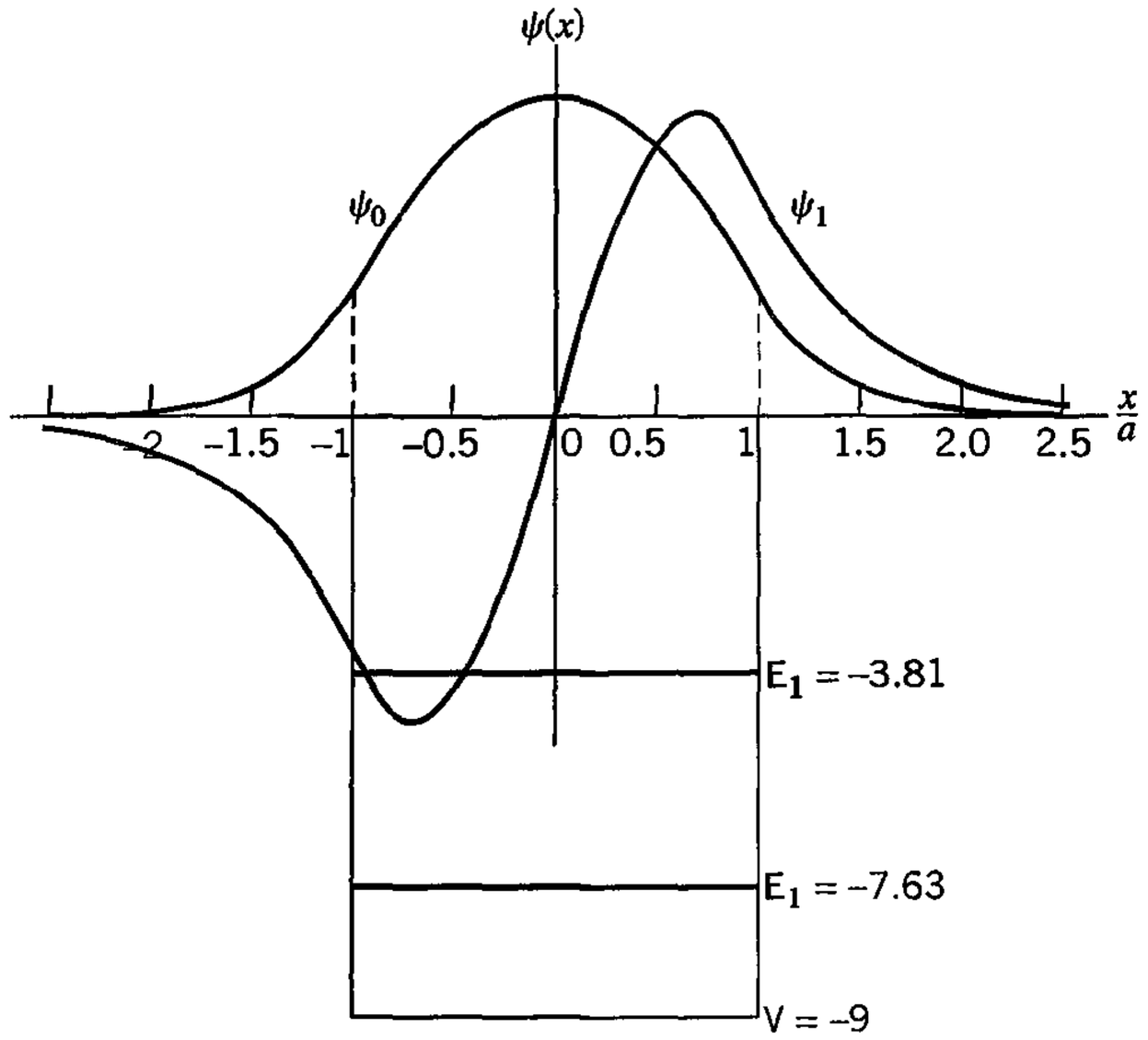


Figure 6.5. Square well potential of width $2a$ and depth V_0. For the choice of the parameter $\beta = 3$ there are two bound states. In units of $\hbar^2/2ma^2$, the well depth is $V_0 = 9$, and the energies for the two discrete states are $E_0 = -7.63$ and $E_1 = -3.81$. The ground and excited state eigenfunctions are equally normalized, but the scale is arbitrary.

The Schrödinger equation takes the form

$$\begin{aligned} \frac{d^2\psi(x)}{dx^2} + k'^2\psi(x) &= 0 \qquad \text{inside the well} \\ \frac{d^2\psi(x)}{dx^2} - \kappa^2\psi(x) &= 0 \qquad \text{outside the well} \end{aligned} \qquad (E < 0) \qquad (6.45)$$

As for any even potential, we may restrict the search for eigenfunctions to those of definite parity.

Inside the well we have

$$\begin{aligned} \psi(x) &= A' \cos k'x \qquad \text{for even parity} \\ \psi(x) &= B' \sin k'x \qquad \text{for odd parity} \end{aligned} \qquad (6.46)$$

Outside the well we have only the decreasing exponential

$$\psi(x) = C' e^{-\kappa|x|} \qquad (6.47)$$

since the wave function must not become infinite at large distances.

It is necessary to join the wave function and its first derivative smoothly at $x = a$, that is, to require,

$$\begin{aligned} \lim_{\varepsilon\to 0} \psi(a - \varepsilon) &= \lim_{\varepsilon\to 0} \psi(a + \varepsilon) \\ \lim_{\varepsilon\to 0} \psi'(a - \varepsilon) &= \lim_{\varepsilon\to 0} \psi'(a + \varepsilon) \end{aligned}$$

Since an overall constant factor remains arbitrary until determined by normalization, these two conditions are equivalent to demanding that the *logarithmic derivative* of ψ,

$$\frac{1}{\psi(x)}\frac{d\psi(x)}{dx} = \frac{d \log \psi(x)}{dx} \tag{6.48}$$

be continuous at $x = a$. This is a very common way of phrasing the smooth joining conditions. Because of the reflection symmetry, the smooth joining conditions are automatically satisfied at $x = -a$ for both even and odd eigenfunctions.

The logarithmic derivative of the outside wave function, evaluated at $x = a$, is $-\kappa$; that of the inside wave functions is $-k' \tan k'a$ for the even case and $k' \cot k'a$ for the odd case.

The transcendental equations

$$\begin{aligned} k' \tan k'a &= \kappa \qquad \text{(even)} \\ k' \cot k'a &= -\kappa \qquad \text{(odd)} \end{aligned} \tag{6.49}$$

permit us to determine the allowed eigenvalues of the energy E.

The general symmetry considerations of Section 6.3 can also be extended to the solutions of the Schrödinger equation with negative values of k^2 and E. In (6.10) we need only replace k by $i\kappa$ and κ by ik'. The solution then takes the form

$$\psi(x) = \begin{cases} Ae^{-\kappa x} + Be^{\kappa x} & (x < -a) \\ Ce^{-ik'x} + De^{ik'x} & (-a < x < a) \\ Fe^{-\kappa x} + Ge^{\kappa x} & (a < x) \end{cases} \tag{6.50}$$

By requiring invariance under time reversal and imposing the principle of conservation of probability, we see that the matrix M, defined as in (6.24), must now be a *real* matrix with det $M = 1$. The boundary conditions at large distances require that $A = G = 0$. In terms of the matrix M, we must thus demand that

$$M_{11} = 0 \tag{6.51}$$

and this equation yields the energy eigenvalues. Defining again an S matrix as in (6.25), it follows that the bound-state energies are *poles of the S matrix.*

Exercise 6.11. Show that for the square well, reflection symmetry implies that the off-diagonal matrix elements of M are ± 1, giving us the even and odd solutions, respectively.

Exercise 6.12. By changing V_0 into $-V_0$ in (6.12), show that, for a square well, (6.51) is equivalent to the eigenvalue conditions (6.49).

A simple graphical method aids in visualizing the roots of (6.49). We set

$$\chi = (k'a)^2, \qquad \beta = \sqrt{\frac{2mV_0a^2}{\hbar^2}}, \qquad (\kappa a)^2 + \chi = \beta^2$$

In Figure 6.6 we plot $F(\chi) = \beta^2 - \chi$ and

$f_{\text{even}}(\chi) = \chi \tan^2 \sqrt{\chi}$ (if $\tan \sqrt{\chi} \geq 0$) $\qquad$ $f_{\text{odd}}(\chi) = \chi \cot^2 \sqrt{\chi}$ (if $\tan \sqrt{\chi} \leq 0$)

as functions of the positive independent variable χ. For the square well, the only pertinent parameter is the value of the dimensionless quantity β. The required roots are found by determining the intercepts of the straight line $F(\chi)$ with the curves f_{even} and f_{odd}. The ordinates $(\kappa a)^2$ of the intersection points are the scaled values of the bound-state energies.

By inspection of Figure 6.6, we can immediately draw several conclusions: All bound states of the well are nondegenerate; even and odd solutions alternate as the energy increases; the number of bound states is finite and equal to $N + 1$, if $N\pi < 2\beta \leq (N + 1)\pi$; if the bound states are labeled in order of increasing energy by a quantum number $n = 0, 1, \ldots N$, even values of n correspond to even parity, odd values of n correspond to odd parity, and n denotes the number of nodes; for any one-dimensional square well there is always at least one even state, but there can be no odd states unless $\beta > \pi/2$; and, the level spacing increases with increasing n.

As V_0 is allowed to increase beyond all bounds, two special cases merit discussion:

(a) Here we let $V_0 \to \infty$ while keeping the width of the square well finite, so that $\beta \to \infty$. For this infinitely deep potential well, the roots of the equations

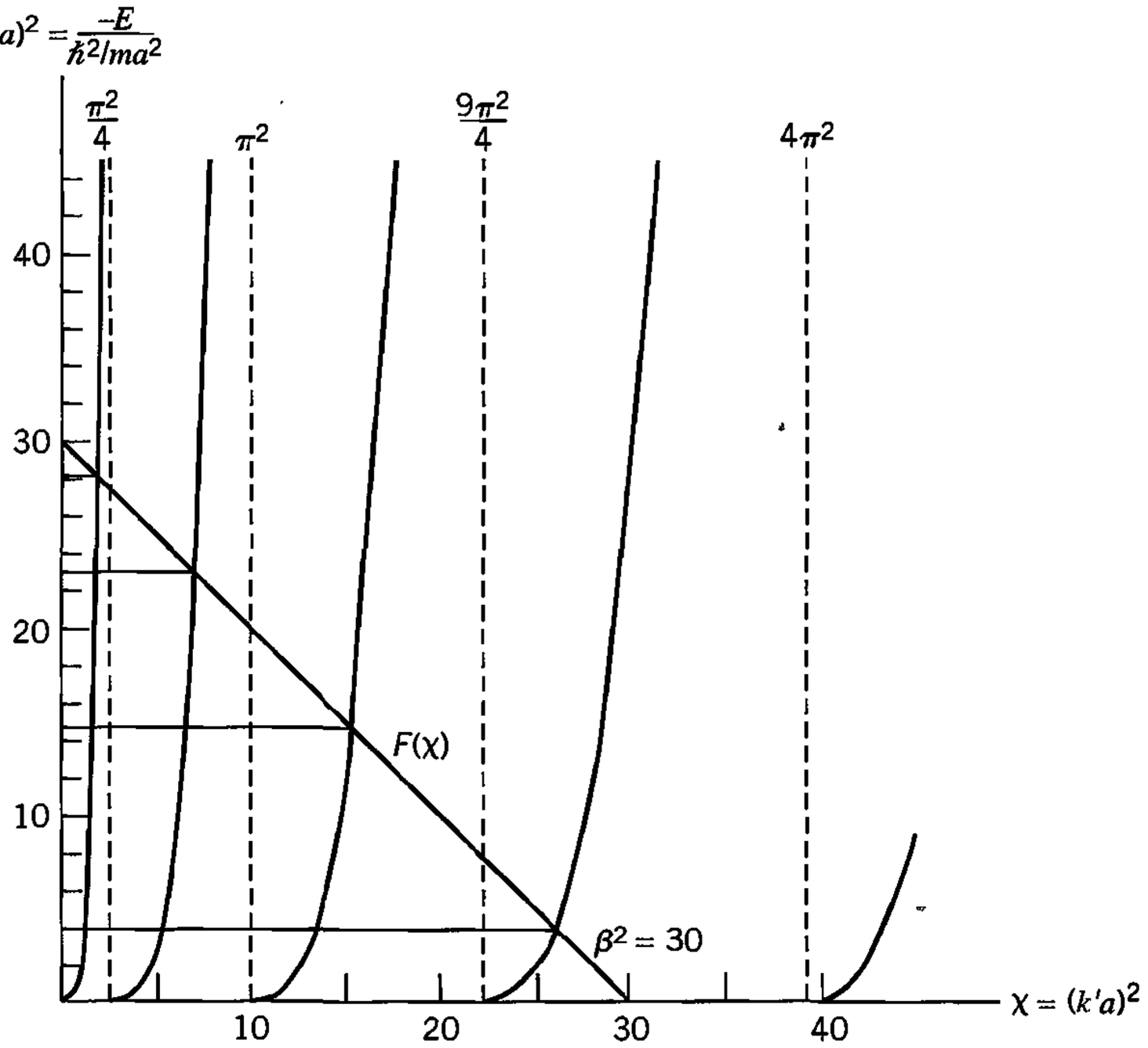


Figure 6.6. Graphic determination of the energy levels in a square well with $\beta^2 = 30$. The curves for $\tan\sqrt{\chi} \geq 0$ (f_{even}) alternate with those for $\tan\sqrt{\chi} \leq 0$ (f_{odd}). The ordinates of the intercepts are the binding energies in units of $\hbar^2/2ma^2$. The dashed-line asymptotes intersect the abscissa at the energy eigenvalues (6.52), again in units $\hbar^2/2ma^2$, for a particle confined to the box $-a \leq x \leq a$. This figure may be used as a template for estimating the bound-state energies for any one-dimensional square well: merely draw a parallel to the diagonal straight line $F(\chi) = \beta^2 - \chi$ for the desired value of β.

expressing the boundary conditions are now simply the asymptotes in Figure 6.6, or

$$\sqrt{\chi} = (n + 1)\,\pi/2 \qquad (n = 0, 1, 2, \ldots)$$

Hence,

$$\boxed{E + V_0 = (n + 1)^2 \frac{\pi^2\hbar^2}{8ma^2}} \tag{6.52}$$

The left-hand side of this expression, $E + V_0$, is the distance in energy from the bottom of the well and represents the kinetic energy of the particle in the well. Since $E \to -\infty$ as $V_0 \to \infty$, it follows that $\kappa \to +\infty$; therefore, the wave function itself must vanish outside the well and at the endpoints $x = \pm a$. There is in this limit no condition involving the slope, which for an infinite potential jump can be discontinuous. Taking into account a shift V_0 of the zero of energy and making the identification $2a = L$, we see that the energy levels (6.52) for odd values of n coincide with the energy spectrum for a free particle whose wave function is subject to periodic boundary conditions (see Section 4.4). Note that the number of states is essentially the same in either case, since there is double degeneracy in (4.59) for all but the lowest level, whereas (6.52) has no degeneracy, but between any two levels (4.59) there lies one given by (6.52) corresponding to even values of n. There is, however, no eigenstate of the infinitely deep well at $E + V_0 = 0$ [corresponding to $n = 0$ in (4.59)], because the corresponding eigenfunction vanishes.

(b) Another interesting special case arises if V_0 tends to infinity as a tends to zero, but in such a way that the product $V_0 a$ remains finite. As in (6.17), we denote the area under the potential by $g = \lim(2V_0 a)$, but instead of being a repulsive barrier the potential is now the attractive delta function well,

$$V(x) = -g\delta(x) \tag{6.53}$$

In this limit $k' \to \infty$, but $k'a \to 0$, and $k'^2 a \to 2mV_0 a/\hbar^2$ and remains finite. There are no odd solutions of (6.49) in this case, but there is one even root given by

$$k'^2 a = \kappa$$

or

$$\boxed{E = -\frac{mg^2}{2\hbar^2}} \tag{6.54}$$

Thus, the attractive one-dimensional delta function well supports only one bound state.

This conclusion can be verified directly from the Schrödinger equation,

$$-\frac{\hbar}{2m}\frac{d^2\psi(x)}{dx^2} - g\delta(x) = E\psi(x) \tag{6.55}$$

if we integrate this equation from $x = -\varepsilon$ to $x = +\varepsilon$ and then take the limit $\varepsilon \to 0$. The result is

$$\lim_{\varepsilon\to 0} \psi'(x)\Big|_{-\varepsilon}^{+\varepsilon} + \frac{2mg}{\hbar^2}\psi(0) = 0 \tag{6.56}$$

From (6.47) we see that $\psi(0) = C'$ and

$$\lim_{\varepsilon \to 0} \psi'(x)\Big|_{-\varepsilon}^{+\varepsilon} = -2C'\kappa$$

Inserting these values in (6.56), we obtain

$$\kappa = \frac{mg}{\hbar^2} \tag{6.57}$$

which is equivalent to the energy equation (6.54).

Exercise 6.13. Use Eq. (6.56) to derive the transmission coefficient T for the delta-function potential as a function of energy, for $E > 0$. Compare with (6.19).

Exercise 6.14. Show that the energy eigenvalue equations (6.49) can be cast in the alternate form

$$\cos\left(k'a - n\frac{\pi}{2}\right) = \frac{k'a}{\beta} \quad \text{for} \quad n\frac{\pi}{2} < k'a < (n+1)\frac{\pi}{2} \tag{6.58}$$

Devise a simple graphical method for obtaining the roots of (6.58). If $\beta = 20$, compute approximate values for the bound-state energy levels in units of $\hbar^2/ma^2$.

To conclude this chapter we discuss briefly what happens to a particle incident from a great distance when it is scattered by a square well. Here $E > 0$. Actually, this problem has already been solved. We may carry over the results for the potential barrier, replacing V_0 by $-V_0$ and κ by ik', where $\hbar k' = \sqrt{2m(E + V_0)}$. Equation (6.12) becomes

$$\begin{pmatrix} A \\ B \end{pmatrix} = \begin{pmatrix} \left(\cos 2k'a - \frac{i\varepsilon'}{2}\sin 2k'a\right)e^{2ika} & -\frac{i\eta'}{2}\sin 2k'a \\ \frac{i\eta'}{2}\sin 2k'a & \left(\cos 2k'a + \frac{i\varepsilon'}{2}\sin 2k'a\right)e^{-2ika} \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix} \tag{6.59}$$

where

$$\varepsilon' = \frac{k'}{k} + \frac{k}{k'}, \qquad \eta' = \frac{k'}{k} - \frac{k}{k'} \tag{6.60}$$

Equation (6.59) defines the matrix M for the square well if the energy is positive.

The transmission coefficient T is obtained from (6.59) by choosing stationary states with $G = 0$ (no wave incident from the right), and writing

$$\frac{F}{A} = \frac{1}{M_{11}} = S_{12} = \sqrt{T}e^{-i\phi} \tag{6.61}$$

This expression defines a phase shift ϕ between the transmitted and the incident wave. From the properties of the S matrix in Section 6.3, or directly from (6.59), we find for the relation between the reflected and the incident wave:

$$\frac{B}{A} = \frac{M_{21}}{M_{11}} = S_{11} = \sqrt{1 - T}e^{-i(\phi \pm \pi/2)} \tag{6.62}$$

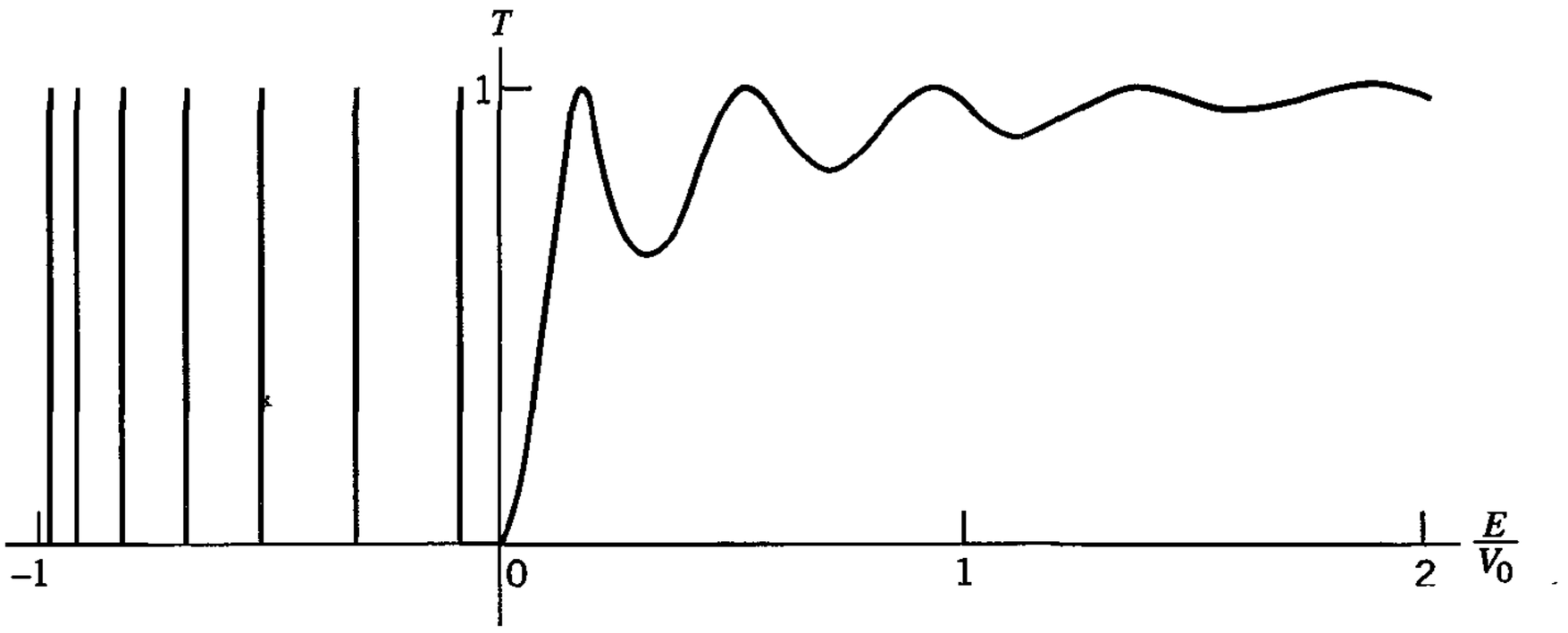


Figure 6.7. Transmission coefficient T versus E/V_0 for a square well with $13\pi/4$. The spikes on the left are at the positions of the seven discrete bound-state energy levels.

For the square well,

$$T = \frac{1}{|M_{11}|^2} = \frac{1}{\cos^2 2k'a + \dfrac{\varepsilon'^2}{4} \sin^2 2k'a} \tag{6.63}$$

As $E \to \infty$, $\varepsilon' \to 2$, and $T \to 1$, as expected. As a function of energy, the transmission coefficient rises from zero, fluctuates between maxima ($T = 1$) at $2k'a = n\pi$ and minima near $2k'a = (2n + 1)\dfrac{\pi}{2}$, and approaches the classical value $T = 1$ at the higher energies. Figures 6.7 and 6.8 show this behavior for two different values of

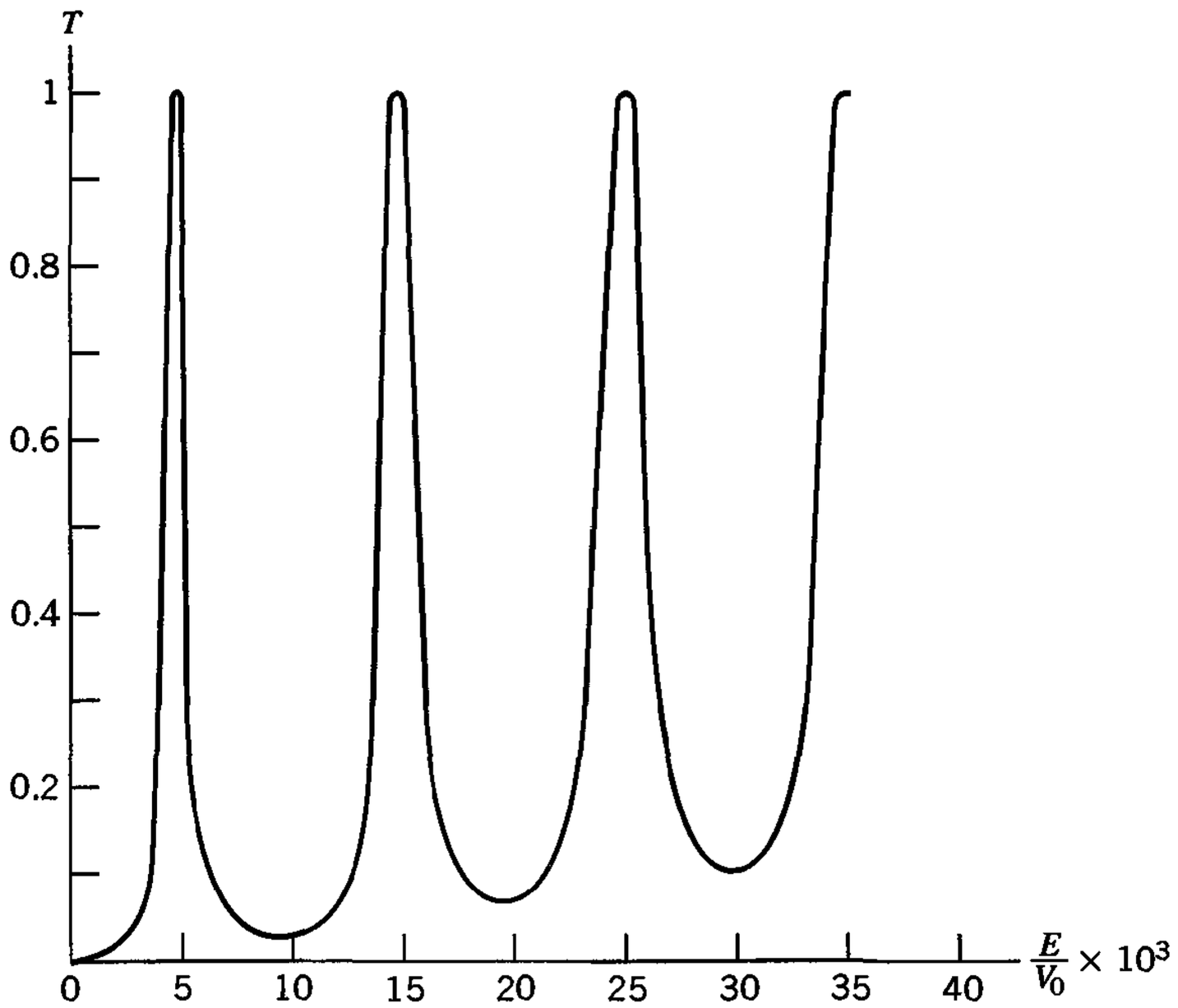


Figure 6.8. Transmission coefficient T versus E/V_0 for a deep square well with $\beta = 315$. As E increases, the resonances become broader.

β. The maxima occur when the distance $4a$ that a particle covers in traversing the well and back equals an integral number of de Broglie wavelengths, so that the incident wave and the waves that are reflected inside the well are in phase reinforcing each other. If the well is deep and the energy E low (β and $\varepsilon' \gg 1$), the peaks stand out sharply between comparatively flat minima (see Figure 6.8). When the peaks in the transmission curve are pronounced they are said to represent *resonances.*

The *phase shift* ϕ can also be calculated from (6.59). We find

$$\phi = 2ka - \arctan\left(\frac{\varepsilon'}{2}\tan 2k'a\right) \tag{6.64}$$

Figure 6.9 portrays the energy derivative of the phase shift, $\frac{d\phi}{dE}$ in units of V_0, as a function of energy, for the same square well as in Figure 6.8. The resonances show up as pronounced points of inflection in the function $\phi(E)$.[1]

Exercise 6.15. Show that the expressions (6.61) and (6.62) for the elements of the S matrix follow from the general properties derived in Section 6.3. From the matrix M for the square well, derive the expression (6.64) for the phase shift. For a square well with $\beta = 315$, as in Figures 6.8 and 6.9, compute numerically and sketch graphically the energy dependence of the phase shift $\phi(E)$ for $E > 0$ in the resonance domain.

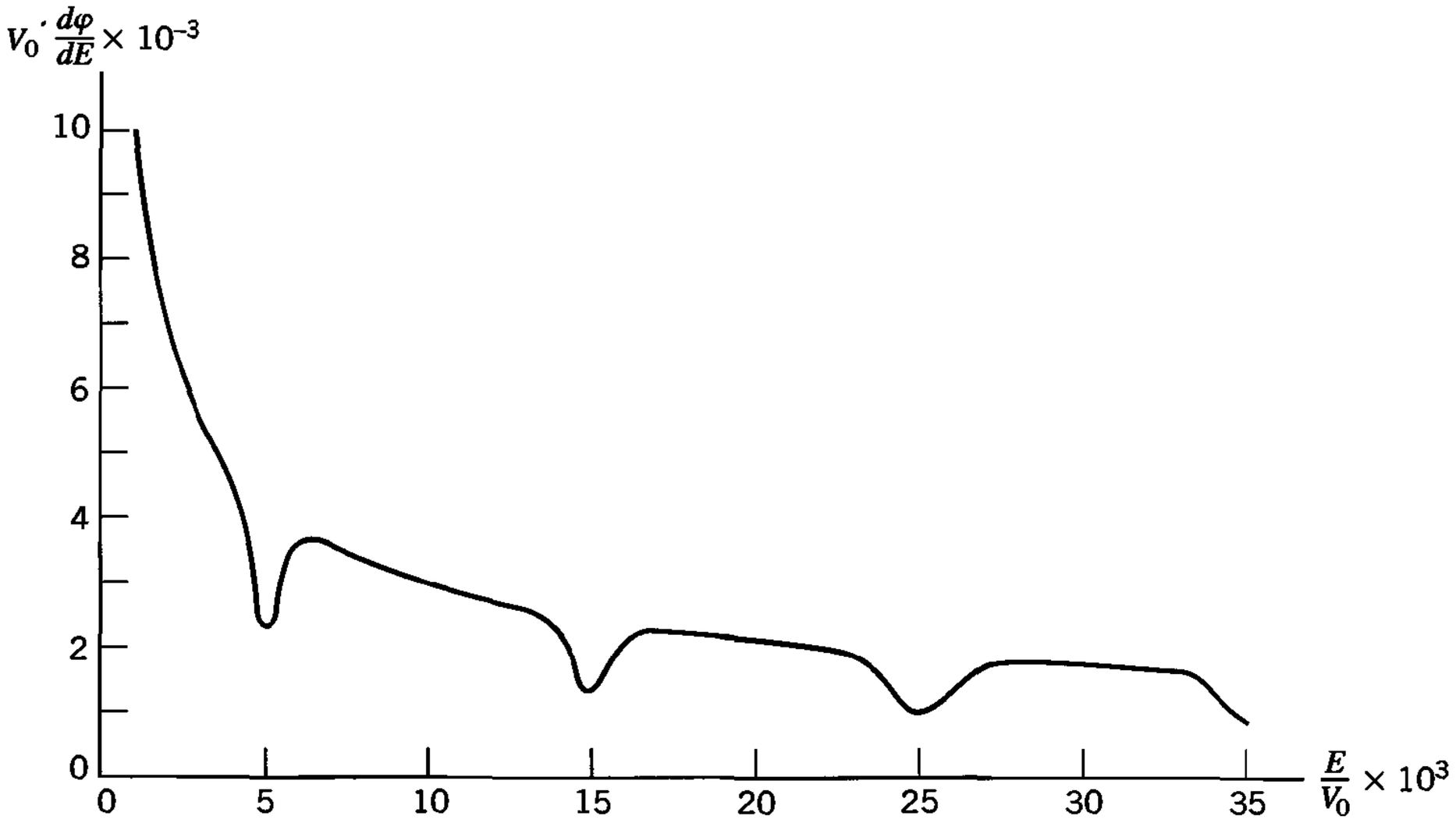


Figure 6.9. The energy derivative of the phase shift, $\frac{d\phi}{dE}$, in units of V_0 for a deep square well with $\beta = 315$.

[1]The energy derivative of the phase shift can be related to the time delay suffered by a particle at resonance inside the potential well. See Merzbacher (1970), Section 6.8.

Resonance peaks in the transmission of particles are typical quantum features, and the classical picture is not capable of giving a simple account of such strong but smooth energy variations. Classically, depending on the available energy, T can only be zero or one, whereas in quantum mechanics T changes continuously between these limits. In the example of the potential barrier of Figure 6.4, although the transmission coefficient (6.15) is numerically small for $E < V_0$, it is different from zero and varies continuously with energy. Classically, for such a barrier. T jumps from 0 to 1 at $E = V_0$. Thus, in a certain sense quantum mechanics attributes to matter more continuous and less abrupt characteristics than classical mechanics.

While these observations have general validity, their verification by extending the solutions of the Schrödinger for discontinuous potentials to the classical limit meets with some obstacles. For example, the reflection coefficient (6.8) does not depend on $\hbar$ and is a function of the particle momentum only. Hence, it is apparently applicable to a particle moving under classical conditions. Yet classically, R is either 0 or 1. This paradox is resolved if we recognize that the correct classical limit of quantum equations is obtained only if care is taken to keep the de Broglie wavelength short in comparison with the distance over which the fractional change of the potential is appreciable. The Schrödinger equation for the piecewise constant potential patently violates this requirement, but the next chapter will deal with potentials for which this condition is well satisfied.

The transmission resonance theory outlined in this section cannot be expected to provide quantitative estimates for phenomena in the atomic and nuclear domain. Not only is the square well unrealistic as a representation of the forces, but also the limitation to one dimension is a gross distortion of the real systems. To appreciate the distinction, one only needs to be aware of the totally different energy spectra of the bound states in a square well in one and three dimensions. From formula (6.52), as well as from the analogous discussion in Section 4.4 for the particle in a box, we know that in one dimension the levels are spaced farther and farther apart with increasing energy; in three dimensions, however, the density of energy levels in a well increases rapidly with increasing energy. Resonances in three dimensions will be discussed in Chapter 13.

Problems

1. Obtain the transmission coefficient for a rectangular potential barrier of width $2a$ if the energy exceeds the height V_0 of the barrier. Plot the transmission coefficient as a function of E/V_0 (up to $E/V_0 = 3$), choosing $(2ma^2V_0)^{1/2} = (3\pi/2)\,\hbar$.
2. Consider a potential $V = 0$ for $x > a$, $V = -V_0$ for $a \geq x \geq 0$, and $V = +\infty$ for $x < 0$. Show that for $x > a$ the positive energy solutions of the Schrödinger equation have the form

$$e^{i(kx+2\delta)} - e^{-ikx}$$

   Calculate the *scattering coefficient* $|1 - e^{2i\delta}|^2$ and show that it exhibits maxima (resonances) at certain discrete energies if the potential is sufficiently deep and broad.
3. A particle of mass m moves in the one-dimensional double well potential

$$V(x) = -g\delta(x - a) - g\delta(x + a)$$

   If $g > 0$, obtain transcendental equations for the bound-state energy eigenvalues of the system. Compute and plot the energy levels in units of $\hbar^2/ma^2$ as a function of

the dimensionless parameter $mag/\hbar^2$. Explain the features of this plot. In the limi of large separation, $2a$, between the wells, obtain a simple formula for the splitting ΔE between the ground state (even parity) energy level, E_+, and the excited (odd parity) energy level, E_-.

4. Problem 3 provides a primitive model for a one-electron linear diatomic molecule with interatomic distance $2a = |X|$, if the potential energy of the "molecule" is taken as $E_\pm(|X|)$, supplemented by a repulsive interaction $\lambda g/|X|$ between the wells ("atoms"). Show that, for a sufficiently small value of λ, the system ("molecule") is stable if the particle ("electron") is in the even parity state. Sketch the total potential energy of the system as a function of $|X|$.

5. If the potential in Problem 3 has $g < 0$ (double barrier), calculate the transmission coefficient and show that it exhibits resonances. (Note the analogy between the system and the Fabry-Perot étalon in optics.)

6. A particle moves in one dimension with energy E in the field of a potential defined as the sum of a Heaviside step function and a delta function:

$$V(x) = V_0\,\eta(x) + g\delta(x) \quad \text{(with } V_0 \text{ and } g > 0\text{)}$$

The particle is assumed to have energy $E > V_0$.

(a) Work out the matrix M, which relates the amplitudes of the incident and reflected plane waves on the left of the origin ($x < 0$) to the amplitudes on the right ($x > 0$).

(b) Derive the elements of the matrix S, which relates incoming and outgoing amplitudes.

(c) Show that the S matrix is unitary and that the elements of the S matrix satisfy the properties expected from the applicable symmetry considerations.

(d) Calculate the transmission coefficients for particles incident from the right and for particles incident from the left, which have the same energy (but different velocities).

7. For the potentials in Problems 5 and 6, verify the identity

$$|r|^2 + |t|^2 = \text{Im } t$$

for the complex-valued amplitudes r and t, if the elements of the S matrix are expressed as $S_{11} = 2ir$ and $S_{21} = 1 + 2it$.

CHAPTER 7

The WKB Approximation

If the potential energy does not have a very simple form, the solution of the Schrödinger equation even in one dimension is usually a complicated mathematical problem that requires the use of approximation methods. Instead of starting with a simplified potential, as perturbation theory (Chapter 8) does, the WKB approximation assumes that the potential varies slowly as a function of x. The solution of the Schrödinger equation is represented as a modulated constant-potential wave function. The method is useful to advance our understanding of tunneling through a potential barrier, resonance behavior in the continuum, and exponential decay of an unstable system.

1. ***The Method.*** The WKB method for obtaining approximate solutions of the Schrödinger equation

$$\boxed{\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}(E - V)\psi = 0} \tag{7.1}$$

is named after its proponents in quantum mechanics, G. Wentzel, H. A. Kramers, and L. Brillouin, but has its roots in the theory of ordinary differential equations. It can also be applied to three-dimensional problems, if the potential is spherically symmetric and a radial differential equation can be established.

The basic idea is simple. If $V = \text{const}$, (7.1) has the solutions $e^{\pm ikx}$. This suggests that if V, while no longer constant, varies only slowly with x, we might try a solution of the form

$$\psi(x) = e^{iu(x)} \tag{7.2}$$

except that the function $u(x)$ now is not simply linear in the variable x. The same idea was already used in Sections 2.5 and 3.1 for the time-dependent wave equation in order to make the connection between quantum mechanics and Hamilton's theory of classical mechanics. The function $u(x)$ in (7.2) is related to the function $S(x, t)$ in (3.2) as

$$S(x, t) = \hbar u(x) - Et \tag{7.3}$$

which explains why the WKB method is occasionally referred to as a *semiclassical* version of quantum mechanics.

Substitution of (7.2) into (7.1) gives us an equation for the x-dependent "phase," $u(x)$. This equation becomes particularly simple if we use the abbreviations

$$k(x) = \left\{\frac{2m}{\hbar^2}[E - V(x)]\right\}^{1/2} \qquad \text{if } E > V(x) \tag{7.4}$$

and

$$k(x) = -i\left\{\frac{2m}{\hbar^2}[V(x) - E]\right\}^{1/2} = -i\kappa(x) \qquad \text{if } E < V(x) \tag{7.5}$$

We find that $u(x)$ satisfies the equation

$$\boxed{i\frac{d^2u}{dx^2} - \left(\frac{du}{dx}\right)^2 + [k(x)]^2 = 0} \tag{7.6}$$

This differential equation is entirely equivalent to (7.1), but the boundary conditions are more easily expressed in terms of $\psi(x)$ than $u(x)$. Although the Schrödinger equation is linear, (7.6), like the classical Hamilton-Jacobi equation, is a nonlinear equation. This would usually be regarded as a drawback, but in this chapter we will take advantage of the nonlinearity to develop a simple approximation method for solving (7.6). Indeed, an iteration procedure is suggested by the fact that u'' is zero for the free particle. We are led to suspect that this second derivative remains relatively small if the potential does not vary too violently. When we omit this term from the equation entirely, we obtain the first crude approximation, u_0, to u:

$$u_0'^2 = [k(x)]^2 \tag{7.7}$$

or, integrating this,

$$u_0 = \pm\int^x k(x)dx + C \tag{7.8}$$

If V is constant, (7.8) is an exact solution. If V varies with x, a successive approximation can be set up by casting (7.6) in the form

$$\left(\frac{du}{dx}\right)^2 = k^2(x) + i\frac{d^2u}{dx^2} \tag{7.9}$$

If we substitute the n-th approximation on the right-hand side of this equation, we obtain the $(n + 1)$-th approximation by a mere quadrature:

$$u_{n+1}(x) = \pm\int^x \sqrt{k^2(x) + i\,u_n''(x)}\,dx + C_{n+1} \tag{7.10}$$

Thus, we have for $n = 0$,

$$\begin{aligned} u_1(x) &= \pm\int^x \sqrt{k^2(x) + i\,u_0''(x)}\,dx + C_1 \\ &= \pm\int^x \sqrt{k^2(x) \pm i\,k'(x)}\,dx + C_1 \end{aligned} \tag{7.11}$$

The two different signs in (7.8), (7.10), and (7.11) give approximations to two particular solutions of (7.6). If we denote these by u_+ and u_-, the general solution of (7.6) is expressible as

$$\begin{aligned} u(x) &= u_+ - i\ln[1 + Ae^{i(u_- - u_+)}] + B \\ &= u_- - i\ln[A + e^{i(u_+ - u_-)}] + B \end{aligned} \tag{7.12}$$

where A and B are arbitrary constants, as befits a second-order ordinary differential equation. The corresponding solution of the Schrödinger equation is

$$\psi(x) = e^{iu(x)} = e^{iB}e^{iu_+(x)} + Ae^{iB}e^{iu_-(x)} \tag{7.13}$$

which is a simple superposition of two particular approximate solutions of (7.1).

Our hope that the approximation procedure (7.10) will tend toward the correct $u(x)$ is baseless unless $u_1(x)$ is close to $u_0(x)$, that is, unless

$$\boxed{|k'(x)| \ll |k^2(x)|} \tag{7.14}$$

In (7.11) both signs must be chosen the same as in the u_0 on which u_1 is supposed to be an improvement. If condition (7.14) holds, we may expand the integrand and obtain

$$\begin{aligned} u_1(x) &= \int^x \left[\pm k(x) + \frac{i}{2} \frac{k'(x)}{k(x)} \right] dx + C_1 \\ &= \pm \int^x k(x)\, dx + \frac{i}{2} \ln k(x) + C_1 \end{aligned} \tag{7.15}$$

The constant of integration is of no significance, because it only affects the normalization of $\psi(x)$, which, if needed at all, is best accomplished after the desired approximations have been made.

The approximation (7.15) to (7.6) is known as *WKB approximation*. It leads to the approximate *WKB wave function*

$$\boxed{\psi(x) \approx \frac{1}{\sqrt{k(x)}} \exp\left[\pm i \int^x k(x) dx \right]} \tag{7.16}$$

In a classically accessible region where $E > V(x)$ and $k(x)$ is real, the two waves (7.16) propagate independently in opposite directions. If the WKB approximation is valid, the potential changes so slowly with x that as the waves propagate no reflected (scattered) wave is generated.

Condition (7.14) for the validity of the WKB approximation can be formulated in ways that are better suited to physical interpretation. If $k(x)$ is regarded as the effective wave number, we may for $E > V(x)$ define an effective wavelength

$$\lambda(x) = \frac{2\pi}{k(x)}$$

The convergence criterion (7.14) can then be cast in the form

$$\left| \frac{d\lambda}{dx} \right| \ll 1 \tag{7.17}$$

requiring the wavelength to vary only slowly. Condition (7.14) can also be written as

$$\lambda(x) \left| \frac{dp}{dx} \right| \ll |p(x)|$$

or

$$\lambda(x) \left| \frac{dV}{dx} \right| \ll \frac{[p(x)]^2}{m}$$

implying that the change of the ''local momentum'' $p(x) = \hbar k(x)$ over a wavelength must be small by comparison with the momentum itself, or that the potential energy change over a wavelength is much less than the local kinetic energy.

These conditions obviously break down if $k(x)$ vanishes or if $k(x)$ varies very rapidly. This certainly happens at the *classical turning points* for which

$$V(x) = E \tag{7.18}$$

or whenever $V(x)$ has a very steep behavior. Generally, in these regions we expect that waves propagating in one direction will generate reflected (scattered) waves. A more accurate solution must be used in a region where (7.14) breaks down. The WKB method is not particularly useful unless we find ways to extend the wave function into and through these regions.

In the nonclassical domain, where $E < V(x)$, it is appropriate to rewrite the WKB wave function (7.16) in its more recognizably real form

$$\boxed{\psi(x) = \frac{1}{\sqrt{\kappa(x)}} \exp\left[\pm \int^x \kappa(x)dx\right]} \tag{7.19}$$

The so-called *connection formulas* serve to link WKB solutions of the type (7.16) in the classically accessible region of x with solutions of type (7.19) in the classically inaccessible region.

2. ***The Connection Formulas.*** Suppose that $x = a$ is a classical turning point for the motion with the given energy, E, as shown in Figure 7.1*a*. The point $x = a$ separates the two regions where $E > V$ and $E < V$, when the classically impenetrable barrier lies to the right of the classical turning point. Analogous considerations hold if the barrier is to the left of the turning point $x = b$ (Figure 7.1b). The results for the two cases will be summarized in Eqs. (7.34) and (7.35).

If the WKB approximation can be assumed to be applicable except in the immediate neighborhood of the turning point, we have

$$\psi(x) \approx \frac{A}{\sqrt{\kappa(x)}} \exp\left[-\int_a^x \kappa(x)dx\right] + \frac{B}{\sqrt{\kappa(x)}} \exp\left[\int_a^x \kappa(x)dx\right] \qquad \text{for } x \gg a \tag{7.20}$$

and

$$\psi(x) \approx \frac{C}{\sqrt{k(x)}} \exp\left[-i\int_a^x k(x)dx\right] + \frac{D}{\sqrt{k(x)}} \exp\left[i\int_a^x k(x)dx\right] \qquad \text{for } x \ll a \tag{7.21}$$

The lower limits of the integrals in the exponents have been arbitrarily, but conveniently, chosen to make the meaning of the amplitudes A, B, C, and D unambiguous.

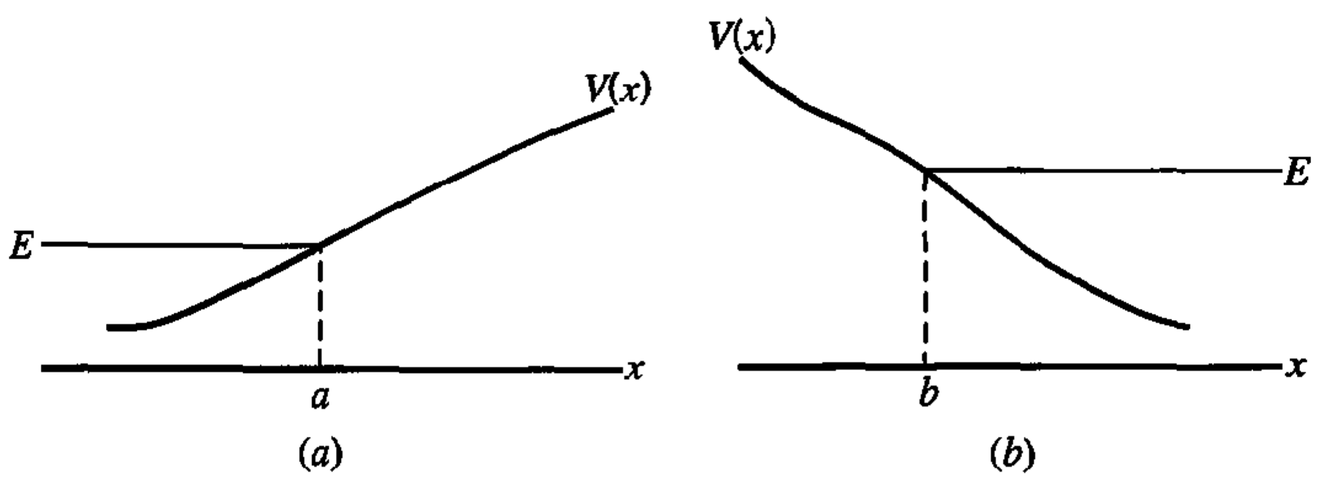


Figure 7.1. (a) Classical turning point at $x = a$, to the right of the classically accessible region. (b) Classical turning point at $x = b$, to the left of the classically accessible region.

We now ask the fundamental question: How are the coefficients C and D related to A and B if (7.20) and (7.21) are to represent the *same state*, albeit in different regions? The inadequacy of the WKB approximation near the turning point is evident, since $k(x) \to 0$ implies an unphysical divergence of $\psi(x)$. To establish the connection between the two separated regions, we must solve the Schrödinger equation more accurately than by the WKB approximation. This can always be done numerically, but an analytic approach works if a somewhat special, yet often appropriate, assumption is made about the behavior of the potential energy near the turning point.[1]

We suppose that in the neighborhood of $x = a$, we may write

$$V(x) - E \approx g(x - a) \tag{7.22}$$

where $g > 0$. The Schrödinger equation for this linear potential,

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + g(x - a)\psi = 0 \tag{7.23}$$

is conveniently transformed by substitution of the dimensionless variable z,

$$z = \left(\frac{2mg}{\hbar^2}\right)^{1/3}(x - a) \tag{7.24}$$

into the form

$$\frac{d^2\psi}{dz^2} - z\psi = 0 \tag{7.25}$$

Note that

$$k^2 = \left(\frac{2mg}{\hbar^2}\right)^{2/3} z \tag{7.26}$$

and that the WKB condition (7.14) implies

$$|z|^{3/2} \gg 1/2$$

Thus, the WKB approximation is simply an asymptotic approximation for the solutions of Eq. (7.23), applicable where $|z|$ becomes large.

The solutions of the differential equation (7.25) are the *Airy functions*[2] $Ai(z)$ and $Bi(z)$. Asymptotically, for large positive z, the leading terms are:

$$Ai(z) \cong \frac{1}{2}\pi^{-1/2}z^{-1/4}e^{-\zeta} \tag{7.27}$$

and

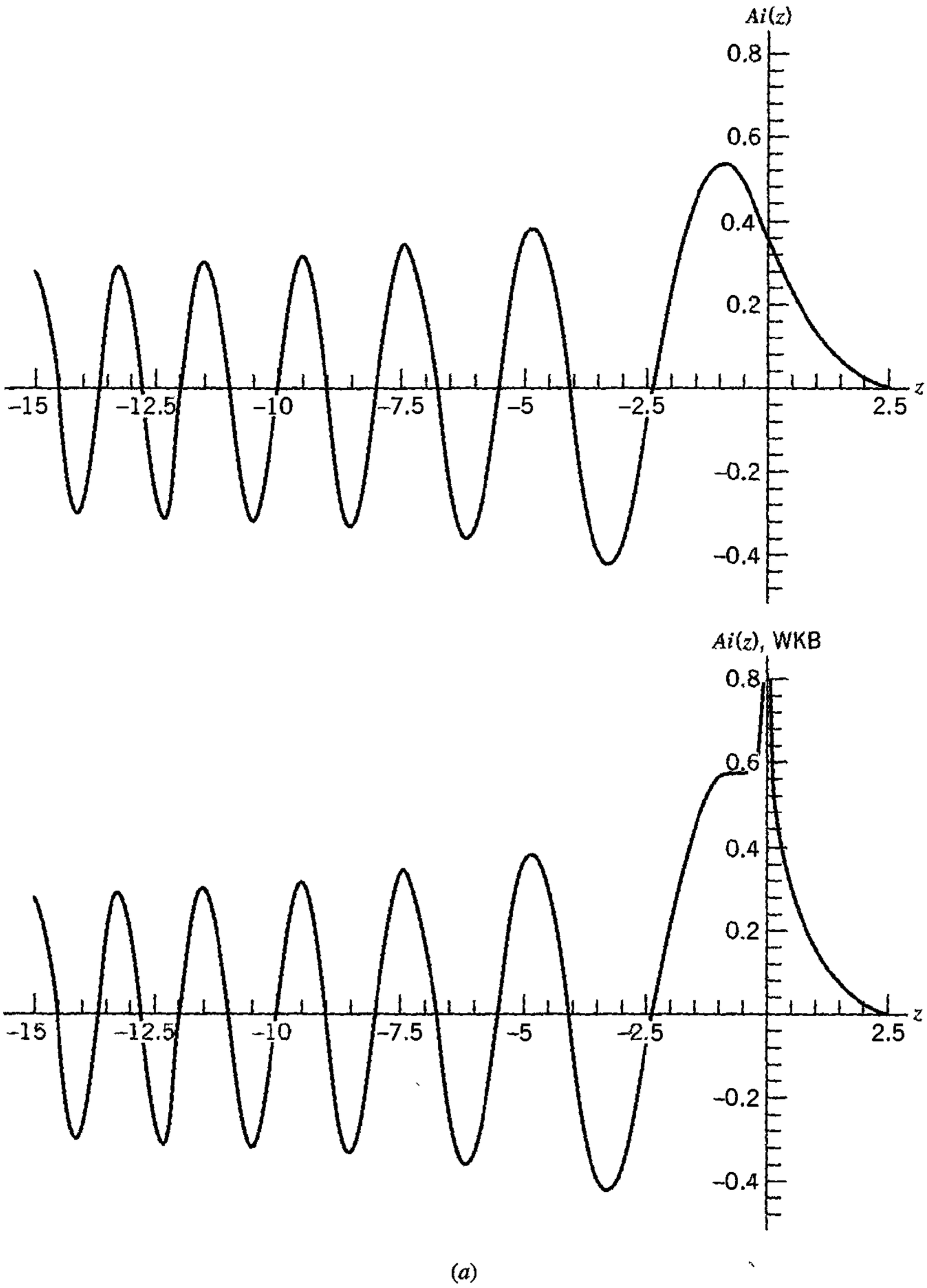
$$Bi(z) \cong \pi^{-1/2}z^{-1/4}e^{\zeta} \tag{7.28}$$

where

$$\zeta = \frac{2}{3}|z|^{3/2} \tag{7.29}$$

[1]For a more general treatment of the WKB approximation at the turning points, and useful historical references, see Schiff (1968), Section 34.

[2]Abramowitz and Stegun (1964), Section 10.4.



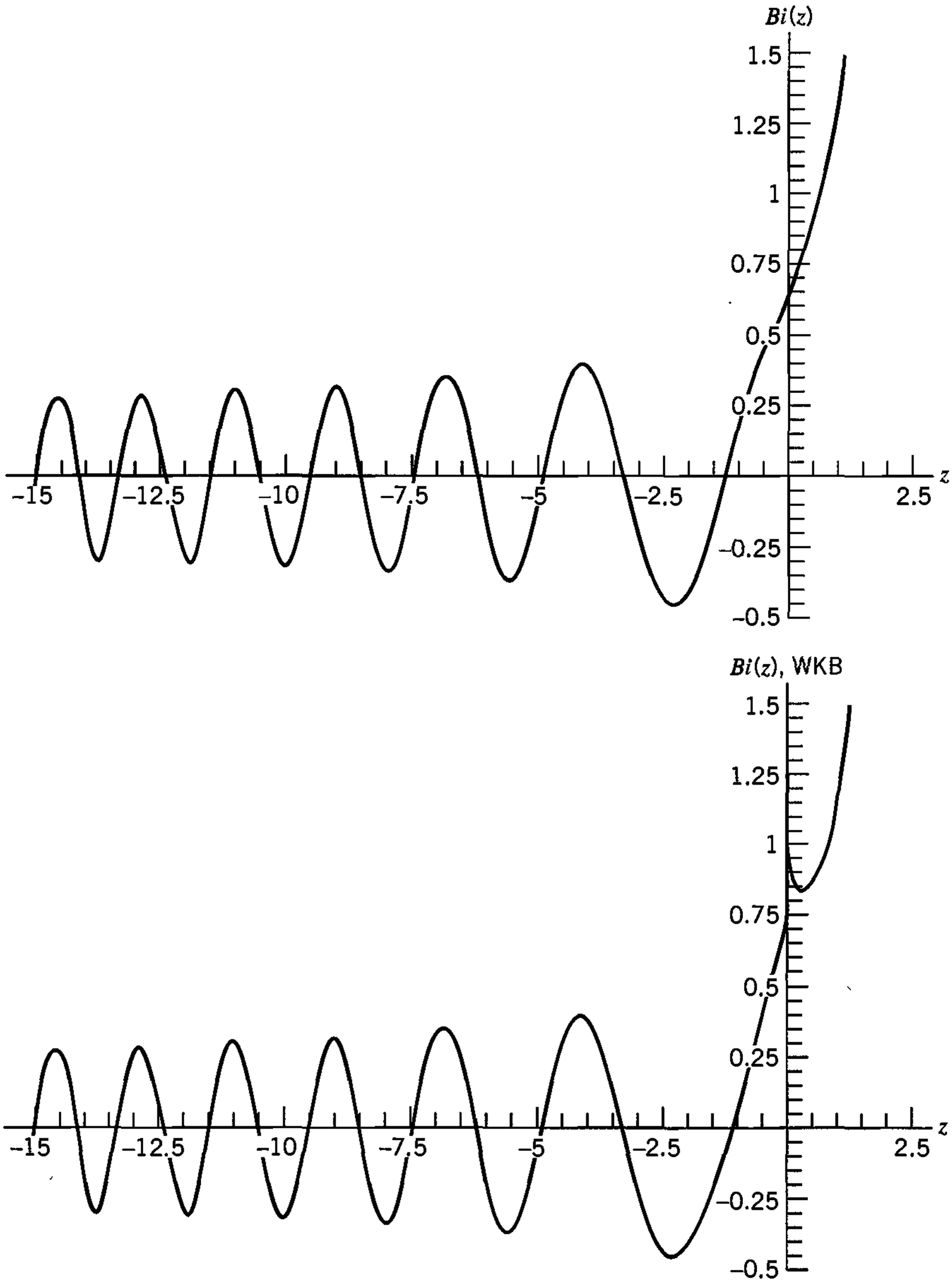
(*a*)

Figure 7.2. (a) The Airy function $Ai(z)$, and (b) the Airy function $Bi(z)$, and their asymptotic (WKB) approximations, for real-valued z. The approximations diverge at $z = 0$.

For large negative values of z:

$$Ai(z) \cong \pi^{-1/2}|z|^{-1/4} \cos\left(\zeta - \frac{\pi}{4}\right) \tag{7.30}$$

$$Bi(z) \cong -\pi^{-1/2}|z|^{-1/4} \sin\left(\zeta - \frac{\pi}{4}\right) \tag{7.3}$$



(*b*)

Figure 7.2. (continued)

Figures 7.2*a* and 7.2*b* are plots of the Airy functions $Ai(z)$ and $Bi(z)$ and their asymptotic (WKB) approximations. Except for $z \approx 0$, where the asymptotic forms diverge, the two functions agree very well.

Since for the assumed linear potential (7.22), for $z < 0$,

$$\int_x^a k(x)dx = \frac{2}{3}\left(\frac{2mg}{\hbar^2}\right)^{1/2}(a-x)^{3/2} = \frac{2}{3}|z|^{3/2} = \zeta \tag{7.32}$$

and for $z > 0$,

$$\int_x^a \kappa(x)dx = \frac{2}{3}\left(\frac{2mg}{\hbar^2}\right)^{1/2}(x-a)^{3/2} = \frac{2}{3}|z|^{3/2} = \zeta \tag{7.33}$$

we verify that the asymptotic forms (7.27)–(7.31) are nothing but WKB wave functions of the type (7.20) and (7.21). By comparing the asymptotic expressions for the Airy functions with (7.20) and (7.21), we learn that the WKB wave functions on the two sides of the turning point $x = a$ are connected as follows:

$$\boxed{\begin{aligned}&\frac{2A}{\sqrt{k(x)}}\cos\left(\int_x^a k(x)dx - \frac{\pi}{4}\right) - \frac{B}{\sqrt{k(x)}}\sin\left(\int_x^a k(x)dx - \frac{\pi}{4}\right) \leftrightarrow \\ &\frac{A}{\sqrt{\kappa(x)}}\exp\left(-\int_a^x \kappa(x)dx\right) + \frac{B}{\sqrt{\kappa(x)}}\exp\left(\int_a^x \kappa(x)dx\right)\end{aligned}} \tag{7.34}$$

The analogous connection formulas for a classical turning point $x = b$, which separates a classically inaccessible region $x < b$ from the accessible region $x > b$, are

$$\boxed{\begin{aligned}&\frac{A}{\sqrt{\kappa(x)}}\exp\left(-\int_x^b \kappa(x)dx\right) + \frac{B}{\sqrt{\kappa(x)}}\exp\left(\int_x^b \kappa(x)dx\right) \leftrightarrow \\ &\frac{2A}{\sqrt{k(x)}}\cos\left(\int_b^x k(x)dx - \frac{\pi}{4}\right) - \frac{B}{\sqrt{k(x)}}\sin\left(\int_b^x k(x)dx - \frac{\pi}{4}\right)\end{aligned}} \tag{7.35}$$

Exercise 7.1. By making the transformation $x - a \to b - x$ which turns Figure 7.1(*a*) into Figure 7.1(*b*), show that the connection formula (7.35) follows from (7.34).

Exercise 7.2. By comparing (7.34) with (7.20) and (7.21), derive the coefficients C and D in terms of A and B.

Caution must be exercised in employing these WKB connection formulas. To see this, suppose that we know that in the region $x \gg a$ in Figure 7.1(*a*) the wave function is adequately represented just by the increasing exponential. Since in this region the decreasing exponential is naturally much smaller than the increasing exponential, the contribution of the decreasing exponential to the wave function may be dwarfed even if the coefficient A is comparable in magnitude to B. Neglecting the decreasing exponential could then lead to a gross error in the WKB estimate in the classically allowed region $x \ll a$.

Conversely, if B is finite but $|B| \ll |A|$, neglecting it entirely is justified in the region $x \ll a$, but as the coefficient of the increasing exponential component of the wave function, this "small" contribution may dominate the behavior of the wave function in the region $x \gg a$.

The connection formulas, which link WKB wave functions between different regions of real and imaginary $k(x)$, break down if two neighboring classical turning points are so close to each other that there is no WKB region between them. This happens, for instance, when the energy E is close to an extremum of the potential. It is then necessary to resort to more accurate analytical or numerical solutions of the Schrödinger equation.

Finally, we observe that the WKB approximation presumes that $k(x)$ is an analytic function. This property fails not only at the classical turning points, but also at the singularities of the potential $V(x)$. Care is required in continuing the WKB wave function through such singularities.

Exercise 7.3. Show that the WKB approximation is consistent with the generalized continuity equation (3.7), and thus with conservation of probability, even across classical turning points.[3]

3. ***Application to Bound States.*** The WKB approximation can be applied to derive an equation for the energies of bound states. The basic idea emerges if we choose a simple well-shaped potential with two classical turning points as shown in Figure 7.3. The WKB approximation will be used in regions 1, 2, and 3 away from the turning points, and the connection formulas will serve near $x = a$ and $x = b$. The usual requirement that ψ must be finite dictates that the solutions which increase exponentially as one moves outward from the turning points must vanish. Thus, to satisfy the boundary condition as $x \to -\infty$, the unnormalized WKB wave function in region 1 is

$$\psi \approx \frac{1}{\sqrt{\kappa}} \exp\left(-\int_x^b \kappa(x)\, dx\right) \qquad \text{for } x < b$$

Hence, by Eq. (7.35), with $B = 0$, in region 2,

$$\psi \approx \frac{2}{\sqrt{\kappa}} \cos\left(\int_b^x k(x)\, dx - \frac{\pi}{4}\right) \qquad \text{for } b < x < a$$

This may also be written as

$$\begin{aligned}\psi &\approx \frac{2}{\sqrt{k(x)}} \cos\left(\int_b^a k(x)\, dx - \int_x^a k(x)\, dx - \frac{\pi}{4}\right) \\ &= -\frac{2}{\sqrt{k(x)}} \cos\left(\int_b^a k(x)\, dx\right) \sin\left(\int_x^a k(x)\, dx - \frac{\pi}{4}\right) \\ &\quad + \frac{2}{\sqrt{k(x)}} \sin\left(\int_b^a k(x)\, dx\right) \cos\left(\int_x^a k(x)\, dx - \frac{\pi}{4}\right)\end{aligned}$$

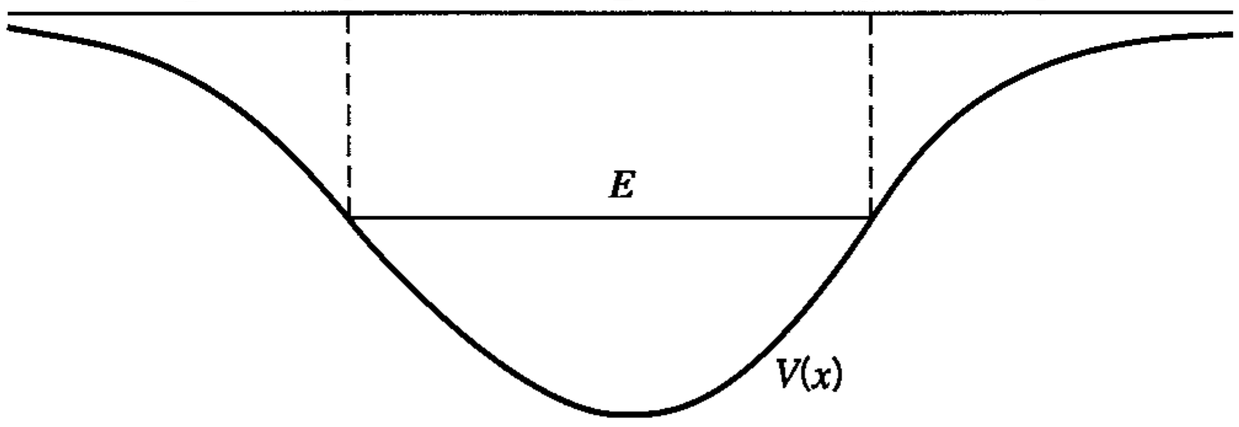


Figure 7.3. Simple one-dimensional potential well, Classically, a particle of energy E is confined to the region between a and b.

[3]This proposition is equivalent to the constancy of the Wronskian for the Schrödinger equation (7.1), which is of the Sturm-Liouville type. See Bradbury (1984), Chapter 7, Section 10.

By (7.34) only the second of these two terms gives rise to a decreasing exponential in region 3 satisfying the boundary conditions as $x \to +\infty$. Hence, the first term must vanish. We obtain the condition

$$\boxed{\int_b^a k(x)\,dx = \left(n + \frac{1}{2}\right)\pi} \tag{7.36}$$

where $n = 0, 1, 2, \ldots$. This equation determines the possible discrete values of E. The energy E appears in the integrand as well as in the limits of integration, since the turning points a and b are determined such that $V(a) = V(b) = E$.

If we introduce the classical momentum $p(x) = \pm\hbar k(x)$ and plot $p(x)$ versus x in phase space, the bounded motion in a potential well can be pictured by a closed curve (Figure 7.4). It is then evident that condition (7.36) may be written as

$$\boxed{J = \oint p(x)\,dx = \left(n + \frac{1}{2}\right)h} \tag{7.37}$$

This equation is very similar to the quantum condition (1.2) in the old quantum theory, which occupied a position intermediate between classical and quantum mechanics.

The expression (7.37) equals the area enclosed by the curve representing the periodic motion in phase space and is called the *phase integral* J in classical terminology. If the WKB approximation is used all the way from b to a, (7.36) measures the phase change that the oscillatory wave function ψ undergoes in region 2 across the well between the two turning points. Dividing this by 2π, we see that according to the WKB approximation $n/2 + 1/4$ quasi-wavelengths fit between b and a. Hence, n represents the number of nodes in the wave function, a fact that helps to visualize the elusive ψ.

In Figure 7.4, the area of enclosed phase space between the closed curves for $n+1$ and n is equal to h. As for free particles (Section 4.4), in the WKB approx-

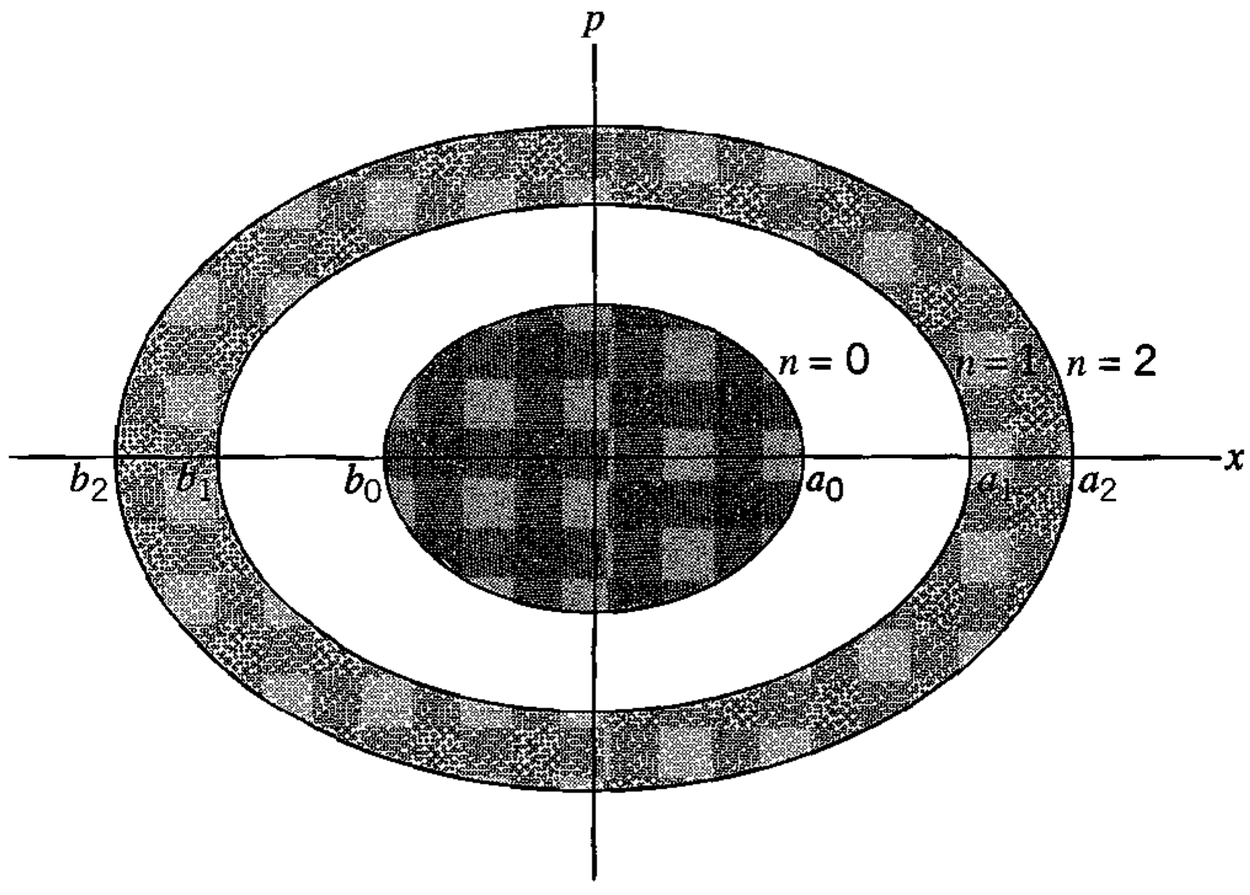


Figure 7.4. Phase space representation of the periodic motion of a particle confined between the classical turning points at $x = a_n$ and $x = b_n$ for the bound-state energies $E_n = V(a_n) = V(b_n)$. The area between two neighboring trajectories is equal to h.

imation each quantum state may be said to occupy a volume h in two-dimensional phase space. In statistical mechanics, this rule is useful in the domain where classical mechanics is applicable but some concession must be made to the quantum structure of matter.

The WKB method is fittingly called a *semiclassical approximation*, because for high energies ψ has a very short wavelength in the classically accessible region. It is a rapidly oscillating function of position, but its maximum amplitude is modulated slowly by a factor $1/\sqrt{k(x)}$. The probability, $|\psi|^2\,dx$, of finding the particle in an interval dx at x is proportional to the reciprocal of the classical velocity, $1/v(x) \propto [E - V(x)]^{-1/2}$. Classically, this is proportional to the length of time (dt) that the particle spends in the interval dx. It thus represents the relative probability of finding the particle in the interval dx if a random (in time) determination of its position is made as the particle shuttles back and forth between the turning points. We thus see that the probability concepts used in quantum and classical mechanics, though basically different, are nevertheless related to each other in the limit in which the rapid phase fluctuations of quantum mechanics can be legitimately averaged to give the approximate classical behavior.

As an illustrative example, we compare the exact solutions of the Schrödinger equation with approximate WKB energies and wave functions for the bound states of a particle of mass m in a potential well defined by $V(x) = g|x|$. The strength of the potential is measured by the positive constant g. Such a V-shaped potential is the one-dimensional analogue of the linear central-force potential, $V(r) = Cr$, to which the confinement of the quark–antiquark constituents of the charmonium ''atom'' is attributed.

It is sufficient to solve the Schrödinger equation for the one-dimensional *linear* potential,

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + g|x|\psi = E\psi \tag{7.38}$$

for $x > 0$ and characterize the even and odd parity solutions by imposing the boundary conditions at the coordinate origin:

$$\psi_e'(0) = 0 \qquad \psi_o(0) = 0 \tag{7.39}$$

For $x > 0$, (7.38) is the same as Eq. (7.23), if we identify the energy $E = ga$. As in Section 7.2, the substitution

$$z = \left(\frac{2mg}{\hbar^2}\right)^{1/3}\left(x - \frac{E}{g}\right) = \left(\frac{2mg}{\hbar^2}\right)^{1/3}(x - a) \qquad (x > 0) \tag{7.40}$$

produces the differential equation for $\psi(z)$:

$$\frac{d^2\psi}{dz^2} - z\psi = 0 \tag{7.41}$$

The boundary condition at large x requires that $\psi(z)$ must vanish asymptotically as $z \to +\infty$. This condition implies that the energy eigenfunction must be the *Airy function* $Ai(z)$. It is interesting to observe the universality of the differential equation (7.41), which is related to the original Schrödinger equation (7.38) by scaling and displacing the independent variable. There are no parameters in (7.41), and the graph of the solution in Figure 7.2 can be used for all values of the potential strength, g, and for all bound-state energies, E.

The *unnormalized* energy eigenfunctions are $\psi(z) = Ai(z)$, or in terms of the original x coordinate ($x > 0$):

$$\psi_n(x) = Ai\left[\left(\frac{2mg}{\hbar^2}\right)^{1/3}\left(x - \frac{E_n}{g}\right)\right] \tag{7.42}$$

The conditions (7.39) at the origin now require that for the even eigenfunctions

$$\psi_n'(0) = Ai'\left[-\left(\frac{2mg}{\hbar^2}\right)^{1/3}\frac{E_n}{g}\right] = 0 \tag{7.43}$$

and for the odd eigenfunctions

$$\psi_n(0) = Ai\left[-\left(\frac{2mg}{\hbar^2}\right)^{1/3}\frac{E_n}{g}\right] = 0 \tag{7.44}$$

The (negative) zeros of $Ai(z)$ and $Ai'(z)$ may be computed or read off Figure 7.2(a).[4] The five lowest energy levels are listed in the second column in Table 7.1.

In the WKB approximation, the approximate energies are extracted from the condition (7.36), which in the present context translates into

$$2\int_0^{E/g}\sqrt{\frac{2m}{\hbar^2}(E - gx)}\,dx = \left(n + \frac{1}{2}\right)\pi$$

From this condition we obtain the WKB estimate:

$$(E_{WKB})^3 = \frac{9}{32}\pi^2\left(n + \frac{1}{2}\right)^2\frac{g^2\hbar^2}{m} \tag{7.45}$$

The values of E_{WKB} for $n = 0$ to 4 are entered in the third column of Table 7.1. Except for the ground state energy, the agreement with the "exact" numerical values is seen to be excellent.

Table 7.1

n	E_n *in units of* $(g^2\hbar^2/m)^{1/3}$	E_{WKB} *in units of* $(g^2\hbar^2/m)^{1/3}$
0	0.8086 (ground state)	0.8853
1	1.8558	1.8416
2	2.5781	2.5888
3	3.2446	3.2397
4	3.8257	3.8306

If the coordinate x and the length $\frac{E_0}{g}$, are scaled in units of $\left(\frac{\hbar^2}{mg}\right)^{1/3}$, and are thus made dimensionless, the unnormalized energy eigenfunctions functions (7.42) are, for $x > 0$, expressed as

$$\psi_n = Ai\left[\sqrt[3]{2}\left(x - \frac{E_n}{g}\right)\right]$$

[4]For numerical tables, see Abramowitz and Stegun (1964), Table 10.11.

These functions can be evaluated from Figure 7.2(a), if the substitution $z = \sqrt[3]{2}\left(x - \frac{E_n}{g}\right)$ is made. (The ground state eigenfunction for the potential $V(x) = g|x|$ is shown in Figure 8.1.)

Exercise 7.4. For the potential $V(x) = g|x|$, compute (or use Figure 7.2(a) to obtain) the energy eigenfunctions for $n = 0$ to 4 and plot them. For $n = 3$ and 4 compare the exact eigenfunctions with the WKB wave functions. For these two energy levels, sketch the exact quantum mechanical, the WKB, and the classical probability densities of finding the particle at position x. Explain why the WKB wave functions have a discontinuous slope at the origin for even values of n.

Exercise 7.5. In a crude model, the S states of the charmonium ''atom'' are regarded as the energy levels of a charmed quark and antiquark, bound by a one-dimensional potential that (except near the origin) is represented by $V(x) = g|x|$. Noting that the two lowest S states have measured rest energies of 3.1 and 3.7 GeV, respectively, and ignoring relativistic effects on the binding, obtain an estimate for the potential strength parameter g in units of GeV/fm. (Treat this two-body problem as an effective one-body problem with a reduced quark mass, i.e., half the rest mass of the charmed quark. It is useful to remember that $\hbar c \approx 0.2$ GeV $\cdot$ fm.)

Exercise 7.6. Show that the WKB approximation gives the energy levels of the linear harmonic oscillator correctly. Compute and plot the WKB approximation to the eigenfunctions for $n = 0$ and 1, and compare with the exact stationary state wave functions.

The WKB method, with its connection formulas, is relatively straightforward for systems that are described by or reducible to a one-dimensional Schrödinger equation. The Hamiltonian of a multidimensional system need not be particularly exotic for the trajectories in phase space to display a far more complicated character than illustrated in Figure 7.4. When the classical system is integrable and its motion multiply periodic, it is possible to generalize the WKB method if due attention is paid to the singularities that are the analogues of classical turning points, but in practice one deals frequently with systems exhibiting classically chaotic motion. Semiclassical quantum mechanics for complex systems has benefited greatly from recent advances in (nonlinear) classical dynamics and constitutes a subject beyond the scope of this book.[5]

4. ***Transmission Through a Barrier.*** The WKB method will now be applied to calculate the transmission coefficient for a barrier on which particles are incident from the left with insufficient energy to pass to the other side classically. This problem is very similar to that of the rectangular potential barrier, Section 6.2, but no special assumption will be made here concerning the shape of the barrier.

[5]Gutzwiller (1990).

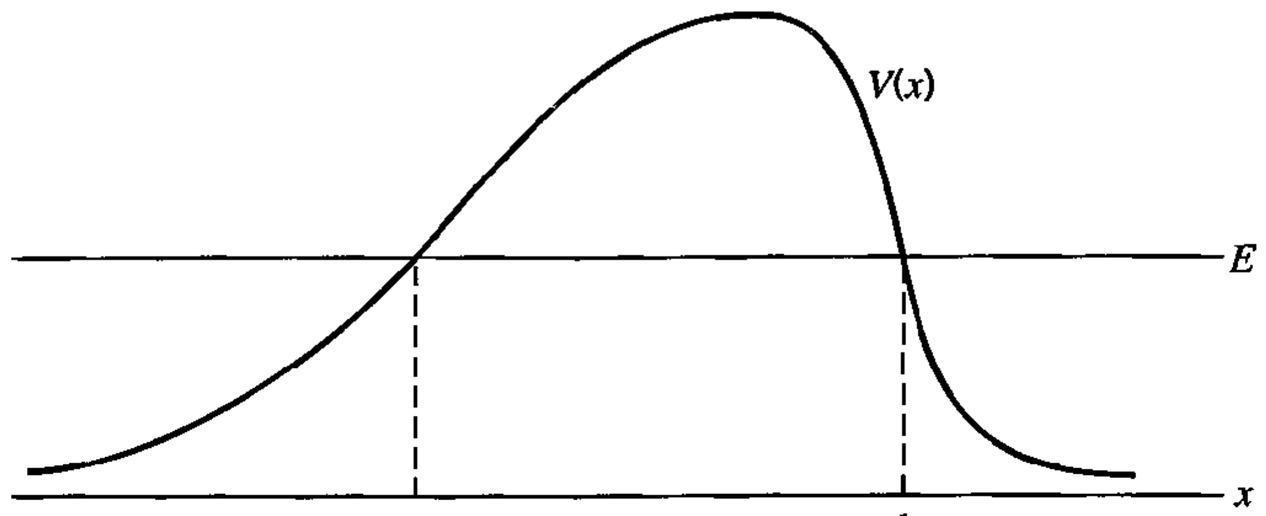


Figure 7.5. Potential barrier.

If the WKB approximation is assumed to hold in the three regions indicated in Figure 7.5, the solution of the Schrödinger equation may be written as

$$\psi(x) = \begin{cases} \dfrac{A}{\sqrt{k(x)}} \exp\left(i \displaystyle\int_a^x k\, dx \right) + \dfrac{B}{\sqrt{k(x)}} \exp\left(-i \displaystyle\int_a^x k\, dx \right) & (x < a) \\ \dfrac{C}{\sqrt{\kappa(x)}} \exp\left(-\displaystyle\int_a^x \kappa\, dx \right) + \dfrac{D}{\sqrt{\kappa(x)}} \exp\left(\displaystyle\int_a^x \kappa\, dx \right) & (a < x < b) \\ \dfrac{F}{\sqrt{k(x)}} \exp\left(i \displaystyle\int_b^x k\, dx \right) + \dfrac{G}{\sqrt{k(x)}} \exp\left(-i \displaystyle\int_b^x k\, dx \right) & (b < x) \end{cases} \tag{7.46}$$

The connection formulas (7.34) and (7.35) can now be used to establish linear relations between the coefficients in (7.46) in much the same way as was done in Chapter 6 for the rectangular barrier. The result of the calculation is remarkably simple and again is best expressed in terms of a matrix M that connects F and G with A and B.

$$\boxed{\begin{pmatrix} A \\ B \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2\theta + \dfrac{1}{2\theta} & i\left(2\theta - \dfrac{1}{2\theta} \right) \\ -i\left(2\theta - \dfrac{1}{2\theta} \right) & 2\theta + \dfrac{1}{2\theta} \end{pmatrix} \begin{pmatrix} F \\ G \end{pmatrix}} \tag{7.47}$$

where the parameter

$$\boxed{\theta = \exp\left(\int_a^b \kappa(x) dx \right)} \tag{7.48}$$

measures the height and thickness of the barrier as a function of energy.

Exercise 7.7. Verify (7.47).

The *transmission coefficient* is defined as

$$T = \frac{|\psi_{\text{trans}}|^2 v_{\text{trans}}}{|\psi_{\text{inc}}|^2 v_{\text{inc}}} = \frac{|\psi_{\text{trans}} \sqrt{k_{\text{trans}}}|^2}{|\psi_{\text{inc}} \sqrt{k_{\text{inc}}}|^2} = \frac{|F|^2}{|A|^2}$$

assuming that there is no wave incident from the right, $G = 0$. From (7.47) we obtain

$$T = \frac{1}{|M_{11}|^2} = \frac{4}{\left(2\theta + \dfrac{1}{2\theta}\right)^2} \tag{7.49}$$

For a high and broad barrier, $\theta \gg 1$, and

$$\boxed{T \approx \frac{1}{\theta^2} = \exp\left(-2\int_a^b \kappa\, dx\right)} \tag{7.50}$$

Hence, θ is a measure of the opacity of the barrier.

As an example, we calculate θ for a one-dimensional model of a repulsive Coulomb barrier (Figure 7.6), which a charged particle such as a proton (charge $Z_1 e$) has to penetrate to reach a nucleus (charge $Z_2 e$). The essence of the calculation survives the generalization to three dimensions (Section 13.8). Thus, let V be defined for $x < 0$ as

$$V = -\frac{Z_2 Z_2 e^2}{x} \tag{7.51}$$

The turning point a is determined by

$$E = -\frac{Z_1 Z_2 e^2}{a}$$

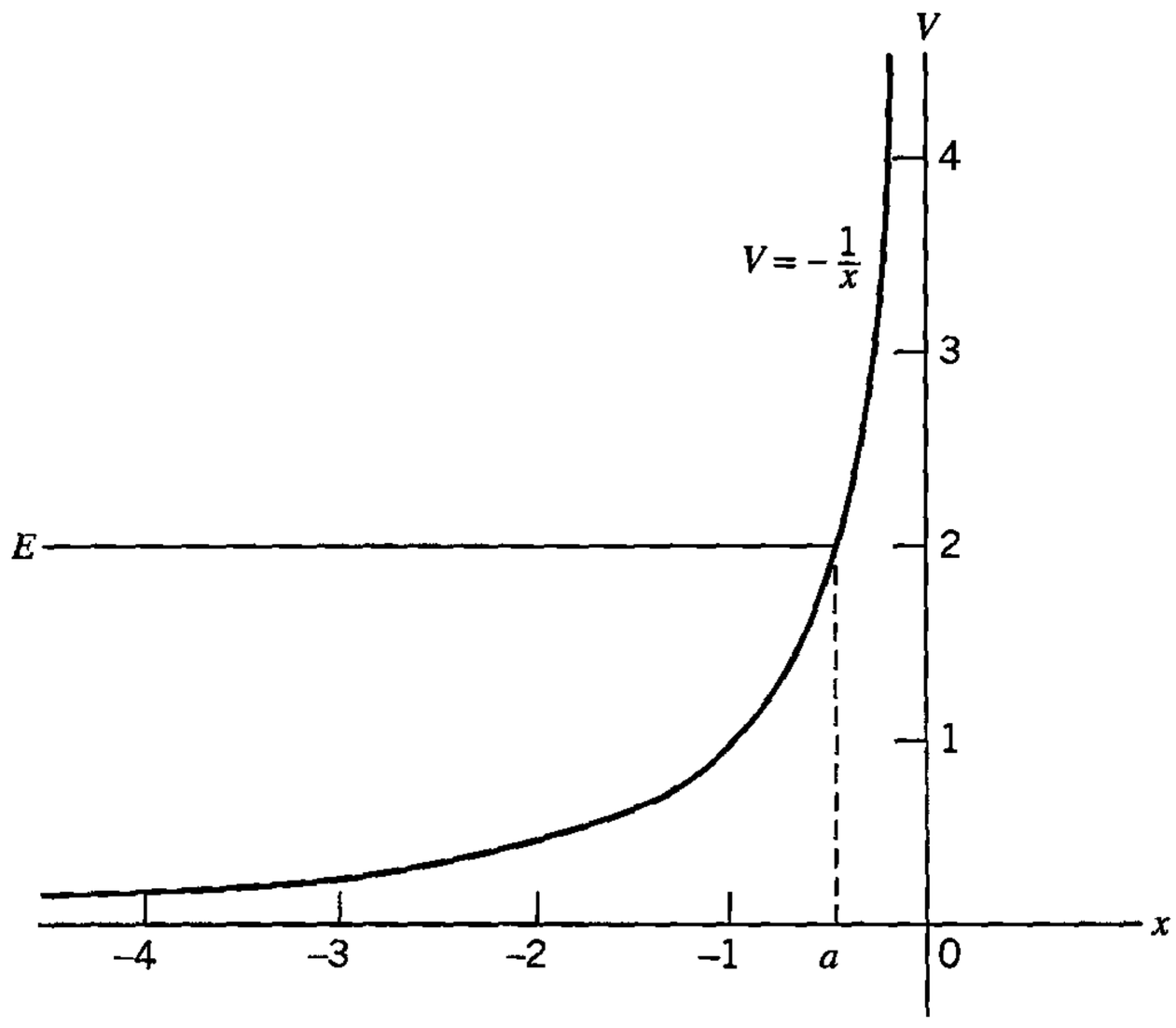


Figure 7.6. One-dimensional analogue of a Coulomb barrier, which repels particles incident from the left.

and we take $b = 0$, callously disregarding the warning signals that caution us about applying the WKB approximation near a singularity of the potential energy. The critical integral is then

$$\int_a^0 \kappa\, dx = \frac{\sqrt{2mE}}{\hbar} \int_a^0 \sqrt{\frac{a}{x} - 1}\, dx = \sqrt{\frac{2m}{E}} \frac{Z_1 Z_2 e^2}{\hbar} \int_0^1 \sqrt{\frac{1}{u} - 1}\, du = \frac{Z_1 Z_2 e^2}{\hbar v}\, \pi$$

where $v = \sqrt{2E/m}$ is the classical particle velocity at $x \to -\infty$. Hence,

$$\boxed{T = \frac{1}{\theta^2} = \exp\left(-\frac{2\pi\, Z_1 Z_2\, e^2}{\hbar v}\right)} \tag{7.52}$$

The barrier inhibits the approach of a positive charged particle to the nucleus, and the transmission coefficient is called the *penetrability*. This quantity determines the probability of nuclear fusion, and it is also decisive in the description of nuclear alpha decay, since the alpha particle, once it is formed inside the nucleus, cannot escape unless it penetrates the surrounding Coulomb barrier.

Exercise 7.8. Calculate the transmission coefficient for the model Coulomb potential (7.51) by assuming, more realistically, that b, instead of being zero is equal to a fraction of a, i.e., $b = \varepsilon a$. Apply the result to the calculation of the Coulomb barrier penetrability for an alpha particle ($Z_1 = 2$) with asymptotic kinetic energy E in the repulsive field of a nucleus (Z_2), with nuclear radius $b = R$. Express E in MeV and R in Fermis.

As a further application of the WKB method, let us consider the passage of a particle through a potential well that is bounded by barriers as shown in Figure 7.7. It will be assumed that $V(x)$ is symmetric about the origin, which is located in the center of the well, and that $V = 0$ outside the interval between $-c$ and c.

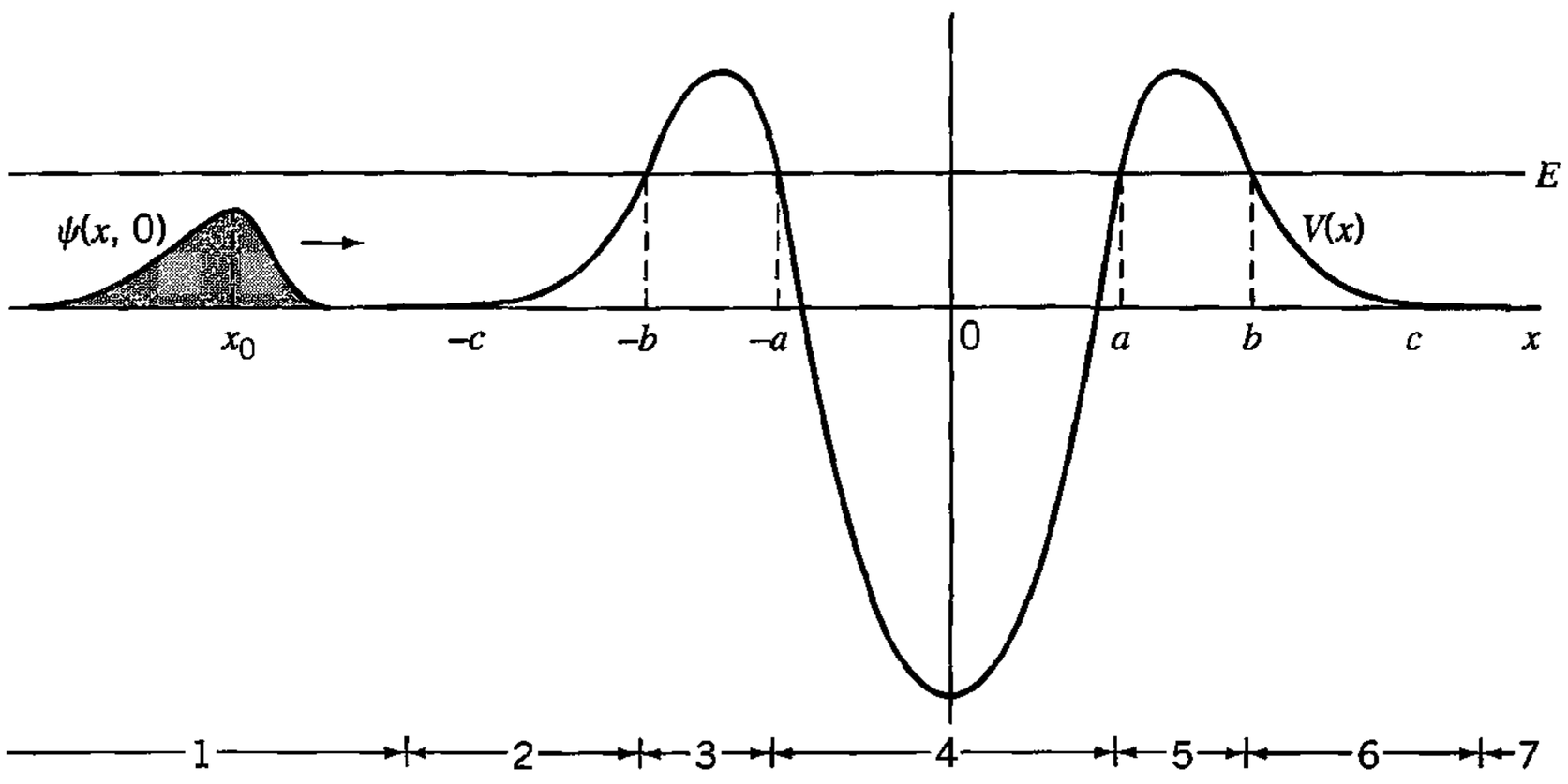


Figure 7.7. Potential barriers surrounding a well are favorable for the occurrence of narrow transmission resonances. Regions 1 through 7 are defined as: (1) $x < -c$; (2) $-c < x < -b$; (3) $-b < x < -a$; (4) $-a < x < a$; (5) $a < x < b$; (6) $b < x < c$; (7) $c < x$. A wave packet is seen to be incident from the left.

In this section, the effect of barrier penetration will be studied for a particle with an energy E below the peak of the barriers. We are particularly interested in the form of the energy eigenfunctions in regions 1 and 7:

$$\psi(x) = \begin{cases} \dfrac{A_1}{\sqrt{k}} \exp(ikx) + \dfrac{B_1}{\sqrt{k}} \exp(-ikx) & (x < -c) \\ \dfrac{A_7}{\sqrt{k}} \exp(ikx) + \dfrac{B_7}{\sqrt{k}} \exp(-ikx) & (x > c) \end{cases} \tag{7.53}$$

When the WKB method is applied to connect the wave function in regions 1 and 7, the relation between the coefficients is again most advantageously recorded in matrix notation:

$$\begin{pmatrix} A_1 \\ B_1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} e^{-2i\rho}\left[\left(4\theta^2 + \dfrac{1}{4\theta^2}\right)\cos L - 2i \sin L\right] & i\left(4\theta^2 - \dfrac{1}{4\theta^2}\right)\cos L \\ -i\left(4\theta^2 - \dfrac{1}{4\theta^2}\right)\cos L & e^{2i\rho}\left[\left(4\theta^2 + \dfrac{1}{4\theta^2}\right)\cos L + 2i \sin L\right] \end{pmatrix} \begin{pmatrix} A_7 \\ B_7 \end{pmatrix} \tag{7.54}$$

In writing these equations, the following abbreviations have been used:[6]

$$L = \int_{-a}^{a} k(x)\, dx, \qquad \rho = \int_{b}^{a} k(x)\, dx - kc \tag{7.55}$$

It follows from the definition of L and from inspection of Figure 7.7 that

$$\frac{\partial L}{\partial E} > 0 \tag{7.56}$$

We will shortly make use of this property.

The final matrix relation (7.54) has the form (6.39) subject to the condition (6.22). This result is expected since, as was pointed out in Section 6.3, the matrix that links the asymptotic parts of the Schrödinger eigenfunction has the same general form for all potentials that are symmetric about the origin.

From Eq. (7.54) we obtain, for $B_7 = 0$,

$$\frac{A_7}{A_1} = \frac{1}{M_{11}} = \sqrt{T}e^{-i\phi} = \frac{e^{2i\rho}}{\dfrac{1}{2}\left(4\theta^2 + \dfrac{1}{4\theta^2}\right)\cos L - i\sin L} \tag{7.57}$$

According to (6.40), the transmission coefficient is

$$T = \frac{1}{|M_{11}|^2} = \frac{4}{\left(4\theta^2 + \dfrac{1}{4\theta^2}\right)^2 \cos^2 L + 4\sin^2 L} \tag{7.58}$$

This quantity reaches its maximum value, unity, whenever $\cos L = 0$, or

$$L = (2n + 1)\, \pi/2 \tag{7.59}$$

The condition determining the location of the transmission peaks is seen to be the same as the quantum condition (7.36) for bound states. If $\theta \gg 1$, so that penetration

[6]Our notation is adapted from a thorough discussion of barrier penetration in Bohm (1951).

through the barriers is strongly inhibited, T has sharp, narrow resonance peaks at these energies. A graph of T in the resonance region will be similar to Figure 6.8.

Under conditions favorable for the occurrence of pronounced resonances ($\theta \gg 1$), it may usually be assumed that in the vicinity of the resonances in a reasonable approximation,

$$\cos L \approx \mp\left(\frac{\partial L}{\partial E}\right)_{E=E_0}(E - E_0), \qquad \sin L \approx \pm 1$$

Substituting these approximations in (7.57) and evaluating the slowly varying quantity θ at $E = E_0$, we get

$$\boxed{\sqrt{T}e^{-i\phi} = \mp\frac{\Gamma/2}{E - E_0 + i\Gamma/2}\, e^{2i\rho}} \tag{7.60}$$

where by definition

$$\boxed{\Gamma = \frac{1}{\theta^2(\partial L/\partial E)_{E=E_0}}} \tag{7.61}$$

Exercise 7.9. Show that the energy spacing D between neighboring resonances is approximately

$$D = \frac{\pi}{\partial L/\partial E} \tag{7.62}$$

and that for low barrier penetration ($\theta \gg 1$), $D \gg \Gamma$, so that the resonances are well separated.

Exercise 7.10. Apply the resonance approximation to the transmission coefficient T, and show that near E_0 it has the characteristic *Breit-Wigner* resonance shape, Γ being its width at half maximum. Compare with Figure 6.8.

Exercise 7.11. A nucleon of energy E is incident on a one-dimensional toy model of a "nucleus" of radius $R = 4$ fm (diameter $2R$). The attractive potential inside the nucleus has a mean depth of $V_0 = 65$ MeV. Approximately rectangular barriers of average 8 fm width and 5 MeV height bound the nucleus symmetrically, so that the potential looks qualitatively like Figure 7.7. Estimate the value of the barrier opacity θ^2 and of L as a function of E. Calculate the energy and width of the lowest nucleon transmission resonance below the barrier. Are there other resonances nearby?

5. ***Motion of a Wave Packet and Exponential Decay.*** It is instructive to consider the motion of a simple broad wave packet incident with well defined positive momentum $\hbar k_0 = \sqrt{2mE_0}$ from the left (from region 1 where $x < -c$) onto the well protected by a barrier, as shown in Figure 7.7. A wave packet, which at $t = 0$ is localized entirely in region 1 near the coordinate $x_0 \ll 0$ far to the left of the barrier and moving toward positive x, may be represented by the wave function in k-space:

$$\phi(k) = |\phi(k)|e^{-ikx_0} \tag{7.62}$$

Here $|\phi(k)|$ is a smoothly varying function with a fairly sharp peak and a width Δk, as exemplified in Figure 2.1, which vanishes for $k < 0$. Though, in the presence of the potential $V(x)$, k is not constant, asymptotically ($|x| \to \infty$) we can use the free-particle relation $\hbar k = \sqrt{2mE}$ to express $\psi(x, 0)$ as an integral over E instead of k. Since the wave packet is narrow in k space, the two variables are related approximately as

$$k = k_0 + (k - k_0) = k_0 + \frac{k^2 - k_0^2}{k + k_0} \approx k_0 + \frac{k^2 - k_0^2}{2k_0} = \frac{k_0}{2} + \frac{E}{\hbar v_0}$$

The initial wave packet can then be written as

$$\begin{aligned} \psi(x, 0) &= \frac{1}{\sqrt{2\pi}} \int_0^\infty |\phi(k)| e^{ik(x-x_0)}\, dk \\ &= \frac{1}{\sqrt{2\pi}\,\hbar v_0} \int_0^\infty |\phi(k)| e^{ik(x-x_0)}\, dE = \int_0^\infty f(E) e^{ik(x-x_0)}\, dE \end{aligned} \tag{7.63}$$

The amplitude $f(E)$, defined in (7.63), is a smoothly varying positive function of E with a fairly sharp peak and a width $\Delta E \approx v_0 \hbar \Delta k$. If $\psi(x, 0)$ is normalized to unity, $f(E)$ satisfies the normalization condition

$$\int_0^\infty |f(E)|^2 v_0\, dE = \frac{1}{2\pi\hbar} \tag{7.64}$$

The representation (7.63) expresses $\psi(x, 0)$ as a superposition of infinite plane waves, but we need an expansion of $\psi(x, 0)$ in terms of the appropriate WKB wave functions, whose asymptotic form is given by (7.53). Since no wave is incident from the right, we include only WKB wave functions with $B_7 = 0$. In the asymptotic region 1, the expansion has the form

$$\psi(x, 0) = \int_0^\infty f(E) e^{-ikx_0} \left(e^{ikx} + \frac{B_1}{A_1} e^{-ikx} \right) dE \qquad (x < -c) \tag{7.65}$$

where the k-dependent coefficients A_1 and B_1 are the same as those that appear in (7.53). Equation (7.65) holds because the integral

$$\int_0^\infty f(E) \frac{B_1}{B_2} e^{-ik(x+x_0)} dE \tag{7.66}$$

differs from zero only when the phase in the integrand is stationary, i.e., for values of x near $x = -x_0$. Elsewhere it vanishes owing to the rapid oscillations of the exponential function. Hence, it vanishes in region 1.

In the asymptotic region 7 to the right of the barrier, the transmitted wave function at arbitrary times $t > 0$ is:

$$\psi(x, t) = \int_0^x f(E) \sqrt{T} \exp(-i\phi) \exp\left[ik(x - x_0) - \frac{i}{\hbar} Et \right] dE \qquad (x > c) \tag{7.67}$$

In order to study the behavior of this transmitted wave packet near a very narrow resonance we assume that the mean energy E_0 of the incident wave packet corresponds to a resonance. We also assume that the width ΔE of the packet considerably exceeds the width of the resonance (but is much smaller than the interval between neighboring resonances). We therefore are entitled to substitute (7.60) into the integrand of (7.67).

Except for uncommonly long-range potential barriers, the phase ρ may be assumed constant, and equal to ρ_0, over the width of the resonance. With these approximations, the wave function in region 7 at positive t becomes

$$\psi(x, t) \approx \mp f(E_0)\exp\left[i\frac{k_0}{2}(x - x_0) + 2i\rho_0 \right] \frac{\Gamma}{2} \int_{-\infty}^{\infty} \frac{\exp\left[\frac{i}{\hbar} E\left(\frac{x - x_0}{v_0} - t \right) \right]}{E - E_0 + i\Gamma/2} dE \tag{7.68}$$

In (7.68), the integration has been extended to $-\infty$ without appreciable error, assuming that t is not too large. (For extremely long times $t \to \infty$, the results to be derived may therefore be suspect.)

The integral in (7.68) is a well-known Fourier integral that is evaluated in the Appendix, Eq. (A.22). The result is that in the asymptotic region $x > c$,

$$\psi(x, t) \approx \begin{cases} \pm \pi i \Gamma f(E_0) e^{2i\rho_0} \exp\left[\frac{\Gamma}{2\hbar} \left(\frac{x - x_0}{v_0} - t \right) \right] \exp[ik_0(x - x_0)] e^{-iE_0 t/\hbar} & \text{if } t > \dfrac{x - x_0}{v_0} \\ 0 & \text{if } t < \dfrac{x - x_0}{v_0} \end{cases} \tag{7.69}$$

This wave function describes a wave packet with a discontinuous front edge at $x = x_0 + v_0 t$ and an exponentially decreasing tail to the left. After the pulse arrives at a point x the probability density decays according to the formula

$$|\psi(x, t)|^2 = \pi^2 \Gamma^2 |f(E_0)|^2 \exp\left[\frac{\Gamma}{\hbar} \left(\frac{x - x_0}{v_0} - t \right) \right] \tag{7.70}$$

Figure 7.8 shows the distorted wave packet at various times after it has reached region 7. We may calculate the total probability that at time t the particle has been

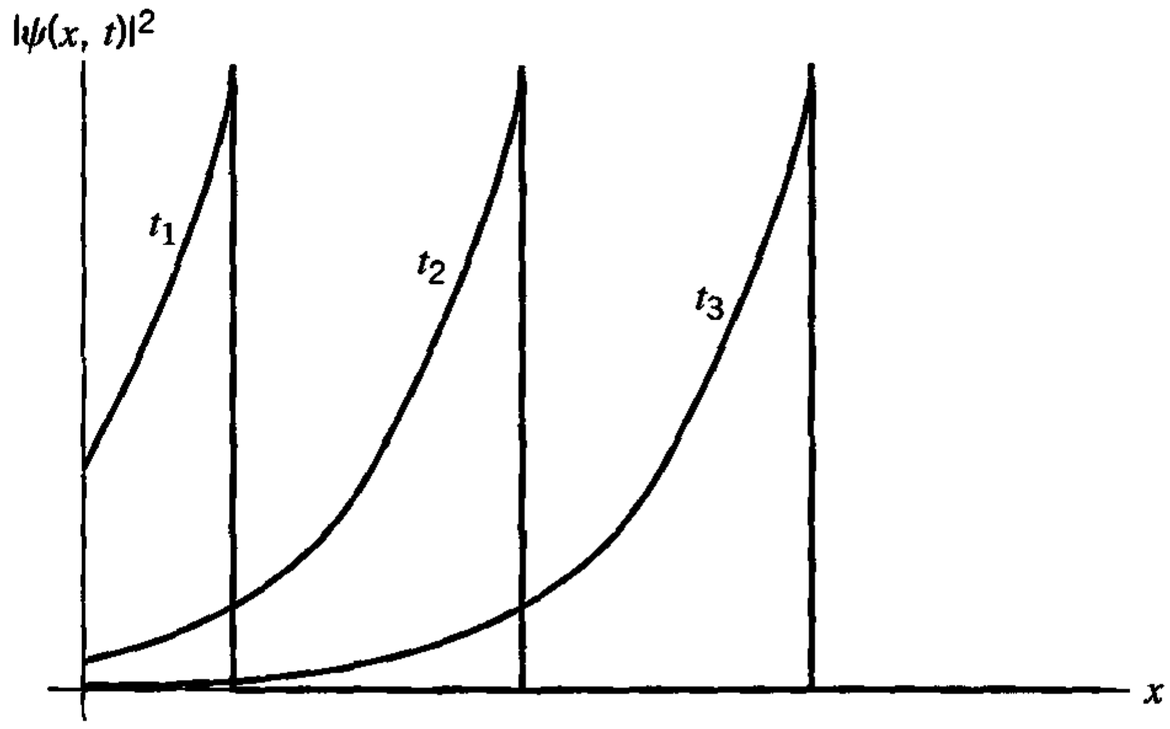


Figure 7.8. The probability density $|\psi(x,t)|^2$ of a decaying state for three different times, $t_1 < t_2 < t_3$. At a fixed position x, after the arrival of the wave front, the probability density decays exponentially in time. At a fixed time, the probability density *increases* exponentially as a function of x (up to the front edge of the wave front).

transmitted and is found anywhere in region 7. For a wave packet whose energy spread ΔE covers a single resonance such that

$$D \gg \Delta E \gg \Gamma \tag{7.71}$$

this probability is

$$\int_0^{x_0+v_0t} |\psi(x, t)|^2 \, dx = \pi^2\hbar\Gamma v_0 |f(E_0)|^2 \left\{1 - \exp\left[-\frac{\Gamma}{\hbar}\left(t + \frac{x_0}{v_0}\right)\right]\right\} \quad \left(t > -\frac{x_0}{v_0}\right)$$

Here we have assumed that $c \ll x_0 + v_0 t$, so that the lower limit of the integral can be set equal to zero. From the normalization (7.64) we obtain as a crude estimate

$$2\pi\hbar v_0 |f(E_0)|^2 \approx \frac{1}{\Delta E}$$

Hence, an order of magnitude estimate for the probability that transmission has occurred is

$$\frac{\Gamma}{\Delta E}\left\{1 - \exp\left[-\frac{\Gamma}{\hbar}\left(t + \frac{x_0}{v_0}\right)\right]\right\} \tag{7.72}$$

The total transmission probability for the incident wave packet (7.63), found by letting $t \to \infty$, is thus approximately equal to $\Gamma/\Delta E$. Equation (7.72) leads to the following simple interpretation: The wave packet reaches the well at time $-x_0/v_0$. A fraction $\Gamma/\Delta E$ of the packet is transmitted according to an exponential time law with a mean lifetime

$$\boxed{\tau = \frac{\hbar}{\Gamma} = \hbar\theta^2\left(\frac{\partial L}{\partial E}\right)_{E=E_0}} \tag{7.73}$$

The remaining portion of the wave packet is reflected promptly.

The study of resonance transmission affords us an example of the familiar *exponential decay law*, and the well with corresponding barriers can serve as a one-dimensional model of nuclear alpha decay. Decay processes will be encountered again in Chapters 13 and 19, but it is well to point out here that the exponential decay law can be derived only as an approximate, and not a rigorous, result of quantum mechanics. It holds only if the decay process is essentially independent of the manner in which the decaying state was formed and of the particular details of the incident wave packet.

Exercise 7.12. Show that condition (7.71) implies that the time it takes the incident wave packet to enter the well must be long compared with the classical period of motion and short compared with the lifetime of the decaying state.

Exercise 7.13. Resonances in the double well (Figure 7.7) may also be defined as quasi-bound states by requiring $A_1 = B_7 = 0$ (no incident wave) or $M_{11} = 0$ in analogy to truly bound states [see Eq. (6.51)]. Show that this condition defines poles in the S matrix as a function of the complex variable E. The real parts of these discrete complex values of E are the resonance energies, whereas the imaginary parts are the half-widths. Construct and interpret the asymptotic solution of a wave equation (time-dependent Schrödinger equation) which corresponds to one of these complex E values (*decaying states*).

Problems

1. Apply the WKB method to a particle that falls with acceleration g in a uniform gravitational field directed along the z axis and that is reflected from a perfectly elastic plane surface at $z = 0$. Compare with the rigorous solutions of this problem.
2. Apply the WKB approximation to the energy levels below the top of the barrier in a symmetric double well, and show that the energy eigenvalues are determined by a condition of the form

$$\tan\left(\int k\,dx + \alpha\right) = \pm 2\theta$$

where θ is the quantity defined in (7.48) for the barrier, α is a constant dependent on the boundary conditions, and the integral $\int k\,dx$ is to be extended between the classical turning points in one of the separate wells. Show that at low transmission the energy levels appear in close pairs with a level splitting approximately equal to $\hbar\omega/\pi\theta$ where ω is the classical frequency of oscillation in one of the single wells.

3. A particle of mass m moves in one dimension between two infinitely high potential walls at $x = a$ and $x = -a$. In this interval the potential energy is $V = -C|x|$, C being a positive constant. In the WKB approximation, obtain an equation determining the energy eigenvalues $E \leq 0$. Estimate the minimum value of C required for an energy level with $E \leq 0$ to exist.
4. Apply the WKB approximation to a particle of mass m moving with energy E in the field of an inverted oscillator potential, $V(x) = -m\omega^2x^2/2$. Determine the WKB wave functions for positive and negative values of the energy, E. Estimate the limits of the region in which the WKB wave functions are expected to be valid approximations to the exact Schrödinger wave function.

CHAPTER 8

Variational Methods and Simple Perturbation Theory

Variational methods are as central to quantum mechanics as to classical mechanics. They also serve as a springboard for numerical computation. When quantum mechanics is applied to realistic physical systems, we must usually employ approximation methods. The WKB method (Chapter 7) is limited to models that tend to be oversimplified. Simple perturbation theory, introduced in this chapter, greatly extends our access to interesting applications. A more systematic study of perturbation methods is the subject of later chapters (18 and 19). The Rayleigh-Ritz variational method provides a bridge from wave mechanics to matrix mechanics. The molecular approximation capitalizes on the mass disparity between electrons and nuclei. This chapter concludes with examples of applications to molecular structure and the band theory of solids.

1. ***The Calculus of Variations in Quantum Mechanics.*** Explicit analytic solutions of the Schrödinger equation can be derived only for a limited number of potential energy functions. The last three chapters dealt with three of these: the harmonic oscillator, the piecewise constant potential, and the linear potential. Another, the Coulomb potential, will be studied in Chapter 12. In the meantime, we will consider some very general variational methods, which make it possible to obtain approximate energy eigenvalues and which also shed light on the physical properties of the stationary states.

The time-independent Schrödinger equation can be regarded as the Euler-Lagrange equation for a variational problem that can be formulated as follows:

Subject to the constraint

$$\boxed{\int \psi^*(\mathbf{r})\psi(\mathbf{r})d^3r = 1} \tag{8.1}$$

find the functions $\psi(\mathbf{r})$ *and* $\psi^*(\mathbf{r})$ *that cause the variation* δI *of the expression*

$$\boxed{I = \int \left(\frac{\hbar^2}{2m}\nabla\psi^* \cdot \nabla\psi + V\psi^*\psi\right)d^3r} \tag{8.2}$$

to vanish.

More graphically, but somewhat less precisely: *Find the functions* $\psi(\mathbf{r})$ *and* $\psi^*(\mathbf{r})$ *which make the integral (8.2) stationary* in the mathematical sense of the term. If $\psi(\mathbf{r})$ and $\psi^*(\mathbf{r})$ are sufficiently smooth and vanish properly at the boundary, the expression (8.2) can be transformed by integration by parts into

$$I = \int \left(-\frac{\hbar^2}{2m}\psi^*\nabla^2\psi + V\psi^*\psi\right)d^3r \tag{8.3}$$

which is the formula for the expectation value of the energy, so that $I = \langle H \rangle$. For the purpose of establishing the variational principle, Eq. (8.2) is more appropriate. For many applications (8.2) and (8.3) are equivalent, but (8.2) has precedence whenever there are ambiguities.

For an understanding of the variational principle and its implications, it is sufficient to work with the one-dimensional case. According to the method of *Lagrangian multipliers*, the constrained problem posed here is equivalent to an unconstrained variational problem for the integral $I - E\int \psi^*\psi\, dx$:

$$\delta \int F(x)\, dx = \delta \int \left(\frac{\hbar^2}{2m} \frac{d\psi^*}{dx} \frac{d\psi}{dx} + V\psi^*\psi - E\psi^*\psi \right) dx = 0 \tag{8.4}$$

where E is an as yet undetermined constant, the Lagrangian multiplier. If $\psi(x)$ and $\psi^*(x)$ are varied independently of each other, as if they were two unrelated functions with their conjugate relationship temporarily ignored, two differential (Euler) equations are obtained:

$$\begin{aligned} \frac{\partial F}{\partial \psi^*} - \frac{d}{dx}\frac{\partial F}{\partial \psi^{*\prime}} &= -\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} + V\psi - E\psi = 0 \\ \frac{\partial F}{\partial \psi} - \frac{d}{dx}\frac{\partial F}{\partial \psi'} &= -\frac{\hbar^2}{2m}\frac{d^2\psi^*}{dx^2} + V\psi^* - E\psi^* = 0 \end{aligned} \tag{8.5}$$

where ψ' and $\psi^{*\prime}$ denote derivatives with respect to x. Both of these equations have the form of the time-independent Schrödinger equation. Since $\psi(x)$ and $\psi^*(x)$ satisfy the appropriate regular boundary conditions, E is assured to be a real number. Hence, all equations are consistent if $\psi^*(x)$ is chosen as the complex conjugate of $\psi(x)$. [The procedure of varying $\psi(x)$ and $\psi^*(x)$ independently is equivalent to expressing ψ in terms of $\mathrm{Re}\psi + i\,\mathrm{Im}\psi$ and varying $\mathrm{Re}\psi$ and $\mathrm{Im}\psi$ independently.]

The simultaneous use of $\psi(\mathbf{r})$ and $\psi^*(\mathbf{r})$ may seem like an unnecessary complication if V is real. It leads, however, to a variational expression with a simple physical significance, since (8.3) denotes the expectation value of the energy. As the energy, this quantity has a lower bound, and the energy of the *ground state* of the system can therefore be determined as the *absolute minimum* of $\langle H \rangle$. The variational procedure is easily generalized to a particle in three dimensions and to the Schrödinger equation for a particle in a magnetic field, which is not invariant under time reversal and has no simple reality properties.

The equivalence of the variational principle and the Schrödinger equation can also be demonstrated without constraining $\psi(\mathbf{r})$ and $\psi^*(\mathbf{r})$ as in (8.1) and using a Lagrangian multiplier, but instead by expressing the expectation value of the Hamiltonian as

$$\boxed{\langle H \rangle = \frac{\int \psi^* H \psi\, d^3r}{\int \psi^*\psi\, d^3r}} \tag{8.6}$$

If the variation of $\langle H \rangle$ is defined by

$$\delta\langle H \rangle = \frac{\int (\psi^* + \delta\psi^*) H (\psi + \delta\psi) d^3r}{\int (\psi^* + \delta\psi^*)(\psi + \delta\psi) d^3r} - \frac{\int \psi^* H \psi\, d^3r}{\int \psi^*\psi\, d^3r} \tag{8.7}$$

for trial functions $\psi + \delta\psi$ and $\psi^* + \delta\psi^*$, upon expanding this expression we obtain:

$$\left(\int \psi^*\psi \, d^3r\right)^2 \delta\langle H\rangle = \int \psi^*\psi \, d^3r\left[\int \delta\psi^* H\psi \, d^3r + \int \psi^* H\delta\psi \, d^3r\right] - \int \psi^* H\psi \, d^3r\left[\int \delta\psi^*\psi \, d^3r + \int \psi^*\delta\psi \, d^3r\right] + O[(\delta\psi)^2] \tag{8.8}$$

where $O[(\delta\psi)^2]$ symbolizes all terms of order higher than the first in the generally independent variations $\delta\psi$ and $\delta\psi^*$. Two conclusions can be drawn from (8.8):

1. If $\psi = \psi_k$ is one of the normalized eigenfunctions of H, such that

$$\int \psi_k^*\psi_k \, d^3r = 1 \quad \text{and} \quad \int \psi_k^* H\psi_k \, d^3r = E_k$$

then $\delta\langle H\rangle = 0$ to first order, and $\langle H\rangle$ is made stationary by the eigenfunction.

2. Conversely, if *all* first-order variations $\delta\langle H\rangle$ are stipulated to vanish, then ψ must be an eigenfunction of H. The proof of this proposition follows if we use the Hermitian property of H and choose as the variation of ψ:

$$\delta\psi = \varepsilon\left[\left(\int \psi^*\psi \, d^3r\right)H\psi - \left(\int \psi^* H\psi \, d^3r\right)\psi\right]$$

with a small real number ε. The condition $\delta\langle H\rangle = 0$ then gives, from the first-order terms in (8.8), the equation

$$H\psi = E\psi \tag{8.9}$$

where

$$E = \langle H\rangle \tag{8.10}$$

Condition (8.9) is the equation whose eigenfunctions and eigenvalues are ψ_k and E_k.

To summarize: If a trial function $\psi_k + \delta\psi$, differing from an eigenfunction ψ_k by $\delta\psi$, is used to calculate the expectation value of the Hamiltonian, $E_k + \delta\langle H\rangle$ is obtained, and for small $\delta\psi$ the change $\delta\langle H\rangle$ is of order $(\delta\psi)^2$.

In practice, then, we can obtain various overestimates of the ground state energy by calculating $\langle H\rangle$ for suitably chosen trial functions. Obviously, altogether erroneous estimates are likely to be obtained unless the trial function simulates the correct ground state wave function reasonably well. Usually, it is necessary to strike a compromise between the desire to improve the estimate of the ground state energy by choosing a "good" wave function and the requirement of ease of calculation.

The variational method is frequently applied to the Schrödinger equation by using for the calculation of $\langle H\rangle$ a trial wave function that contains one or more real variable parameters $\alpha, \beta, \gamma \ldots$. If the expectation value $\langle H\rangle$ is a differentiable function of these parameters, the extrema of $\langle H\rangle$ are found with the help of the equations

$$\frac{\partial\langle H\rangle}{\partial\alpha} = \frac{\partial\langle H\rangle}{\partial\beta} = \frac{\partial\langle H\rangle}{\partial\gamma} = \cdots = 0 \tag{8.11}$$

Evidently, the absolute minimum of $\langle H\rangle$ obtained by this method gives an upper bound for the lowest eigenvalue of H. This bound may even be a fair approximation to the lowest energy eigenvalue if the parametric trial function is flexible enough to

simulate the ground-state eigenfunction ψ_0. The other, relative, extrema of $\langle H \rangle$ may correspond to excited states of the system, provided that the trial function is sufficiently adaptable to represent the desired eigenstate of H to a reasonable approximation.

As an example, we consider the Schrödinger equation for the potential $V(x) = g|x|$. This problem was solved in Section 7.3 by exact methods and by applying the WKB approximation. Here we make use of the variational method by choosing a normalized Gaussian trial function $\psi_t(x)$ with adjustable width,

$$\psi_t(x) = \left(\frac{2\alpha}{\pi}\right)^{1/4} e^{-\alpha x^2} \tag{8.12}$$

The expectation value of H becomes

$$\langle H \rangle = \int_{-\infty}^{+\infty} \psi_t^*(x)\left\{-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + g|x|\right\}\psi_t(x)\,dx = \frac{\hbar^2\alpha}{2m} + \frac{g}{\sqrt{2\pi\alpha}} \tag{8.13}$$

From this expression we obtain the condition for a minimum:

$$\frac{\partial\langle H \rangle}{\partial\alpha} = \frac{\hbar^2}{2m} - \frac{g}{2\sqrt{2\pi}}\alpha^{-3/2} = 0$$

or

$$\alpha = \left(\frac{gm}{\sqrt{2\pi}\hbar^2}\right)^{2/3} \tag{8.14}$$

When this result is substituted in (8.13), the variational estimate

$$\langle H \rangle = \frac{1}{\pi^{1/3}}\left[\frac{1}{2^{4/3}} + \frac{1}{2^{1/3}}\right]\left(\frac{g^2\hbar^2}{m}\right)^{1/3} = 0.813\left(\frac{g^2\hbar^2}{m}\right)^{1/3} \tag{8.15}$$

is obtained. As expected, this estimate exceeds the exact ground state energy

$$E_0 = 0.809\left(\frac{g^2\hbar^2}{m}\right)^{1/3}.$$

but is remarkably close to this value. Figure 8.1 shows a comparison of the exact (Airy) ground state wave function and the optimal Gaussian trial function.

Exercise 8.1. For the Schrödinger equation with a potential $V(x) = g|x|$ (with $g > 0$), use an optimized *exponential* trial function to estimate the ground state energy. How does the result compare with the estimate from the Gaussian trial function?

Exercise 8.2. Calculate the variational estimate for the ground state energy of the potential $V(x) = g|x|$ (with $g > 0$), using the *triangular* trial function

$$\psi(x) = C(\alpha - |x|) \quad \text{for} \quad |x| \leq \alpha, \quad \text{and} \quad \psi(x) = 0 \quad \text{for } |x| > \alpha$$

Why is it safest to use (8.2) rather than (8.3) here? How good is the estimate?

Exercise 8.3. Invent other simple trial functions that can be used to estimate the ground state energy of the particle in the potential $V(x) = g|x|$ (with $g > 0$). By

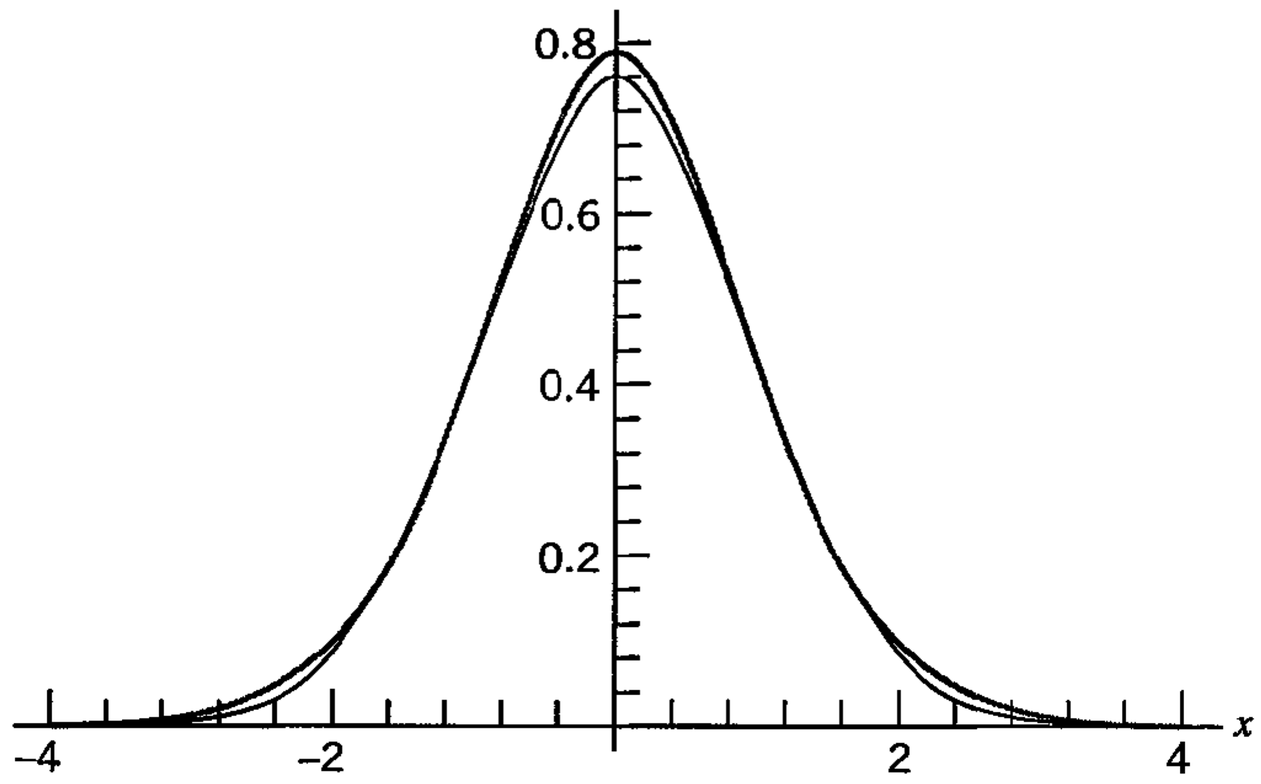


Figure 8.1. The thick line is the normalized ground state energy eigenfunction, $\psi_0(x) = 1.47\ Ai[\sqrt[3]{2}(x - 0.8086)]$, for the potential $V(x) = |x|$, and the thin line represents the optimal normalized Gaussian variational trial function, $\psi_t(x)$. The x coordinate and the energy are made dimensionless by scaling, as described in Section 7.3. The mean-square deviation between the two functions is only $\int_{-\infty}^{+\infty} |\psi_0(x) - \psi_t(x)|^2\, dx \approx 10^{-3}$!

using odd trial functions with only one node, at the origin, obtain numerical estimates for the first excited state and compare with the exact value of E_1 in Table 7.1. Can you proceed to the second excited state and calculate a variational energy estimate for this state?

2. ***The Rayleigh-Ritz Trial Function.*** One of the most prominent applications of the variational principle is the *Rayleigh-Ritz* method. This procedure is applicable to the discrete spectrum of an Hermitian operator and consists of using as a trial function ψ a linear combination of n suitably chosen, linearly independent, quadratically integrable functions χ_i:

$$\boxed{\psi(\mathbf{r}) = \sum_{i=1}^{n} c_i\, \chi_i(\mathbf{r})} \tag{8.16}$$

The functions χ_i are referred to as a set of *basis functions*. The coefficients c_i are to be determined from the variational principle. Their real and imaginary parts may be taken as the variational parameters, or, as in Section 8.1, the coefficients c_i and c_i^* may be varied independently. The $\psi(\mathbf{r})$ can be regarded as vectors in an n-dimensional space that is spanned by the basis $\chi_i(\mathbf{r})$. The coefficients c_i are the components of the vector.

Generally, the basis functions χ_i need not be orthogonal or normalized to unity. If they are not, we know from Section 4.1 how to proceed in order to replace them by an equivalent orthonormal set of n basis functions. This cumbersome orthogonalization procedure can be avoided by keeping the formalism sufficiently general to accommodate the use of nonorthonormal basis functions, much as nonorthogonal coordinate systems may be employed in geometry. The Rayleigh-Ritz method will be extended to arbitrary basis functions in Section 8.4. In this section, we assume that the basis functions are *orthonormal*.

The expectation value $\langle H \rangle$ is expressed in terms of the c_i and c_i^* as

$$\langle H \rangle = \frac{\int \psi^* H \psi \, d^3r}{\int \psi^* \psi \, d^3r} = \frac{\sum_{i,j=1}^{n} c_j^* \langle j|H|i\rangle c_i}{\sum_{i,j=1}^{n} c_j^* c_i} \tag{8.17}$$

where the *matrix elements* of H are defined by

$$\boxed{\langle j|H|i\rangle = \int \psi_j^* H \psi_i \, d^3r} \tag{8.18}$$

Since H is a Hermitian operator,

$$\begin{aligned} \langle i|H|j\rangle &= \int \psi_i^* H \psi_j \, d^3r = \left[\int (H\psi_j)^* \psi_i \, d^3r\right]^* \\ &= \left(\int \psi_j^* H \psi_i \, d^3r\right)^* = \langle j|H|i\rangle^* \end{aligned} \tag{8.19}$$

The variational conditions for making $\langle H \rangle$ stationary,

$$\frac{\partial \langle H \rangle}{\partial c_j^*} = 0 \quad \text{and} \quad \frac{\partial \langle H \rangle}{\partial c_i} = 0 \qquad \text{for all } i, j = 1, 2, \ldots n \tag{8.20}$$

produce the n linear homogeneous equations

$$\sum_{j=1}^{n} \langle i|H|j\rangle c_j = \langle H \rangle c_i \tag{8.21}$$

Exercise 8.4. Derive (8.21) from the variational principle, $\delta\langle H \rangle = 0$. Show that, because H is Hermitian, the second set of linear equations derived from (8.20) is redundant.

If we denote the real expectation value of H corresponding to the (initially unknown) optimal trial function by $\langle H \rangle = E$, the system of equations (8.21) can be written as

$$\boxed{\sum_{j=1}^{n} \langle i|H|j\rangle c_j = E \, c_i} \tag{8.22}$$

or

$$\sum_{j=1}^{n} (\langle i|H|j\rangle - E\delta_{ij}) c_j = 0 \tag{8.23}$$

This system of equations for the coefficients c_i has nontrivial solutions only if the *characteristic value* (or *eigenvalue*) E is one of the n roots of the determinantal (*characteristic* or *secular*) equation:

$$\boxed{D_n(E) \equiv \det(\langle i|H|j\rangle - E\delta_{ij}) = 0} \tag{8.24}$$

If we take the complex conjugate of Eq. (8.22) and use the property (8.19), we obtain

$$\sum_{j=1}^{n} c_j^* \langle j|H|k\rangle = E c_k^* \tag{8.25}$$

If Eqs. (8.22) and (8.25) are written for two different characteristic values, E and $\bar{E}$, corresponding to trial functions ψ and $\bar{\psi}$, it is easy to see that

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \bar{c}_i^* \langle i|H|j\rangle c_j = E \sum_{i=1}^{n} \bar{c}_i^* c_i$$

and

$$\sum_{j=1}^{n}\sum_{k=1}^{n} \bar{c}_j^* \langle j|H|k\rangle c_k = \bar{E} \sum_{k=1}^{n} \bar{c}_k^* c_k$$

from which by subtraction it follows that

$$(E - \bar{E}) \sum_{i=1}^{n} \bar{c}_i^* c_i = 0$$

Hence, the optimal Rayleigh-Ritz trial functions corresponding to *different* characteristic values are orthogonal:

$$\boxed{\int \bar{\psi}^*(\mathbf{r})\, \psi(\mathbf{r})\, d^3r = \sum_{i=1}^{n} \bar{c}_i^* c_i = 0} \tag{8.26}$$

Since the system of equations (8.22) is linear and homogeneous, any linear combination of two or more optimal trial functions belonging to the *same* characteristic value E is also an optimal trial function. Hence, as was shown in Chapter 4, if the optimal trial functions are not initially orthogonal, they may be "orthogonalized" and replaced by an equivalent set of orthogonal functions. It is important to note that all of these results pertain to the *approximate* solutions of the Schrödinger equation, whereas we already know them to be true for the *exact* solutions.

Exercise 8.5. Show that if the n optimal orthonormal trial functions, obtained by the Rayleigh-Ritz method, are used as basis functions, the matrix elements of H are diagonal:

$$\langle i|H|j\rangle = E_i' \delta_{ij} \tag{8.27}$$

In order to help us appreciate the significance and utility of the n characteristic values E produced by the Rayleigh-Ritz method, let us assume that the roots of the determinantal equation (8.24) have been ordered so that

$$E_1' \leq E_2' \leq \cdots \leq E_{n-1}' \leq E_n'$$

The corresponding n optimal trial functions may be assumed to constitute an orthonormal set. If this orthonormal set is used as a basis, (8.27) shows that the determinantal equation takes the simple form

$$D_n(E) \equiv \begin{vmatrix} E_1'-E & 0 & \cdots & 0 \\ 0 & E_2'-E & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdots & E_n'-E \end{vmatrix} = 0 \tag{8.28}$$

Suppose that an $(n+1)$-st basis function is added, which may be chosen to be orthogonal to the previous n basis functions. In this new basis, the determinantal equation is also simple:

$$D_{n+1}(E) = \begin{vmatrix} E_1'-E & 0 & & \langle 1|H|n+1\rangle \\ 0 & E_2'-E & & \langle 2|H|n+1\rangle \\ \cdot & \cdot & \cdots & \cdot \\ \langle n+1|H|1\rangle & \langle n+1|H|2\rangle & \cdots & \langle n+1|H|n+1\rangle - E \end{vmatrix} = 0 \quad (8.29)$$

From this equation a new set of $n+1$ roots or characteristic values can be calculated. The values of the determinant $D_{n+1}(E)$ for $E = E_r'$ and $E = E_{r+1}'$ can be shown to have opposite signs, provided that $E_{r+1}' \neq E_r'$. Hence, between E_r' and E_{r+1}' there must be at least one root, a characteristic value. From the behavior of the determinant $D_{n+1}(E)$ for $E \to \pm\infty$, it can be seen that there must also be one characteristic value that is $\leq E_1'$ and one that is $\geq E_n'$. Hence, since there are exactly $n+1$ roots, each interval between E_r' and E_{r+1}' must contain one new characteristic value, and the two remaining characteristic values lie below and above the old spectrum. If $E_r' = E_{r+1}'$, a new root of $D_{n+1}(E)$ coincides with these two previous characteristic values, but generally it need not be a repeated zero of $D_{n+1}(E)$.

Exercise 8.6. If the roots of $D_3(E) = 0$ are distinct and ordered as $E_1' < E_2' < E_3'$, show that $D_4(E_1')$ and $D_4(E_3')$ are negative and $D_4(E_2')$ is positive. What can you infer about the roots of $D_4(E) = 0$? Try to discern a pattern.

Since the Hamiltonian is an operator with a lower bound for its expectation values, it follows that E_1' lies above the ground state energy, E_2' lies above the first excited level, E_3' is above the second excited level, and so on. As a new basis function is added, all the new variational estimates E_r'' will move down. Figure 8.2 illustrates the situation.

If in the limit as $n \to \infty$ the basis set approaches a *complete set* of linearly independent functions for the domain of H, in the sense of convergence "in the mean," the optimal trial functions approach the true eigenfunctions of the Hamiltonian operator. The determinantal characteristic equation becomes infinitely dimensional, and as n is increased its roots approach the eigenvalues of H from above.

Thus, in practice, the Rayleigh-Ritz method provides a way of estimating upper bounds for the lowest n eigenvalues of the energy operator H, and the bounds are steadily improved by choosing ever larger values of n. There is no assurance that the convergence will be rapid; in fact, generally it will be slow unless the trial function can be made to resemble the actual eigenfunctions closely. An application of the Rayleigh-Ritz is afforded by perturbation theory for degenerate or near-degenerate states in the next section.

The equations derived in this section for the approximate determination of energy eigenvalues and the corresponding eigenfunctions are all familiar from linear algebra and can be written in matrix form. We will return to the matrix representation in Chapter 9, but the present discussion already shows how to link Schrödinger's *wave mechanics* with Heisenberg's *matrix mechanics*.

3. ***Perturbation Theory of the Schrödinger Equation.*** In every physical theory two trends are evident. On the one hand, we strive to formulate exact laws and equations that govern the phenomena; on the other hand, we are confronted with the

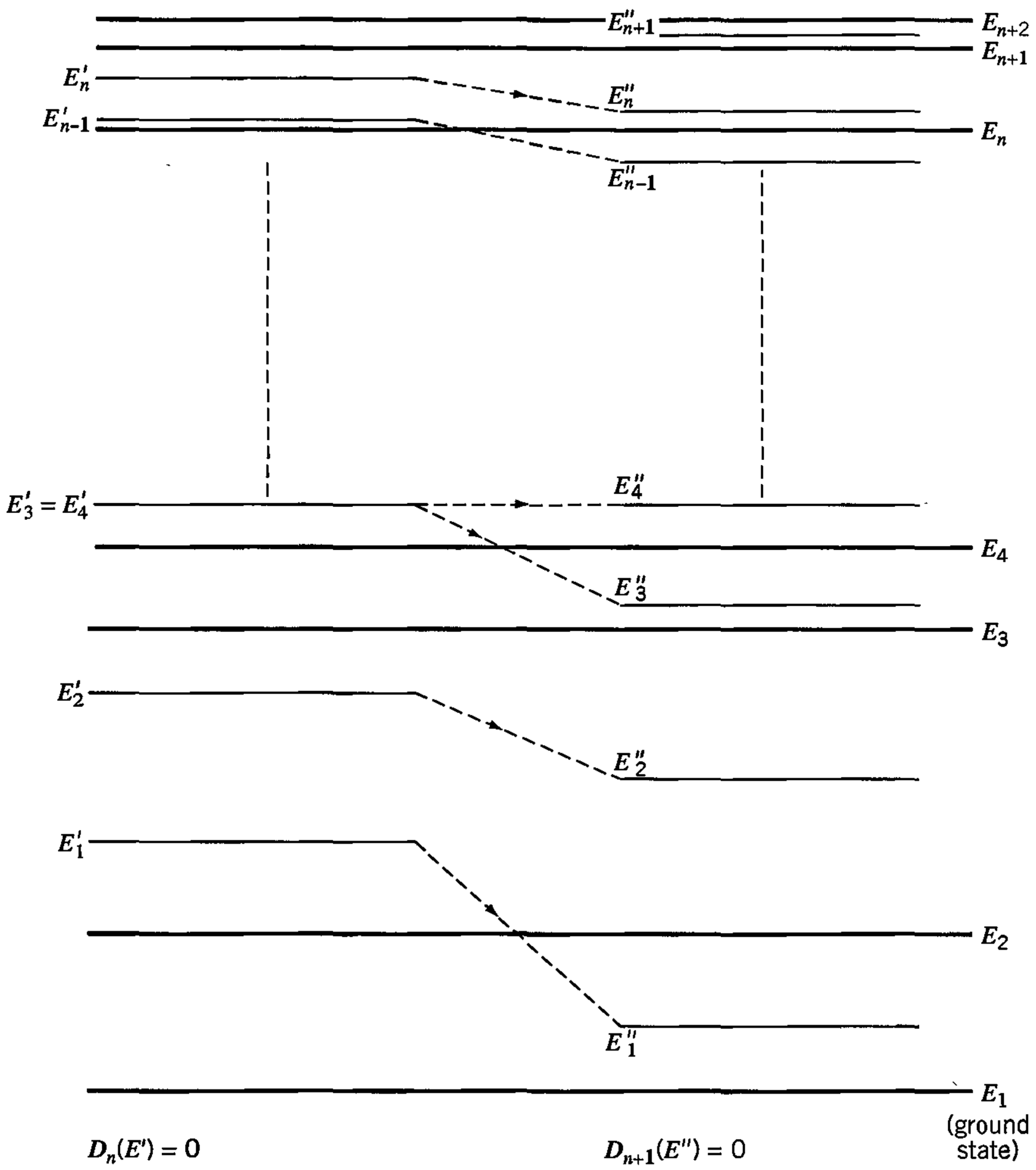


Figure 8.2. Diagram showing schematically the location of the roots of the characteristic determinantal equations $D_n(E') = 0$ and $D_{n+1}(E'') = 0$ in relation to the exact eigenvalues E_r of the $(r\text{-}1)$-st excited energy level of the Hamiltonian H, which the approximate eigenvalues approach from above. When the $(n+1)$-st basis function is added, the degeneracy $E'_3 = E'_4$ is removed and a new approximate eigenvalue, E''_{n+1}, appears.

need to obtain more or less approximate solutions to the equations, because rigorous solutions can usually be found only for oversimplified models of the physical situation. These are nevertheless useful, because they often serve as a starting point for approximate solutions to the complicated equations of the actual system.

In quantum mechanics, the perturbation theories are examples of this approach. Given a complex physical system, we choose, if possible, a similar but simpler comparison system whose quantum equations can be solved exactly. The complicated actual system may then often be described to good approximation in terms of the cognate system.

Examples of the uses of perturbation theory are legion. For instance, in atomic physics the problem of the motion of an electron in a Coulomb field can be solved rigorously (Chapter 12), and we may regard the motion of an electron in a real many-

electron atom as approximated by this simpler motion, perturbed by the interaction with the other electrons. The change of the energy levels of an atom in an applied electric field can be calculated by treating the field as an added perturbation, and the influence of an anharmonic term in the potential energy of an oscillator on the energy spectrum can be assessed by perturbation theory. In a large class of practical problems in the quantum domain, perturbation theory provides at least a first qualitative orientation, even where its quantitative results may be inaccurate. This section is our first introduction to perturbation theory for stationary states. A more systematic treatment, including the calculation of higher-order approximations, will be given in Chapter 18.

Suppose that a Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V_1\psi = E\psi \tag{8.30}$$

with an unmanageable potential energy V_1 is given, but that we know the normalized solutions of a similar Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\psi^{(0)} + V_0\psi^{(0)} = E^{(0)}\psi^{(0)} \tag{8.31}$$

with V_1 differing from V_0 by a small amount $V = V_1 - V_0$. If V is small (in a sense yet to be made precise), we intuitively expect that the eigenvalues $E^{(0)}$ of (8.31) will be reasonably close to the energy eigenvalues E of (8.30), and that the corresponding eigenfunctions $\psi^{(0)}$ will be similar to ψ.

If the *unperturbed* eigenfunction $\psi^{(0)}$ is used as a trial function and if $E^{(0)}$ is *nondegenerate*, the variational expression

$$\begin{aligned}\langle H\rangle &= \int \psi^{(0)*}\left[-\frac{\hbar^2}{2m}\nabla^2 + V_1\right]\psi^{(0)}\,d^3r = \int \psi^{(0)*}(H_0 + V)\psi^{(0)}\,d^3r \\ &= E^{(0)} + \int \psi^{(0)*}V\psi^{(0)}\,d^3r\end{aligned} \tag{8.32}$$

represents a sensible, and often quite accurate, estimate of the energy eigenvalue of the Schrödinger equation (8.30). The difference potential V is called a *perturbation* of the unperturbed potential energy V_0, and the change in the energy eigenvalue is

$$\boxed{\Delta E = E - E^{(0)} \approx \int \psi^{(0)*}V\psi^{(0)}\,d^3r} \tag{8.33}$$

which is one of the most useful formulas in quantum mechanics.

If $E^{(0)}$ in (8.31) is degenerate and corresponds to n linearly independent unperturbed eigenfunctions, $\psi_1^{(0)}, \psi_2^{(0)}, \ldots, \psi_n^{(0)}$, the Rayleigh-Ritz method of the last section can be used to calculate a variational estimate of the perturbed energy. Frequently, this method is also useful if the n unperturbed energies, rather than being all equal as in the case of degeneracy, are merely so close to each other that the perturbation is effective in "mixing" them.

We assume that the n unperturbed degenerate or near-degenerate eigenfunctions are orthonormal. The matrix elements (8.18) of the Hamiltonian $H = H_0 + V$ now take the form

$$\langle i|H|j\rangle = E_i^{(0)}\,\delta_{ij} + \langle i|V|j\rangle \tag{8.34}$$

The determinantal equation for the estimated energy E is

$$\det[(E_i^{(0)} - E)\delta_{ij} + \langle i|V|j\rangle] = 0 \tag{8.35}$$

The n roots of this (secular) equation are approximate perturbed energy values.

As an example, we suppose that two particular eigenvalues $E_1^{(0)}$ and $E_2^{(0)}$ of H_0 are close to each other, or degenerate (in which case $E_1^{(0)} = E_2^{(0)}$), and belong to two orthonormal eigenfunctions $\psi_1^{(0)}$ and $\psi_2^{(0)}$, which will serve as basis functions. Equations (8.35) turns for this case into

$$\begin{vmatrix} \langle 1|V|1\rangle + E_1^{(0)} - E & \langle 1|V|2\rangle \\ \langle 2|V|1\rangle & \langle 2|V|2\rangle + E_2^{(0)} - E \end{vmatrix} = 0 \tag{8.36}$$

with roots

$$\boxed{\begin{aligned} E_{1,2} = \frac{1}{2}\,[&E_1^{(0)} + E_2^{(0)} + \langle 1|V|1\rangle + \langle 2|V|2\rangle \\ &\pm \sqrt{(E_2^{(0)} + \langle 2|V|2\rangle - E_1^{(0)} - \langle 1|V|1\rangle)^2 + 4|\langle 1|V|2\rangle|^2}] \end{aligned}} \tag{8.37}$$

We consider two limiting cases in more detail.

(a) Here we assume that the unperturbed energies $E_1^{(0)}$ and $E_2^{(0)}$ are sufficiently separated so that

$$|E_2^{(0)} + \langle 2|V|2\rangle - E_1^{(0)} - \langle 1|V|1\rangle| \gg |\langle 1|V|2\rangle| \tag{8.38}$$

The magnitude of the "mixing" matrix element $\langle 1|V|2\rangle$ is small compared with the difference between the first-order perturbed energies. In this case, we can expand (8.37) and obtain

$$E_{1,2} = \begin{cases} E_1^{(0)} + \langle 1|V|1\rangle - \dfrac{|\langle 1|V|2\rangle|^2}{E_2^{(0)} + \langle 2|V|2\rangle - E_1^{(0)} - \langle 1|V|1\rangle} \\ E_2^{(0)} + \langle 2|V|2\rangle + \dfrac{|\langle 1|V|2\rangle|^2}{E_2^{(0)} + \langle 2|V|2\rangle - E_1^{(0)} - \langle 1|V|1\rangle} \end{cases} \tag{8.39}$$

This result shows that for the last term to be negligible and the first-order estimate (8.33) for the perturbed energy to be reliable, the perturbation V must be weak *and* the other unperturbed energy levels must be far away on the energy scale.

(b) In the opposite limit of exact degeneracy of the unperturbed levels, $E_1^{(0)} = E_2^{(0)} \equiv E^{(0)}$, and (8.37) reduces to

$$\boxed{E = E^{(0)} + \frac{1}{2}\,[\langle 1|V|1\rangle + \langle 2|V|2\rangle \pm \sqrt{(\langle 1|V|1\rangle - \langle 2|V|2\rangle)^2 + 4|\langle 1|V|2\rangle|^2}]} \tag{8.40}$$

Corresponding to these energies, the simultaneous linear equations (8.22) reduce to

$$\begin{pmatrix} \langle 1|V|1\rangle & \langle 1|V|2\rangle \\ \langle 2|V|1\rangle & \langle 2|V|2\rangle \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = (E - E^{(0)}) \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \tag{8.41}$$

The amplitudes c_1 and c_2 define the optimal trial functions

$$\psi = c_1\psi_1^{(0)} + c_2\psi_2^{(0)} \tag{8.42}$$

Unless $\langle 1|V|1\rangle = \langle 2|V|2\rangle$ *and* $\langle 1|V|2\rangle = \langle 2|V|1\rangle = 0$, the perturbation splits the degenerate levels and the two optimal trial functions are orthogonal. Two states

$\psi_1^{(0)}$ and $\psi_2^{(0)}$, which are connected by a nonvanishing perturbation matrix element $\langle 1|V|2\rangle$, are sometimes said to *interact* with each other.

Exercise 8.7. If the perturbation affecting two degenerate unperturbed states is such that $\langle 1|V|1\rangle = \langle 2|V|2\rangle$, show that $|c_1| = |c_2|$ and that the relative phase of the amplitudes c_1 and c_2 depends on the phase of $\langle 1|V|2\rangle$.

Exercise 8.8. Assume that the unperturbed Hamiltonian has an n-fold degenerate energy eigenvalue $E^{(0)} = 0$, that all diagonal matrix elements $\langle k|V|k\rangle = 0$ $(k = 1, 2, \ldots, n)$, and that all off-diagonal matrix elements $\langle k|V|\ell\rangle$ $(k \neq \ell)$ are equal to a real negative value $-\nu$ $(\nu > 0)$. Show that the degeneracy is partially removed and that the perturbed (ground) state has energy $E = -(n-1)\nu$, while all the others occur at $E = \nu$. Also show that the amplitudes, which define the ground state trial function, all have equal magnitudes and phases.

4. ***The Rayleigh-Ritz Method with Nonorthogonal Basis Functions.*** In Section 8.2, the basis for the Rayleigh-Ritz trial functions was assumed to be orthonormal. This is not necessary and not always desirable. Instead, we may choose basis functions χ_i that are neither orthogonal nor necessarily normalized and again consider trial functions of the form

$$\boxed{\psi(\mathbf{r}) = \sum_{i=1}^{n} c^i \chi_i(\mathbf{r})} \tag{8.43}$$

The index $i = 1, 2, \ldots, n$ on the undetermined coefficients c^i has been elevated to a superscript for purposes of this section only to emphasize the geometrical significance of vector representations like (8.43).

It is convenient to introduce a second related basis $\chi^j(\mathbf{r})$ in this n-dimensional vector space such that

$$\boxed{\chi^j(\mathbf{r}) = \sum_{i=1}^{n} \chi_i(\mathbf{r}) g^{ij}} \tag{8.44}$$

and

$$\boxed{\int \chi^{j*}(\mathbf{r})\chi_i(\mathbf{r}) d^3r = \delta_i^j} \tag{8.45}$$

where δ_i^j is the Kronecker delta: equal to one if $i = j$ and zero otherwise.

From (8.44) and (8.45) it follows that

$$\boxed{g^{ij} = \int \chi^{i*}\chi^j \, d^3r = g^{ji*}} \tag{8.46}$$

The g^{ij} can be taken to be the elements of an $n \times n$ matrix (with the first superscript labeling the rows and the second one the columns). A matrix for which (8.46) holds is said to be *Hermitian*. Similarly, if we define a second Hermitian matrix,

$$\boxed{g_{ij} = \int \chi_i^* \chi_j \, d^3r = g_{ji}^*} \tag{8.47}$$

we obtain, by substitution of (8.44) into (8.45), the connection

$$\boxed{\sum_{k=1}^{n} g^{jk} g_{ki} = \delta_i^j} \tag{8.48}$$

If it is assumed that the determinant of the matrix g_{ki} does not vanish, Eq. (8.48) can be used to calculate the coefficients g^{jk}. The two Hermitian matrices are inverses of each other.

The expressions (8.46) and (8.47) are said to be *overlap integrals*, because for $i \neq j$ their values tell us how nonorthogonal the basis functions are. Equation (8.44) can now also be inverted:

$$\chi_i = \sum_{i=1}^{n} \chi^j g_{ji} \tag{8.49}$$

and an alternative expansion for (8.43) can be derived:

$$\psi(\mathbf{r}) = \sum_{i=1}^{n} \sum_{j=1}^{n} c^i \chi^j g_{ji} = \sum_{j=1}^{n} c_j \chi^j(\mathbf{r}) \tag{8.50}$$

provided we identify

$$\boxed{c_j = \sum_{i=1}^{n} g_{ji} c^i} \tag{8.51}$$

On account of the condition (8.45), the new basis χ^j is called *reciprocal* to the basis χ_i.

If the basis χ_i is orthonormal, then the reciprocal basis coincides with it, and we have $\chi^i = \chi_i$ and $g_{ij} = g^{ij} = \delta_{ij}$. Such a basis is *self-reciprocal.*

The formalism developed in this section is isomorphic to the geometry of an n-dimensional vector space. The use of the g's to raise and lower indices by summation over repeated adjacent indices—one upper and one lower—and the matching of indices on two sides of an equation are the notational means by which the theory is expressed concisely and conveniently. All the relations and rules we have established have the same form independent of the choice of equivalent basis functions, provided only that these span the same space of trial functions (8.43). For more details, we refer to the standard mathematical literature.[1]

Exercise 8.9. On the interval $x = -\infty$ to $+\infty$ the nonorthogonal basis functions $\chi_k(x) = x^k e^{-|x|}$ are introduced with $k = 1, 2, \ldots, n$. For $n = 3$ and 4, construct the reciprocal bases.

The calculation of the variational expectation value $\langle H \rangle$ requires evaluation of the matrix elements

$$H_{ij} = \int \chi_i^* H \chi_j \, d^3r \tag{8.52}$$

[1]Coxeter (1969), Chapter 18.

or the related matrix elements

$$H^{ij} = \int \chi^{i*} H \chi^j \, d^3r, \qquad H^i{}_j = \int \chi^{i*} H \chi_j d^3r, \qquad H_i{}^j = \int \chi_i^* H \chi^j \, d^3r \quad (8.53)$$

Exercise 8.10. Show that any one of the matrix elements defined in (8.52) and (8.53) can be expressed as a linear combination of any of the others, e.g.,

$$H^{ij} = \sum_{\ell=1}^{n} g^{i\ell} H_\ell{}^j = \sum_{\ell=1}^{n} \sum_{k=1}^{n} g^{i\ell} H_{\ell k} g^{kj}$$

Thus, the g's serve to raise and lower indices.

Exercise 8.11. Show that the overlap integrals are the matrix elements of the identity operator in the nonorthogonal basis.

We now implement the variational principle for the functional $\langle H \rangle$ by varying the n components c_i and c^{i*} independently. As in Section 8.2, where the basis functions were orthonormal, we obtain

$$\sum_{j=1}^{n} H^i{}_j c^j = \langle H \rangle c^i \quad (8.54)$$

or their equivalent partners, obtained by raising or lowering indices, such as:

$$\sum_{j=1}^{n} H_{kj} c^j = \langle H \rangle \sum_{i=1}^{n} g_{ki} c^i \quad (8.55)$$

If we again use the notation $\langle H \rangle = E$, these systems of equations can also be written as

$$\boxed{\sum_{j=1}^{n} [H^i{}_j - E\delta^i_j] c^j = 0} \quad (8.56)$$

and

$$\boxed{\sum_{j=1}^{n} [H_{ij} - E g_{ij}] c^j = 0} \quad (8.57)$$

It is a matter of taste which of these sets of homogeneous linear equations one prefers to solve. If the basis is orthonormal and thus self-reciprocal, these equations all reduce to a single system. The system (8.56) is similar to the simultaneous equations (8.23) in Section 8.2. However the matrix $H^i{}_j$ is generally not a Hermitian matrix because

$$H^i{}_j = \int \chi^{i*} H \chi_j \, d^3r = \left[\int \chi_j^* (H \chi^i) d^3r \right]^* = H_j{}^{i*} \quad (8.58)$$

and this generally does *not* equal $H^j{}_i{}^*$ (except if the basis is orthogonal).

Exercise 8.12. Derive the n simultaneous equations (8.57) for the variational parameters c^i from the condition $\delta\langle H \rangle = 0$.

The system (8.57) contains Hermitian matrices, but the overlap integral matrix g appears with the eigenvalue. Both (8.56) and (8.57) give the same characteristic

values and the same optimal trial functions. The characteristic values are the n roots of the determinantal equations

$$\det(H^i{}_j - E\delta^i_j) = 0 \tag{8.59}$$

or

$$\det(H_{ij} - Eg_{ij}) = 0 \tag{8.60}$$

Exercise 8.13. From (8.55) deduce another convenient form of the variational equation:

$$\sum_{j=1}^{n} H^j_k c_j = Ec_k \tag{8.61}$$

and prove that the optimal trial functions belonging to two different roots of the characteristic equation are orthogonal.

Exercise 8.14. Show explicitly that the determinantal equations (8.59) and (8.60) have the same roots.

This brief introduction to nonorthogonal basis functions draws attention to a variational technique that has wide applicability. Nonorthogonal basis functions are particularly useful and popular in molecular physics and quantum chemistry. An elementary example in the next section will serve as an illustration.

5. ***The Double Oscillator.*** As a further illustration of approximation methods, we now supplement the discussion of the simple harmonic oscillator (Chapter 5) by study of a more complicated potential, pieced together from two harmonic oscillators. The example comes from molecular physics. There we frequently encounter motion in the neighborhood of a stable equilibrium configuration, approximated by a harmonic potential. To be sure, one-dimensional models are of limited utility. Even diatomic molecules rotate in space, besides vibrating along a straight line. Nevertheless, important qualitative features can be exhibited with a linear model, and some quantitative estimates can also be obtained.

Let us consider, as a model, two masses m_1 and m_2, constrained to move in a straight line and connected with each other by an elastic spring whose length at equilibrium is a. If x_1 and x_2 are the coordinates of the two mass points, and p_1 and p_2 their momenta, it is shown in classical mechanics that the nonrelativistic two-body problem can be separated into the motion of the center of mass and an equivalent one-body motion, executed by a fictitious particle of mass $m = m_1m_2/(m_1 + m_2)$ with a coordinate $x = x_1 - x_2$ about a fixed center under the action of the elastic force. The correspondence principle suggests that these general dynamical features of any two-particle system, which is subject only to an interaction potential between the two particles depending on their relative coordinates, survive the transition to quantum mechanics. Only the relative motion of the *reduced mass* m will be considered in this section. A full discussion of the separation of the relative motion from the motion of the center of mass will be given in Section 15.4.

The potential representing the relative motion of the reduced mass m is

$$V(x) = \frac{1}{2} m\omega^2(|x| - a)^2 \tag{8.62}$$

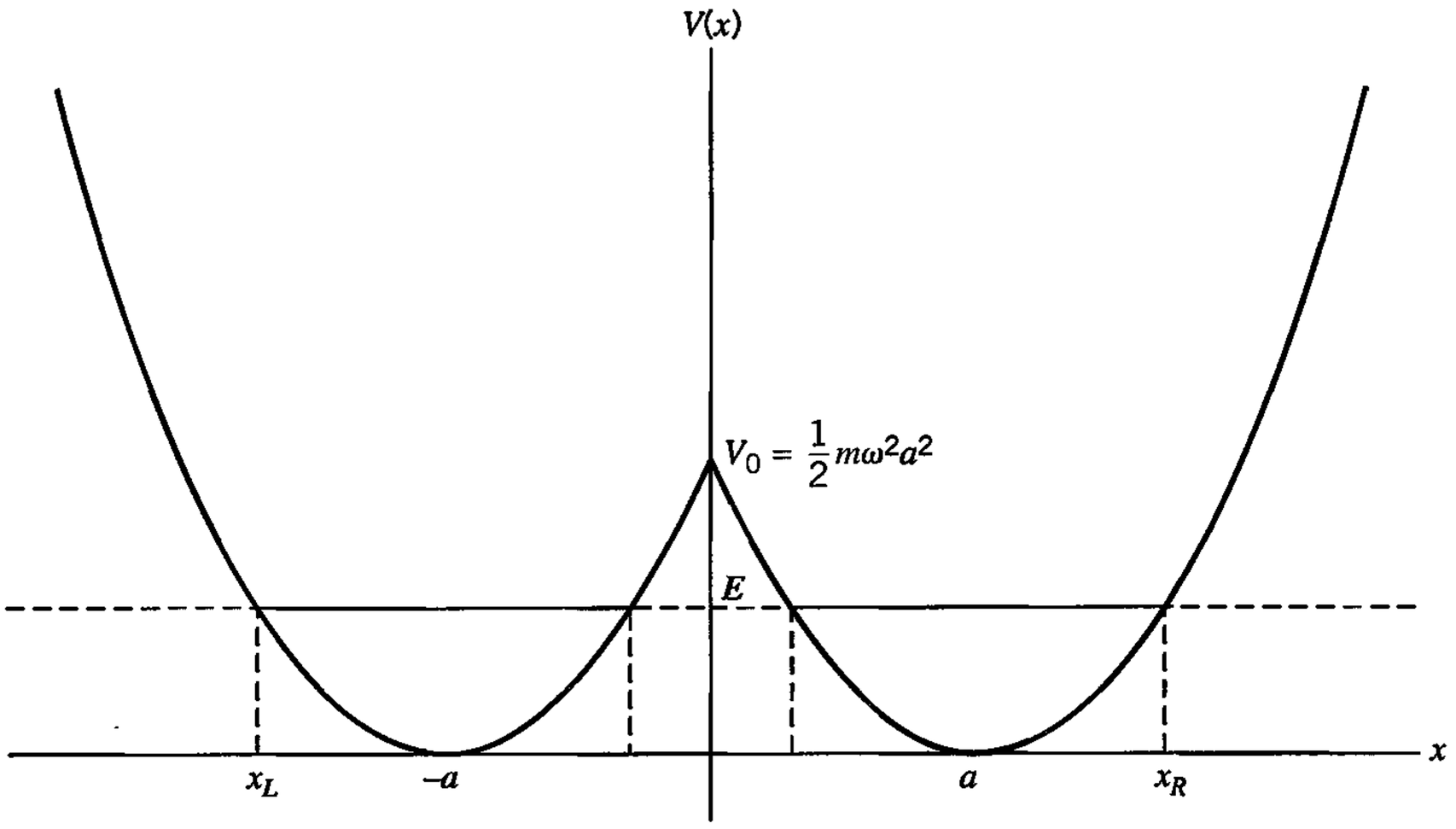


Figure 8.3. The double oscillator potential $V(x) = m\omega^2(|x|-a)^2/2$. The outer classical turning points are at x_L and x_R. If $E < V_0$, there are also two inner turning points, but the particle can tunnel through the barrier.

If $a = 0$, (8.62) reduces to the potential for the simple linear harmonic oscillator. If $a \neq 0$, it is almost the equation of a harmonic oscillator whose equilibrium position is shifted by an amount a, but it is not quite that. For it is important to note that the *absolute value* of x appears in the potential energy (8.62), because Hooke's law is assumed to hold for all values of the interparticle coordinate x. As shown in Figure 8.3, there are two equilibrium positions ($x = \pm a$), and we have two parabolic potentials, one when particle 1 is to the right of particle 2 ($x > 0$), and the other when the particles are in reverse order ($x < 0$). The two parabolas are joined at $x = 0$ with the common value $V(0) = V_0 = m\omega^2a^2/2$. Classically, if $E < V_0$, we can assume that only one of these potential wells is present, for no penetration of the barrier is possible. In quantum mechanics the barrier can be penetrated. Even if $E < V_0$, the wave functions may have a finite value at $x = 0$, which measures the probability that particles 1 and 2 are found in the same place.

The wave equation corresponding to the equivalent one-body problem is

$$i\hbar \frac{\partial \psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x, t)}{\partial x^2} + \frac{1}{2} m\omega^2(|x| - a)^2 \psi(x, t) \tag{8.63}$$

and the Schrödinger equation is

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x)}{\partial x^2} + \frac{1}{2} m\omega^2(|x| - a)^2 \psi(x) = E\psi(x) \tag{8.64}$$

For $|x| \gg a$, (8.64) approaches the Schrödinger equation for the simple harmonic oscillator; hence, the physically acceptable eigenfunctions must be required to vanish as $|x| \to \infty$.

Before attempting to solve (8.64), we note that as the parameter a is varied from 0 to $+\infty$, the potential changes from the limit (I) of a single harmonic oscillator well (of frequency ω) to the other limit (II) of two separate oscillator wells (also of

frequency ω), divided by an infinitely high and broad potential barrier. In case I, we have nondegenerate energy eigenvalues

$$E = \hbar\omega\left(n + \frac{1}{2}\right) \qquad (n = 0, 1, 2, \ldots) \tag{8.65}$$

In case II, the energy values are the same as those given by (8.65), but each is *doubly degenerate*, since the system may occupy an eigenstate at either the harmonic oscillator well on the left or the one on the right.

The potential energy for the double well, $V(x)$, is an even function and invariant under x reflection. The eigenstates have even or odd parity. The probability distribution for every energy eigenstate is symmetric about the origin, and in these states there is equal probability of finding the system in either of the two potential wells. If the state is to favor one side of the potential over the other, we must superpose even (symmetric) and odd (antisymmetric) stationary states. The superposition is generally not stationary.

As the limit of case II ($a \to \infty$) is approached, however, the two degenerate ground state wave functions are concentrated in the separate wells and do not have definite parity. Thus, the reflection symmetry of the double well is said to be hidden, or broken spontaneously by the ground state energy eigenfunctions, without any external influences. Case II serves to illustrate the concept of *spontaneous symmetry breaking* which arises in many physical systems, particularly in quantum field theory and many-body physics.

As a is varied continuously, energies and eigenfunctions change from case I to case II. It is customary to call this kind of continuous variation of an external parameter an *adiabatic* change of the system, because these changes describe the response of the system in time to changes of the external parameter performed infinitesimally slowly. As the potential is being distorted continuously, certain features of the eigenfunctions remain unaltered. Such stable features are known as *adiabatic invariants*.

An example of an adiabatic invariant is provided by the number of zeros, or nodes, of the eigenfunctions. If an eigenfunction has n nodes, as the eigenfunction of potential I belonging to the eigenvalue E_n does, this number cannot change in the course of the transition to II. We prove this assertion in two steps:

(a) No two adjacent nodes disappear by coalescing as a changes, nor can new nodes ever be created by the inverse process. If two nodes did coalesce at $x = x_0$, the extremum of ψ between them would also have to coincide with the nodes. Hence, both ψ and its first derivative ψ' would vanish at this point. By the differential equation (8.64), this would imply that all higher derivatives also vanish at this point. But a function all of whose derivatives vanish at a point can only be $\psi \equiv 0$. The presence of isolated singularities in V (as at $x = 0$ for the double oscillator) does not affect this conclusion, since ψ and ψ' must be continuous.

(b) No node can wander off to or in from infinity as a changes. To show this, we only need to prove that all nodes are confined to the region between the two extreme classical turning points of the motion, i.e., the region between x_L and x_R in Figure 8.3. Classically, the coordinate of a particle with energy E is restricted by $x_L \leq x \leq x_R$, where x_L and x_R are the smallest and largest roots, respectively, of $V(x) = E$. From Schrödinger's equation we infer that

$$\frac{\psi''}{\psi} = \frac{2m}{\hbar^2}\,[V(x) - E] \tag{8.66}$$

Then the expression on the right is positive definite in the classically inaccessible region. If there were a node in the region $x \leq x_L$, it would have to be a point of inflection for the function $\psi(x)$. Conversely, a point of inflection would also have to be a node. The existence of such a point is incompatible with the asymptotic condition: $\psi(x) \to 0 \quad \text{as} \quad x \to -\infty$. The same reasoning holds for the region $x \geq x_R$. It follows that *outside the classical region* there can be no node and no extremum (see Figure 6.3).

Being an adiabatic invariant, the number of nodes n characterizes the eigenfunctions of the double oscillator for any finite value of a. Figure 8.4 shows this for the two lowest energy eigenvalues ($n = 0, 1$).

The two eigenfunctions ψ_0 and ψ_1 correspond to $E = \hbar\omega/2$ and $3\hbar\omega/2$, if $a = 0$ (case I). For $a \to \infty$ (case II), they become degenerate with the common energy $\hbar\omega/2$. When a is very large, the linear combinations $\psi_0 + \psi_1$ and $\psi_0 - \psi_1$ are *approximate* eigenfunctions corresponding to the wave function being concentrated at $x = +a$ and at $x = -a$, respectively; the reflection symmetry is broken.

If a is finite but still large compared with the characteristic amplitude of the two separate oscillators, so that

$$a \gg \sqrt{\frac{\hbar}{m\omega}} \tag{8.67}$$

the harmonic oscillator eigenfunctions $\psi_n(x - a) \quad \text{and} \quad \psi_n(x + a)$, obtained from (5.27) by shifting the coordinate origin, may be used as basis functions for a variational calculation of energy eigenvalues of the Hamiltonian

$$H = \frac{p_x^2}{2m} + \frac{1}{2} m\omega^2(|x| - a)^2 \tag{8.68}$$

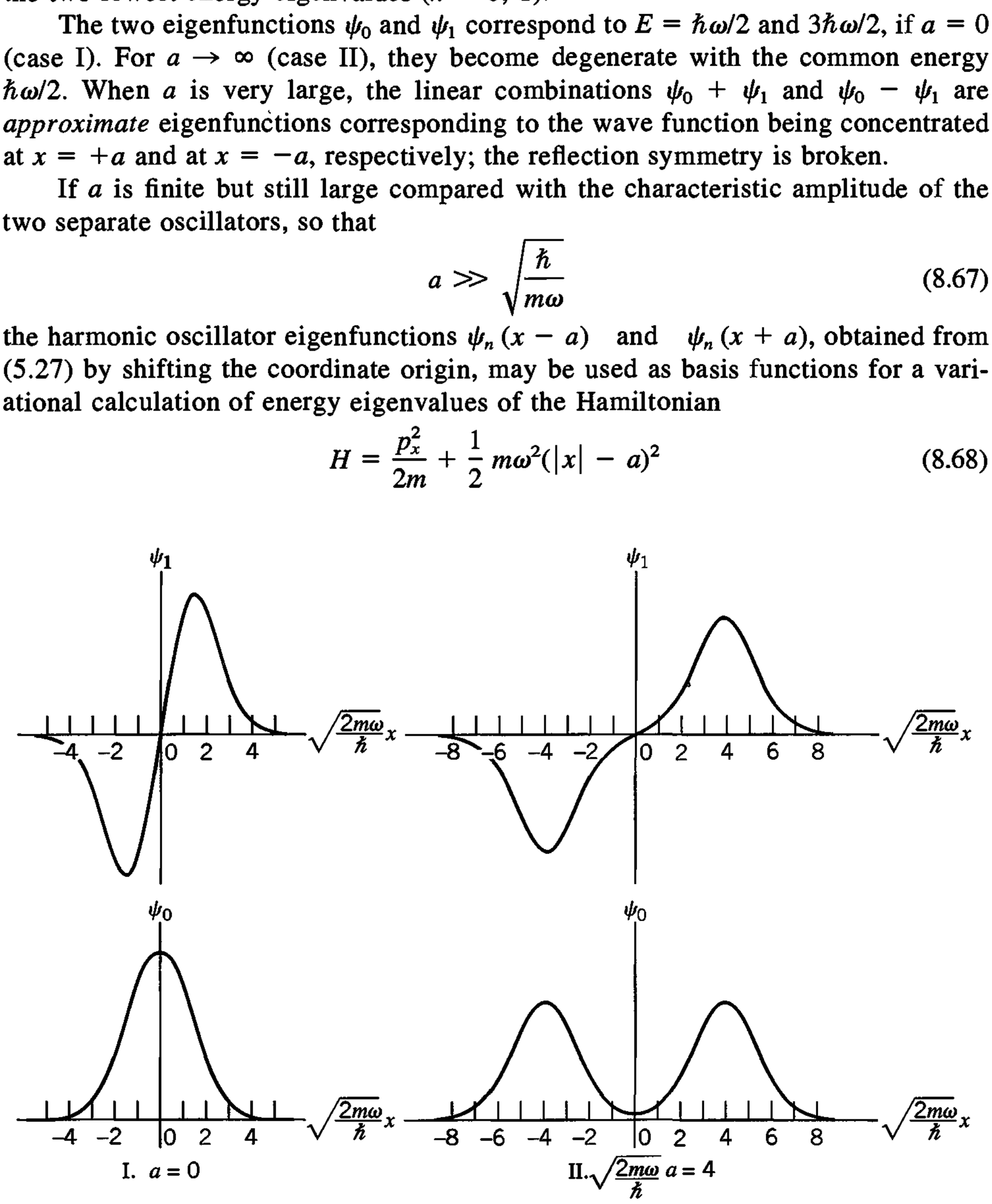


Figure 8.4. The two lowest energy eigenfunctions ψ_0 and ψ_1 for the double oscillator. In case I ($a = 0$) we have a simple harmonic oscilator, no barrier, and $E_1 - E_0 = \hbar\omega$. In case II ($\sqrt{2m\omega/\hbar}\, a = 4$), a high barrier separates the wells, and $E_1 - E_0 \approx 10^{-3}\, \hbar\omega$. The number of nodes characterizes the eigenfunctions.

at least for not too large values of n. Since the exact eigenfunctions of H, for finite a, have definite parity, reflection symmetry suggests that the even and odd trial functions

$$\boxed{\psi_{\pm} = N_{\pm}[\psi_n(x - a) \pm \psi_n(x + a)]} \tag{8.69}$$

be used. Here $N_{\pm}$ is a normalization constant. Except in the limit of case II ($a \to \infty$), the two components in (8.69) are not orthogonal; they overlap. For even n, the plus sign gives an even (or symmetric) function, and the minus sign an odd (or antisymmetric) function. These assignments are reversed for odd n. Since the even operator H does not connect states of opposite parity, so that $\langle even|H|odd\rangle = 0$, it is optimally efficient to choose $\psi_{\pm}$ as trial functions. Linear combinations of these, à la Rayleigh-Ritz, give nothing new. For the normalized trial functions, the variational estimate of energies of the double oscillator is then

$$\langle H\rangle = \int \psi_{\pm}^{*} H \psi_{\pm}\, d^3r$$

If we substitute (8.69) into this formula, we obtain for the real-valued oscillator eigenfunctions,

$$\boxed{\langle H\rangle = \frac{A_n \pm B_n}{1 \pm C_n}} \tag{8.70}$$

where

$$A_n = \int \psi_n(x - a) H \psi_n(x - a)\, dx = \int \psi_n(x + a) H \psi_n(x + a)\, dx \tag{8.71}$$

$$B_n = \int \psi_n(x - a) H \psi_n(x + a)\, dx = \int \psi_n(x + a) H \psi_n(x - a)\, dx \tag{8.72}$$

and C_n is the *overlap integral*

$$C_n = \int \psi_n(x + a)\psi_n(x - a)\, dx \tag{8.73}$$

Also

$$N_{\pm} = \frac{1}{\sqrt{2(1 \pm C_n)}} \tag{8.74}$$

Using the scaled dimensionless quantity,

$$\alpha = \sqrt{\frac{m\omega}{\hbar}}\, a \tag{8.75}$$

we have for $n = 0$,

$$A_0 = \hbar\omega\left[\frac{1}{2} + \frac{2\alpha}{\sqrt{\pi}} \int_{-\infty}^{0} \xi\, e^{-(\xi - \alpha)^2}\, d\xi\right]$$

and for large α:

$$A_0 \cong \hbar\omega\left[\frac{1}{2} - \frac{1}{2\sqrt{\pi}\,\alpha} e^{-\alpha^2} + O\left(\frac{e^{-\alpha^2}}{\alpha^2}\right)\right]$$

Similarly,

$$B_0 = \hbar\omega\left[-\frac{\alpha}{\sqrt{\pi}} + \frac{1}{2}\right]e^{-\alpha^2} \quad \text{and} \quad C_0 = e^{-\alpha^2}$$

For $\alpha \gg 1$, the leading terms give $A_0 = \hbar\omega/2$, $B_0 = -(\alpha/\sqrt{\pi})\, e^{-\alpha^2}\, \hbar\omega$, $C_0 = 0$, and thus

$$\langle H\rangle \approx \hbar\omega\left(\frac{1}{2} \mp \frac{\alpha}{\sqrt{\pi}}\, e^{-\alpha^2}\right) \tag{8.76}$$

showing that the degenerate ground state for $a \to \infty$ splits symmetrically as a decreases. The even state is the ground state, and the odd state is the first excited state. If the height of the potential barrier between the two wells,

$$V_0 = \frac{1}{2}\, m\omega^2 a^2 = \frac{1}{2}\, \hbar\omega\alpha^2 \tag{8.77}$$

is introduced, the splitting of the two lowest energy levels, for the double oscillator with $V_0 \gg \hbar\omega$, is

$$\boxed{E_- - E_+ = \Delta E = 2\hbar\omega\sqrt{\frac{2V_0}{\pi\hbar\omega}}\exp\left(-\frac{2V_0}{\hbar\omega}\right)} \tag{8.78}$$

Exercise 8.15. Work out the above expressions for A_0, B_0, and C_0 for the double oscillator (8.68), and derive the asymptotic value of $\langle H\rangle$ as given in (8.76).

The frequency corresponding to the energy splitting (8.78) is

$$\omega_s = \frac{\Delta E}{\hbar} = 2\omega\sqrt{\frac{2V_0}{\hbar\omega\pi}}\exp\left(-\frac{2V_0}{\hbar\omega}\right) \tag{8.79}$$

The physical significance of the frequency ω_s is best appreciated if we consider a system that is initially represented by a wave function (again assuming $n = 0$)

$$\psi(x, 0) = \frac{1}{\sqrt{2}}\,[\psi_+(x) + \psi_-(x)] \tag{8.80}$$

If the two wells are widely separated ($\alpha \gg 1$), so that the overlap integral C_0 is very small, ψ_+ and ψ_- tend to cancel for $x < 0$ and the initial wave packet is a single peak around $x = +a$. If $x = x_1 - x_2$ is the coordinate difference for two elastically bound particles, as described at the beginning of this section, the initial condition (8.80) implies that at $t = 0$ particle 1 is definitely to the right of particle 2, breaking the reflection symmetry. The time development of the system is then

$$\begin{aligned}\psi(x, t) &= \frac{1}{\sqrt{2}}\left[\exp\left(-\frac{i}{\hbar}\, E_+ t\right)\psi_+(x) + \exp\left(-\frac{i}{\hbar}\, E_- t\right)\psi_-(x)\right] \\ &= \exp\left[-\frac{i}{2\hbar}\,(E_+ + E_-)t\right]\left[\frac{\psi_+ + \psi_-}{\sqrt{2}}\cos\frac{\omega_s}{2}\, t + i\,\frac{\psi_+ - \psi_-}{\sqrt{2}}\sin\frac{\omega_s}{2}\, t\right]\end{aligned} \tag{8.81}$$

This last form shows that the system shuttles back and forth between $(\psi_+ + \psi_-)/\sqrt{2}$ and $(\psi_+ - \psi_-)/\sqrt{2}$ with the frequency $\omega_s/2$. It takes a time $\tau_s = \pi/\omega_s$ for the system to change from its initial state $(\psi_+ + \psi_-)/\sqrt{2}$ to the state

$(\psi_+ - \psi_-)/\sqrt{2}$, which means that the particles have exchanged their places in an entirely unclassical fashion, and the wave function is now peaked around $x = -a$. Since E was assumed to be much less than V_0, this exchange requires that the system tunnel with frequency $\omega_s/2$ through the classically inaccessible barrier region indicated in Figure 8.3. The ratio of the time τ_s during which the exchange takes place and the period $2\pi/\omega$ of the harmonic oscillator is given approximately by

$$\frac{\omega\tau_s}{2\pi} = \frac{\omega}{2\omega_s} = \frac{1}{4}\sqrt{\frac{\hbar\omega\pi}{2V_0}}\exp\left(\frac{2V_0}{\hbar\omega}\right) \tag{8.82}$$

If the barrier V_0 is high compared to $\hbar\omega$, the tunneling is strongly inhibited by the presence of the exponential factor.

Exercise 8.16. Show explicitly that the expectation value of the parity operator for the wave packet (8.81) is zero at all times. [The parity operator changes $\psi(x)$ into $\psi(-x)$.]

The situation is similar for the higher energy levels (Fig. 8.5). Asymptotically, at $a \to \infty$, the doubly degenerate spectrum is $E = \hbar\omega(n + 1/2)$. As the separation a of the two wells decreases from ∞ to 0, the lower (even) energy level in each split pair of levels first decreases and eventually increases toward $E = \hbar\omega(2n + 1/2)$.

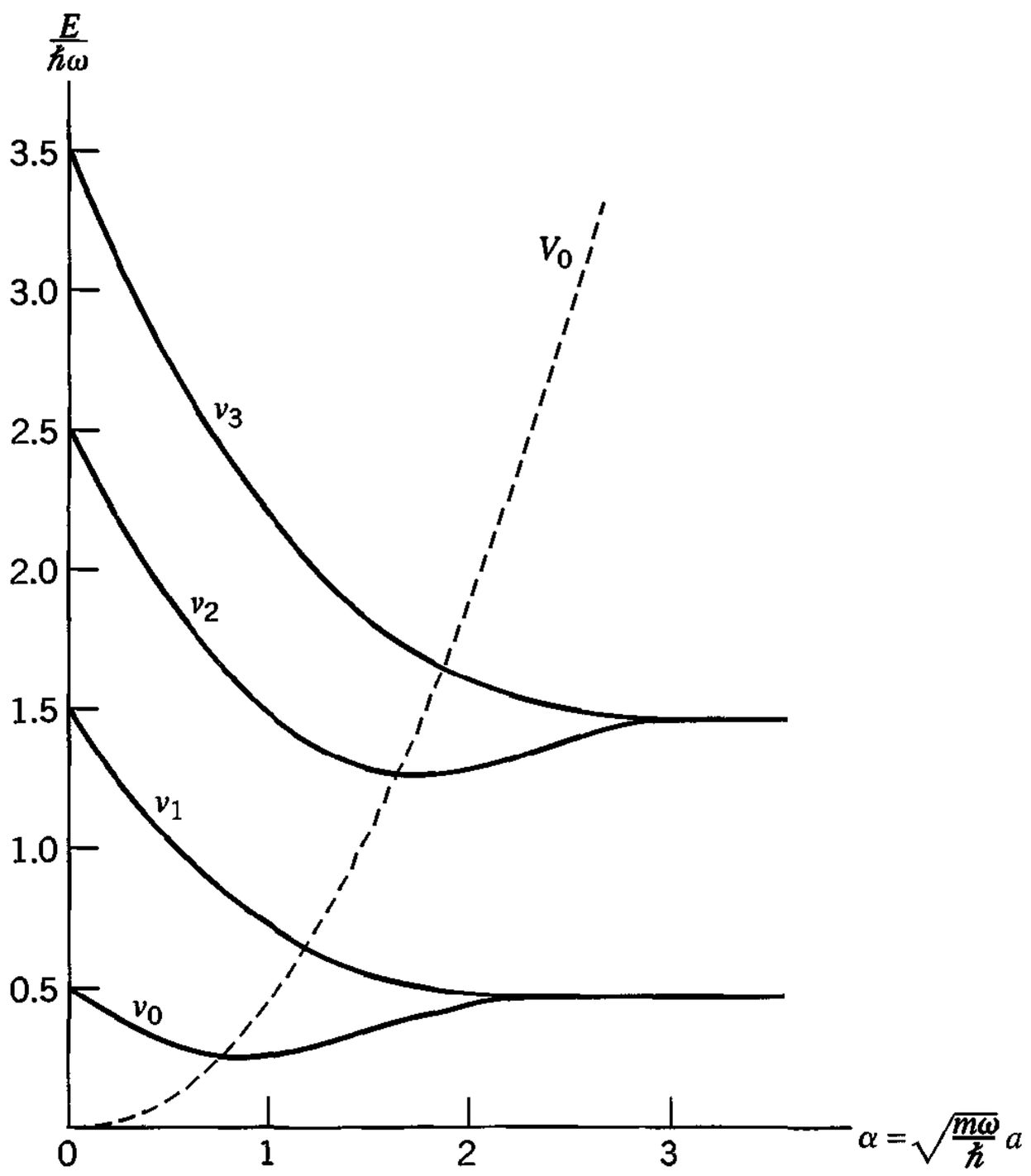


Figure 8.5. The energy $E = \hbar\omega\left(\nu + \frac{1}{2}\right)$ in units of $\hbar\omega$ versus $a = \sqrt{m\omega/\hbar}\, a$ for the four lowest energy eigenstates of the double oscillator, $V(x) = m\omega^2(|x| - a)^2/2$. For comparison with the energy levels, the dashed curve shows the barrier height V_0.

The upper (odd) level increases monotonically toward $E = \hbar\omega(2n + 3/2)$. To obtain the energy eigenvalues for an arbitrary value of a, the Schrödinger equation (8.64) must be solved accurately.

The exact solution of (8.64) is facilitated by substituting for positive x the new variable,

$$z = \sqrt{\frac{2m\omega}{\hbar}}\,(x - a) \qquad (x > 0)$$

and expressing the energy as

$$E = \hbar\omega\left(\nu + \frac{1}{2}\right) \tag{8.83}$$

in terms of the new quantum number ν, which is generally not an integer. We obtain for positive x the differential equation

$$\frac{d^2\psi}{dz^2} + \left(\nu + \frac{1}{2} - \frac{z^2}{4}\right)\psi = 0 \tag{8.84}$$

For negative x the same equation is obtained, except that now the substitution

$$z' = \sqrt{\frac{2m\omega}{\hbar}}\,(x + a) \qquad (x < 0)$$

must be used. The differential equation valid for negative x is

$$\frac{d^2\psi}{dz'^2} + \left(\nu + \frac{1}{2} - \frac{z'^2}{4}\right)\psi = 0 \tag{8.85}$$

which has the same form as (8.84). The boundary condition $\psi \to 0$ as $x \to \pm\infty$ implies that we must seek solutions of (8.84) and (8.85) which vanish as $z \to +\infty$ and $z' \to -\infty$. For $a = 0$, which is the special case of the simple linear harmonic oscillator, we have $z' = z$, and the two equations become identical.

Instead of proceeding to a detailed treatment of differential equation (8.84), we refer to the standard treatises on mathematical analysis for the solutions.[2] The particular solution of (8.84) which vanishes for very large positive values of z is called a *parabolic cylinder function*. It is denoted by $D_\nu(z)$ and is defined as

$$\begin{aligned} D_\nu(z) = 2^{\nu/2}e^{-z^2/4}\Bigg[&\frac{\Gamma(1/2)}{\Gamma[(1-\nu)/2)}\,{}_1F_1\left(-\frac{\nu}{2};\frac{1}{2};\frac{z^2}{2}\right) \\ &+ \frac{z}{\sqrt{2}}\,\frac{\Gamma(-1/2)}{\Gamma(-\nu/2)}\,{}_1F_1\left(\frac{1-\nu}{2};\frac{3}{2};\frac{z^2}{2}\right)\Bigg] \end{aligned} \tag{8.86}$$

The function ${}_1F_1$ is the *confluent hypergeometric* (or *Kummer*) *function*. Its power series expansion is

$${}_1F_1(a;c;z) = 1 + \frac{a}{c}\,\frac{z}{1!} + \frac{a(a+1)}{c(c+1)}\,\frac{z^2}{2!} + \cdots \tag{8.87}$$

[2]Magnus and Oberhettinger (1949), Morse and Feshbach (1953), Abramowitz and Stegun (1964), Mathews and Walker (1964), Thompson (1997).

If z is large and positive ($z \gg 1$ and $z \gg |\nu|$)

$$D_\nu(z) \cong e^{-z^2/4} z^\nu \left[1 - \frac{\nu(\nu-1)}{2 \cdot z^2} + \frac{\nu(\nu-1)(\nu-2)(\nu-3)}{2 \cdot 4 \cdot z^4} - \cdots \right] \tag{8.88}$$

and if z is large and negative ($z \ll -1$ and $z \ll -|\nu|$),

$$\begin{aligned} D_\nu(z) \cong e^{-z^2/4} z^\nu & \left[1 - \frac{\nu(\nu-1)}{2 \cdot z^2} + \frac{\nu(\nu-1)(\nu-2)(\nu-3)}{2 \cdot 4 \cdot z^4} - \cdots \right] \\ - \frac{\sqrt{2\pi}}{\Gamma(-\nu)} e^{\nu\pi i} e^{z^2/4} z^{-\nu-1} & \left[1 + \frac{(\nu+1)(\nu+2)}{2 \cdot z^2} \right. \\ & \left. + \frac{(\nu+1)(\nu+2)(\nu+3)(\nu+4)}{2 \cdot 4 \cdot z^4} + \cdots \right] \end{aligned} \tag{8.89}$$

Although the series in the brackets all diverge for any finite value of z, (8.88) and (8.89) are useful asymptotic expansions of $D_\nu(z)$.

Exercise 8.17. Using the identity

$$\Gamma(1+\nu)\Gamma(-\nu) = -\frac{\pi}{\sin \nu\pi} \tag{8.90}$$

obtain the eigenvalues of the simple harmonic oscillator from (8.89). Show the connection between the parabolic cylinder functions (8.86) and the eigenfunctions (5.27).

If $D_\nu(z)$ is a solution of (8.84), $D_\nu(-z)$ is also a solution of the same equation, and these two solutions are linearly independent unless ν is a nonnegative integer. Inspection of the asymptotic behavior shows that the particular solution of (8.85) which vanishes for very large negative values of z' must be $D_\nu(-z')$. It follows that a double oscillator eigenfunction must be proportional to $D_\nu(z)$ for positive values of x and proportional to $D_\nu(-z')$ for negative values of x. The remaining task is to join these two solutions smoothly at $x = 0$, the point where the two parabolic potentials meet with discontinuous slope.

As was discussed earlier, the eigenfunctions can be assumed to have definite parity, even or odd. The smooth joining condition requires that at $x = 0$ the even functions have zero slope and the odd functions must vanish. Matching ψ and ψ' at $x = 0$ thus leads to

$$D'_\nu\left(-\sqrt{\frac{2m\omega}{\hbar}}\, a \right) = 0 \tag{8.91}$$

for even ψ, and

$$D_\nu\left(-\sqrt{\frac{2m\omega}{\hbar}}\, a \right) = 0 \tag{8.92}$$

for odd ψ. These are transcendental equations for ν.

Exercise 8.18. Show that if $a = 0$, the roots of (8.91) and (8.92) give the eigenvalues of the simple harmonic oscillator. [Use the properties of the gamma function embodied in the identity (8.90).]

In general, the roots ν of (8.91) and (8.92) are obtained by numerical computation. Figure 8.5 shows how the lowest eigenvalues depend on the parameter α defined in (8.75). A few of the corresponding eigenfunctions are plotted in Figure 8.4. The unnormalized eigenfunctions are

$$\psi(x) = \begin{cases} D_\nu\left(\sqrt{\dfrac{2m\omega}{\hbar}}\,(x - a)\right) & (x \geq 0) \\ \pm D_\nu\left(-\sqrt{\dfrac{2m\omega}{\hbar}}\,(x + a)\right) & (x \leq 0) \end{cases} \tag{8.93}$$

where the upper sign is to be used if ν is a root of $D'_\nu(-\sqrt{2}\ \alpha) = 0$, and the lower sign if ν is a root of $D_\nu(-\sqrt{2}\ \alpha) = 0$.

Exercise 8.19. A convenient representation of a different anharmonic double oscillator is the potential (Figure 8.6):

$$V(x) = \frac{m\omega^2}{8a^2}\,(x - a)^2(x + a)^2 = \frac{m\omega^2}{8a^2}\,(x^2 - a^2)^2 \tag{8.94}$$

For large values of the distance between the minima, estimate the splitting of the two lowest energy levels.

Two (or more) potential wells separated by a potential barrier as in (8.62) and (8.94) occur frequently in many branches of physics. The dynamics of the vibrations of atoms in molecules is the most prominent example. Hindered rotations of a molecule, described by an angle-dependent potential such as

$$V(\theta) = V_0(\cos 2\theta - 1) \qquad (V_0 > 0) \tag{8.95}$$

also represent typical features of a double-well potential, but the boundary conditions for the Schrödinger equation are different from those for a potential that depends on the linear coordinate x.

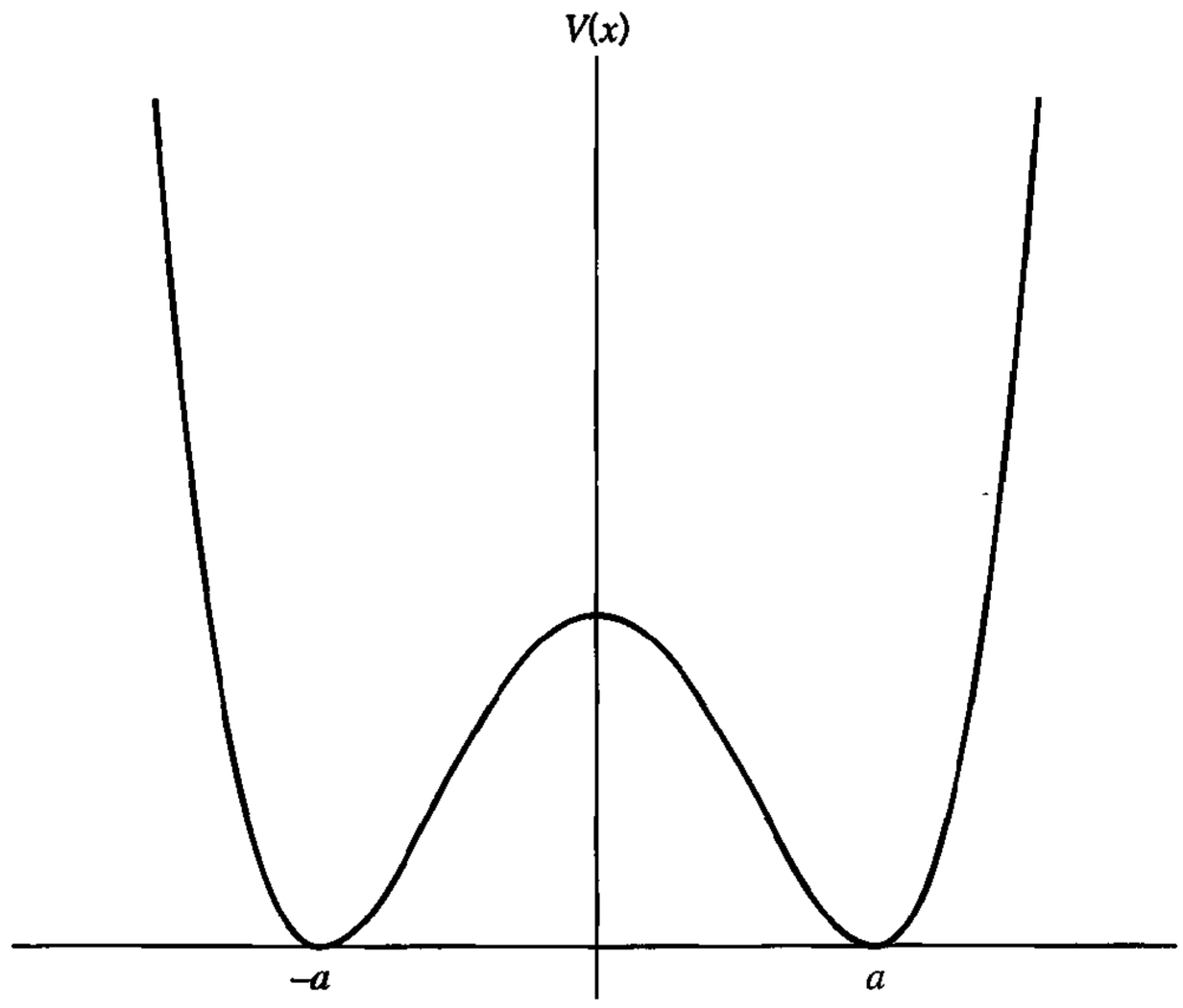


Figure 8.6. Quartic double-well potential, defined by Eq. (8.94).

As a numerical example of molecular vibrations, consider a molecule in which the equilibrium distance between the atoms is $a = 1$ Å, and the reduced mass $m = 10^{-27}$ kg. For $\omega = 2 \times 10^{15}$ Hz (corresponding to an infrared vibration spectrum common in many molecules), we have $V_0/\hbar\omega \approx 10^3$. From (8.82), the shuttling period τ_s turns out to be superastronomical (about 10^{400} years); hence, the exchange "hardly ever" takes place. On the other hand, if $\omega = 10^{12}$ Hz, corresponding to microwave vibrations of 0.2 cm wavelength, the value of ω_s obtained from Figure 8.5 is approximately 8×10^{11} Hz, which is again in the microwave region. Transitions corresponding to such barrier tunneling oscillations are commonly observed in microwave spectroscopy.[3]

6. ***The Molecular Approximation.*** Although many physical systems are usefully modeled by simple one-particle Hamiltonians of the form

$$H = \frac{p^2}{2m} + V(\mathbf{r})$$

with a prescribed external potential $V(\mathbf{r})$, this fiction cannot always be maintained. In most real physical processes two or more partial systems interact dynamically in significant ways. Since it is generally far too difficult to treat such a composite system exactly, we must resort to various approximation methods, based on simplifying features of the system and its components. In this section we consider an approximation scheme that yields useful zeroth-order energy eigenfunctions for starting perturbation calculations in composite systems where the masses of the interacting subsystems are grossly different.

The dynamics of interacting atoms offers a valuable example of a complex composite system whose subsystems (the valence electrons and the atomic cores) can be treated as if they were autonomous though they interact strongly. Molecular structure, ion–atom collisions, and condensed matter physics are among the important applications. Because of the large (three to four orders of magnitude) disparity in mass between the electrons and atoms, we think of the atomic cores as moving relatively slowly under the influence of their mutual interaction and in the average field produced by the fast-moving electrons. Conversely, the electrons respond to the nearly static field of the cores, in addition to their own mutual potential energy.

We consider the Hamiltonian for a model of a molecule composed of two (or more) massive atomic cores (nuclei plus electrons in tightly bound inert inner shells) and some loosely bound atomic valence electrons that move in the common field of the cores. This is indicated schematically in Figure 8.7.

In the simplest case of two positively charged atomic cores and just one shared valence electron, the Hamiltonian may be written as

$$\boxed{H = \frac{P^2}{2M} + \frac{p^2}{2m} + V(\mathbf{R}, \mathbf{r}) + V_n(\mathbf{R})} \tag{8.96}$$

where $\mathbf{R}$ and $\mathbf{P}$ are the relative canonical coordinates of the slowly moving nuclear framework with reduced mass M, and $\mathbf{r}$ and $\mathbf{p}$ refer to the electron of mass m which moves much more swiftly. The potential $V(\mathbf{R}, \mathbf{r})$ stands for the mutual interaction

[3]Townes and Schawlow (1955).

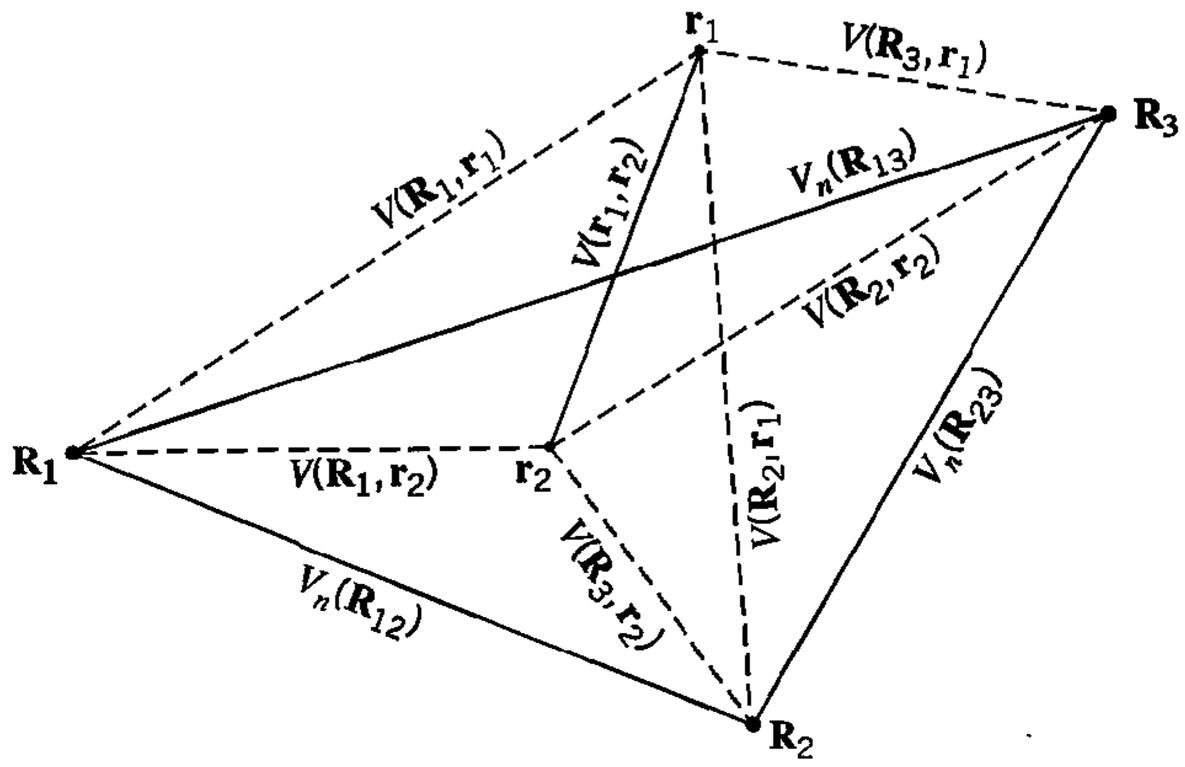


Figure 8.7. Two valence electrons ($\mathbf{r}_1$ and $\mathbf{r}_2$) in a molecule moving in the field of three heavy atomic cores ($\mathbf{R}_1$,$\mathbf{R}_2$,$\mathbf{R}_3$). The ten interactions between the five centers of force are indicated.

between the cores and the electrons, and for the Coulomb repulsion between the electrons. The core-core interaction is explicitly accounted for by the potential $V_n(\mathbf{R})$, which is chiefly due to the Coulomb forces between the charges.

In the limit $M \to \infty$ the nuclear skeleton is frozen in its position, and the electrons move in the static potential field $V(\mathbf{R}, \mathbf{r})$, where the external parameter $\mathbf{R}$ is fixed. The Schrödinger equation for the (quasi-) stationary states of the electron,

$$\boxed{\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{R}, \mathbf{r})\right]\phi_{\mathbf{R}}^{(i)}(\mathbf{r}) = E_e^{(i)}(\mathbf{R})\,\phi_{\mathbf{R}}^{(i)}(\mathbf{r})} \tag{8.97}$$

corresponds to the subsystem Hamiltonian for the electron:

$$H_e = \frac{p^2}{2m} + V(\mathbf{R}, \mathbf{r}) \tag{8.98}$$

In principle, once (8.97) is solved for the complete set of quasi-stationary states for all $\mathbf{R}$, the energy eigenfunction of the complete molecular Hamiltonians (8.96) may be expanded as

$$\psi(\mathbf{R}, \mathbf{r}) = \sum_i \eta^{(i)}(\mathbf{R})\phi_{\mathbf{R}}^{(i)}(\mathbf{r}) \tag{8.99}$$

Substitution of this expression in the Schrödinger equation for the composite system,

$$H\psi(\mathbf{R}, \mathbf{r}) = E\psi(\mathbf{R}, \mathbf{r}) \tag{8.100}$$

yields a set of infinitely many coupled equations for the wave functions $\eta^{(i)}(\mathbf{R})$. How does one avoid having to cope with such an unwieldy problem?

The simplest idea is to assume that the eigenstates of (8.100), at least near the ground state, can be reasonably well represented by *single* terms in the expansion (8.99) and to use the product

$$\psi(\mathbf{R}, \mathbf{r}) = \eta^{(i)}(\mathbf{R})\phi_{\mathbf{R}}^{(i)}(\mathbf{r}) \tag{8.101}$$

with yet-to-be-determined $\eta^{(i)}(\mathbf{R})$, as a trial function in the variational integral $\langle H \rangle$, based on the complete Hamiltonian (8.96). To keep the discussion focused on the

essentials, we assume the solution $\phi_{\mathbf{R}}^{(i)}(\mathbf{r})$ of (8.97) to be nondegenerate; otherwise, one must work in the subspace of degenerate eigenfunctions of (8.97) as in Section 8.3, using matrices. The approximate eigenfunction (8.101) is said to constitute an *adiabatic* or *Born-Oppenheimer* approximation to the exact molecular wave function. (The term *adiabatic* refers to the semiclassical picture of the parameter $\mathbf{R}$ slowly evolving in time, but our treatment is fully quantum mechanical.)

The variational procedure requires that we add to the expectation value of the electron Hamiltonian,

$$\langle H_e \rangle = \int \eta^{(i)*}(\mathbf{R}) E_e^{(i)}(\mathbf{R}) \eta^{(i)}(\mathbf{R})\, d^3R \tag{8.102}$$

the expectation values of the kinetic and potential energies of the heavy structure. The former is

$$\left\langle \frac{P^2}{2M} \right\rangle = -\frac{\hbar^2}{2M} \int \eta^{(i)*}(\mathbf{R}) \phi_{\mathbf{R}}^{(i)*}(\mathbf{r}) \nabla_{\mathbf{R}}^2 [\eta^{(i)}(\mathbf{R}) \phi_{\mathbf{R}}^{(i)}(\mathbf{r})]\, d^3R\, d^3r$$

which can be transformed into

$$\begin{aligned} \left\langle \frac{P^2}{2M} \right\rangle = \int \eta^{(i)*}(\mathbf{R}) \Bigg\{ & \frac{1}{2M} (\mathbf{P} - \mathbf{A}_{\mathbf{R}})^2 \\ & + \frac{\hbar^2}{2M} \int [\nabla_{\mathbf{R}} \phi_{\mathbf{R}}^{(i)*}(\mathbf{r})] \cdot [\nabla_{\mathbf{R}} \phi_{\mathbf{R}}^{(i)}(\mathbf{r})]\, d^3r - \frac{1}{2M} \mathbf{A}_{\mathbf{R}}^2 \Bigg\} \eta^{(i)}(\mathbf{R})\, d^3R \end{aligned} \tag{8.103}$$

provided that the electron eigenfunctions $\phi_{\mathbf{R}}^{(i)}(\mathbf{r})$, for each value of $\mathbf{R}$, and the trial functions $\eta^{(i)}(\mathbf{R})$ are normalized to unity. The normalization ensures that the vector function $\mathbf{A}_{\mathbf{R}}$ defined as

$$\mathbf{A}_{\mathbf{R}} = i\hbar \int \phi_{\mathbf{R}}^{(i)*}(\mathbf{r}) \nabla_{\mathbf{R}} \phi_{\mathbf{R}}^{(i)}(\mathbf{r})\, d^3r \tag{8.104}$$

is real-valued. (We omit the label i, on which $\mathbf{A}_{\mathbf{R}}$ depends.)

The expectation value of H is obtained by adding (8.102) and (8.103) to the expectation value of $V_n(\mathbf{R})$ to give

$$\boxed{\langle H \rangle = \int \eta^{(i)*}(\mathbf{R}) \left[\frac{1}{2M} (\mathbf{P} - \mathbf{A}_{\mathbf{R}})^2 + U^{(i)}(\mathbf{R}) \right] \eta^{(i)}(\mathbf{R})\, d^3R} \tag{8.105}$$

where the *effective potential* $U^{(i)}(\mathbf{R})$ is defined as

$$\boxed{U^{(i)}(\mathbf{R}) = E^{(i)}(\mathbf{R}) + \frac{\hbar^2}{2M} \int |\nabla_{\mathbf{R}} \phi_{\mathbf{R}}^{(i)}(\mathbf{r})|^2\, d^3r - \frac{1}{2M} \mathbf{A}_{\mathbf{R}}^2 + V_n(\mathbf{R})} \tag{8.106}$$

Applied to (8.105), the variational principle $\delta\langle H \rangle = 0$ leads to an effective Schrödinger equation for the trial function $\eta^{(i)}(\mathbf{R})$:

$$\boxed{\left[-\frac{\hbar^2}{2M} \left(\nabla_{\mathbf{R}} + \frac{1}{i\hbar} \mathbf{A}_{\mathbf{R}} \right)^2 + U^{(i)}(\mathbf{R}) \right] \eta^{(i)}(\mathbf{R}) = E \eta^{(i)}(\mathbf{R})} \tag{8.107}$$

In solving this equation, care must be taken to ensure that the wave function $\psi(\mathbf{R}, \mathbf{r})$ for the complete system, of which $\eta^{(i)}(\mathbf{R})$ is but one factor, satisfies the correct boundary (and single-valuedness) conditions.

Exercise 8.20. Derive (8.103) and verify (8.104).

The notation $\mathbf{A}_\mathbf{R}$ for the vector field defined by (8.104) was chosen to emphasize the formal resemblance of Eq. (8.107) to the Schrödinger equation for a charged particle in the presence of an external electromagnetic field (Section 4.6). If, as is the case for simple molecules, it is possible to choose the electronic eigenfunction $\phi_\mathbf{R}^{(i)}(\mathbf{r})$ real-valued for all $\mathbf{R}$, the vector $\mathbf{A}_\mathbf{R}$ vanishes. More generally, if

$$\nabla_\mathbf{R} \times \mathbf{A}_\mathbf{R} = 0 \tag{8.108}$$

(corresponding to zero magnetic field in the electromagnetic analogue), the gauge transformation

$$\begin{aligned} \eta^{(i)}(\mathbf{R}) &\to \exp\left(-\frac{i}{\hbar}\int_{\mathbf{R}_0}^{\mathbf{R}} \mathbf{A}_\mathbf{R}\cdot d\mathbf{R}\right)\eta^{(i)}(\mathbf{R}) \\ \phi_\mathbf{R}^{(i)}(\mathbf{r}) &\to \exp\left(\frac{i}{\hbar}\int_{\mathbf{R}_0}^{\mathbf{R}} \mathbf{A}_\mathbf{R}\cdot d\mathbf{R}\right)\phi_\mathbf{R}^{(i)}(\mathbf{r}) \\ \mathbf{A}_\mathbf{R} &\to 0 \end{aligned} \tag{8.109}$$

produces no change in the product wave function $\psi(\mathbf{R}, \mathbf{r})$ of the composite system, but reduces the Schrödinger equation (8.107) to the simpler form

$$\boxed{\left[-\frac{\hbar^2}{2M}\nabla_\mathbf{R}^2 + U^{(i)}(\mathbf{R})\right]\eta^{(i)}(\mathbf{R}) = E\eta^{(i)}(\mathbf{R})} \tag{8.110}$$

Exercise 8.21. If $\phi_\mathbf{R}^{(i)}(\mathbf{r})$ is a real-valued function for all $\mathbf{R}$, show that $\mathbf{A}_\mathbf{R}$ is the gradient of a function and therefore can be eliminated by a gauge transformation (8.109).

If the Born-Oppenheimer method is applied to more complex nonlinear molecules, with $\mathbf{R}$ symbolizing a set of generalized atom core coordinates, the condition (8.108) may be violated for isolated singular values of $\mathbf{R}$, requiring the use of (8.107) rather than (8.110). The phase integral

$$\Phi = \oint \mathbf{A}_\mathbf{R}\cdot d\mathbf{R} \tag{8.111}$$

which is gauge invariant, may acquire different values for various classes of topologically distinct closed loops in the parameter space. This nonintegrability may result in the appearance of characteristic spectral features, not expected from the simple form (8.110). The vector $\mathbf{A}_\mathbf{R}$ is known as (Berry's) *connection* and the integral (8.111) as *Berry's* (geometric) *phase*.[4]

The success of any calculation in the Born-Oppenheimer approximation is contingent on the availability of good solutions $\phi_\mathbf{R}^{(i)}(\mathbf{r})$ for the parametric Schrödinger equation (8.97) for the electron(s). In quantum chemistry, these solutions are gen-

[4]See Shapere and Wilczek (1989).

erally known as *molecular orbitals*. In practice, (8.97) is again treated as a variational problem, and $\phi_{\mathbf{R}}^{(i)}(\mathbf{r})$ is represented by trial functions that are linear combinations of energy eigenfunctions of the hypothetically separated atoms, which constitute the molecule (linear combination of atomic orbitals, or LCAO). If the constituent atomic orbitals are nonorthogonal, their mutual overlap integrals play an important role in the calculations. Many-body methods (Chapter 22) are used to include the interaction between elections. From all of this, effective energy eigenvalues, $E^{(i)}(\mathbf{R})$, for the low-lying electronic states are eventually derived and, with the inclusion or neglect of the usually small correction terms in (8.106), the *potential energy surfaces* $U^{(i)}(\mathbf{R})$ are computed as functions of the nuclear configuration coordinates $\mathbf{R}$.

Exercise 8.22. Assuming the molecular orbital functions $\phi_{\mathbf{R}}^{(i)}(\mathbf{r})$ in the expansion (8.99) to be orthonormal and real, derive the exact coupled integro-differential equations for the nuclear configuration wave functions $\eta^{(i)}(\mathbf{R})$ and show that Eq. (8.110) results if certain off-diagonal matrix elements are neglected.

To illustrate the approximation method just described, we consider the schematic model of a one-dimensional homonuclear diatomic molecule, which can vibrate but not rotate and in which effectively only one electron is orbiting in the field of both atomic cores (Figure 8.8). The H_2^+ molecular ion is the simplest example. If M is the reduced mass of the molecule, X is the relative coordinate of the two nuclei, m is the mass of the electron, and x is the electron coordinate relative to the center of mass, the Schrödinger equation for this example may be written as

$$\left[-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial X^2} - \frac{\hbar}{2m}\frac{\partial^2}{\partial x^2} + V(x, X) + V_n(X)\right]\psi(x, X) = E\psi(x, X) \qquad (8.112)$$

provided that the center of mass of the entire molecule is assumed to be at rest. We also assume that $X > 0$ and thus now neglect the possibility of exchange between the two identical nuclei on the grounds that the barrier tunneling period τ_s tends to infinity. In effect, we assume that the interaction $V(x, X)$ is not invariant under the nuclear reflection $X \to -X$. The potential $V_n(X)$ represents the repulsive interaction between the two atomic cores.

The electron-core interaction is an electrostatic attraction, but we crudely simulate it by another double oscillator potential,

$$V(x, X) = \frac{m\omega^2}{2}\left(|x| - \frac{X}{2}\right)^2 \qquad (8.113)$$

with $X \approx 10^{-10}$ m being the separation of the two atomic wells. The characteristic frequency of the electron motion in the separated atom is of order $\omega = 5 \times 10^{15}$ Hz. The critical ratio $V_0/\hbar\omega = (m\omega/8\hbar)\,X^2$ is about 0.05 for this case, indicating a

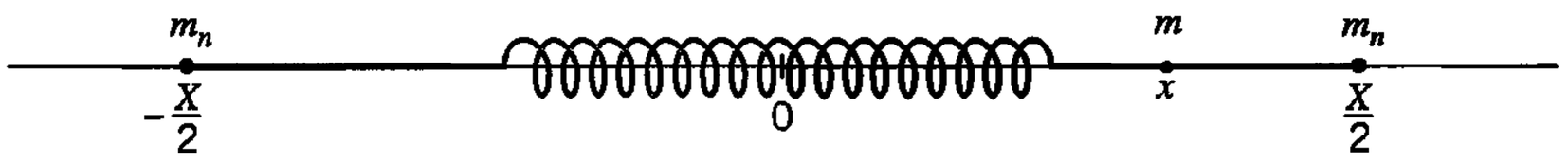


Figure 8.8. Model of a one-dimensional one-electron homonuclear diatomic molecule, with interatomic distance X. The electron (mass m) interacts with the atomic cores of mass m_n (and reduced mass $M \approx m_n/2$). The spring symbolizes the binding force between the cores, attributable in the Born-Oppenheimer approximation to rapid exchange of the electron between the slowly moving atoms.

low barrier. The electron motion cannot be localized in just one of the two wells and is shared by both of them. (For an even more primitive model, see Problems 3 and 4 in Chapter 6.)

The Born-Oppenheimer approximation for the wave function consists of assuming that the variational trial solution has the approximate simple product form

$$\psi(x, X) = \eta^{(i)}(X)\phi_X^{(i)}(x) \tag{8.114}$$

as in Eq. (8.101). Equation (8.112) separates approximately into a stationary state equation (8.97) for the molecular-orbital electron:

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{m\omega^2}{2}\left(|x| - \frac{X}{2}\right)^2\right]\phi_X^{(i)}(x) = E_e^{(i)}(X)\phi_X^{(i)}(x) \tag{8.115}$$

corresponding to a fixed value of the parameter X, and a second Schrödinger equation describing the motion of the nuclei in the presence of the potential $E_e^{(i)}(X)$:

$$\left[-\frac{\hbar^2}{2M}\frac{\partial^2}{\partial X^2} + U^{(i)}(X)\right]\eta^{(i)}(X) = E\eta^{(i)}(X) \tag{8.116}$$

In (8.116), the potential is

$$U^{(i)}(X) = E^{(i)}(X) + V_n(X) + \frac{\hbar^2}{2M}\int_{-\infty}^{+\infty}\left|\frac{\partial\phi_X^{(i)}(x)}{\partial X}\right|^2 dx \tag{8.117}$$

since the connection $\mathbf{A_R}$ vanishes for a one-dimensional system.

By inspection of the scaled Figure 8.5, we see that the potential function $E_e^{(i)}(X)$ for the two atoms has a minimum corresponding to a stable equilibrium configuration if the electron is in one of the symmetric eigenstates. Only in these electronic eigenstates is it possible for the attractive exchange interaction mediated by the shared electron to overcome the core-core repulsion, allowing the vibrating atoms to be bound in the molecule. In this simplistic model, if the system found itself in an antisymmetric state it would promptly dissociate.

Confining ourselves to the electronic ground state ($i = 0$), we see that near the minimum E^0 of $E_e^{(0)}(X)$ at $X = X_0 \approx 2\sqrt{\hbar/m\omega}$ we may write approximately

$$U^{(0)}(X) \approx E_e^{(0)}(X) + V_n(X) \approx E^0 + \frac{m\omega^2}{2}C(X - X_0)^2 + V_n(X_0)$$

where C is a number of the order of unity. It follows that the nuclei perform harmonic oscillations with a vibration frequency of the order of $\sqrt{m/M}\ \omega$. The expectation, underlying the Born-Oppenheimer approximation, of relatively slow nuclear motion, is thus confirmed. Since actual molecules move in three dimensions and can rotate, besides vibrating along the line joining the nuclei, their spectra exhibit more complex features, but the general nature of the approximations used is the same.

The electronic Schrödinger equation (8.115) with a potential that is invariant under the reflection $x \to -x$ models the force by which two similar atoms are bound together in a diatomic molecule. As we saw in Section 8.5, the bond in the symmetric eigenstates has its origin in the fact that an electron can be exchanged between the atoms and is shared by them. An understanding of the more general *covalent* chemical binding between dissimilar atoms also relies on the concept of exchange, but the mechanism depends on *pairs* of electrons with spin, rather than on single electrons as considered here.

Exercise 8.23. For a semiquantitative estimate of the properties of the one-dimensional model of a homonuclear diatomic molecule with one shared electron, assume that the atoms have mass $M = 10$ u and that the electronic energy scale of the separate oscillator "atoms" is $\hbar\omega = 10$ eV. From Figure 8.5, deduce the equilibrium distance between the atoms and the corresponding dissociation energy. Estimate the vibration frequency of the molecule.

7. ***The Periodic Potential.*** In a solid, as in a molecule, we deal with slowly moving heavy atomic cores and swift valence and conduction electrons, justifying the use of the adiabatic approximation. In this section, we apply perturbation and matrix techniques to the Schrödinger equation for a particle in the presence of a one-dimensional *periodic potential* composed of a succession of potential wells. As a useful idealization of the potential to which an electron in a crystal lattice is exposed, we assume that the lattice of potential wells extends indefinitely in both directions, although in reality the number of atoms in a crystal is large but finite.

Our experience with the Schrödinger equation for potentials that are even functions of x (harmonic oscillator, square well, etc.) has taught us that in order to derive and understand the energy eigenvalues and eigenfunctions, it is helpful first to consider the symmetry properties of the Hamiltonian. *Group theory* is the mathematical discipline that provides the tools for a systematic approach to symmetry in quantum mechanics. An introduction to the use of group representations in quantum mechanics will be given in Chapter 17, but the one-dimensional example of motions in a periodic potential can motivate and illustrate the group theoretic treatment. The relevant group for the dynamics of a particle in a periodic potential is the group of finite displacements or translations, which was introduced in Section 4.5.

If the potential $V(x)$ is periodic, such that

$$D_\xi V(x) = V(x - \xi) = V(x)$$

the Hamiltonian possesses symmetry under finite translation D_ξ by the displacement $x \to x - \xi$, for all x. Since the kinetic energy, being a derivative operator, is invariant under arbitrary translations, the symmetry of the Schrödinger equation and the Hamiltonian is expressed by

$$D_\xi H\psi_E(x) = D_\xi\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right)\psi_E(x) = HD_\xi\psi_E(x) = ED_\xi\psi_E(x) \quad (8.118)$$

From this equation, two conclusions can be drawn: The translation operator D_ξ commutes with the Hamiltonian,

$$D_\xi H - HD_\xi = [D_\xi, H] = 0$$

and if $\psi(x)$ is an eigenfunction of H, with eigenvalue E, then $D_\xi\psi_E(x)$ is also an eigenfunction of H, with the *same* eigenvalue E.

As a second-order linear homogeneous differential equation, the Schrödinger equation for a periodic potential generally admits two linearly independent eigenfunctions, $\psi_1(x)$ and $\psi_2(x)$, corresponding to an energy eigenvalue E. We may assume these to be orthogonal and suitably normalized. As eigenfunctions of H, with eigenvalue E, the displaced functions $D_\xi\psi_1(x)$ and $D_\xi\psi_2(x)$ are linear combinations of $\psi_1(x)$ and $\psi_2(x)$. By the standard methods of linear algebra, we may construct those linear

combinations of $\psi_1(x)$ and $\psi_2(x)$, which are eigenfunctions of the translation operator D_ξ. From Section 4.5 we know that the eigenfunctions of D_ξ are the Bloch functions

$$\boxed{\psi_k(x) = e^{ikx}\, u_k(x)} \tag{8.119}$$

where $u_k(x)$ is a periodic function:

$$u_k(x + \xi) = u_k(x) \tag{8.120}$$

The corresponding eigenvalues have modulus one and can be expressed as $e^{-ik\xi}$. (In Section 4.5, this eigenvalue was written as $e^{-ik'\xi}$, and k was reserved for an operator. In this section, we have no need to refer to k as an operator, and we therefore omit the prime on the real number k.) From the preceding discussion we conclude that all eigenfunctions of H may be assumed to be Bloch functions. This property of the solutions of the Schrödinger equation with a periodic potential is known as *Floquet's theorem.*

Since the Schrödinger equation is real (invariant under time reversal), if $\psi_k(x)$ is an eigenfunction, its complex conjugate,

$$[\psi_k(x)]^* = e^{-ikx} u_k^*(x) \tag{8.121}$$

is also an eigenfunction of both H and D_ξ, corresponding to the eigenvalues E and $e^{ik\xi}$, respectively. The two eigenfunctions, which physically correspond to modulated plane waves propagating in opposite directions, are certain to be linearly independent, except when $e^{ik\xi} = e^{-ik\xi}$ or $k\xi = \pi n$. In the language of group theory, the quantum number k (mod $2\pi/\xi$) labels the irreducible representations of the one-dimensional group of finite translations.

If the function (8.119) is substituted in the one-dimensional Schrödinger equation, the periodic function, which we now more specifically designate by $u_{k,E}(x)$, must satisfy the differential equation

$$-\frac{\hbar^2}{2m}\frac{d^2u_{k,E}(x)}{dx^2} - \frac{i\hbar^2 k}{m}\frac{du_{k,E}(x)}{dx} + \frac{\hbar^2k^2}{2m}u_{k,E}(x) + V(x)u_{k,E}(x) = E(k)u_{k,E}(x)$$

or, more compactly and intelligibly,

$$\boxed{H_k u_{k,E}(x) \equiv \frac{1}{2m}\left(\frac{\hbar}{i}\frac{d}{dx} + \hbar k\right)^2 u_{k,E}(x) + V(x)u_{k,E}(x) = E(k)u_{k,E}(x)} \tag{8.122}$$

In this equation the k-dependence of the eigenfunctions and eigenvalues has been made explicit. Not surprisingly in view of the connection (8.119) between $\psi_k(x)$ and $u_{k,E}(x)$, these two functions are related as in a simple gauge transformation, discussed in Section 4.7. When the periodicity condition (8.120) is imposed on the solutions of (8.122), a *dispersion function* $E = E(k)$ between the energy eigenvalues E and the quantum number k is obtained. As k is varied continuously over the real axis, the energy E ranges over continuous ''allowed'' bands. The extension of these ideas to three-dimensional periodic lattices forms the band theory of solids.

Exercise 8.24. Derive (8.122) from the Schrödinger equation for $\psi_k(x)$.

Exercise 8.25. In Eq. (8.122) make the replacement

$$k \to k + \frac{2\pi}{\xi} n$$

(changing k by a *reciprocal lattice vector*), and show that the new equation has solutions periodic in x, corresponding to the old energy $E(k)$. Also show that time-reversal invariance implies the degeneracy

$$E(-k) = E(k) \tag{8.123}$$

Thus, $E(k)$ is a periodic (but also generally multivalued) function of k with period $2\pi/\xi$, which is expressed by saying that "$E(k)$ has the same symmetry as the reciprocal lattice." Each interval defined by $(2n - 1)\pi/\xi < k \leq (2n + 1)\pi/\xi$ is known as a *Brillouin zone*. Summarizing the results of Exercise 8.25, we conclude that

$$E\left(\frac{\pi n}{\xi} - k\right) = E\left(\frac{\pi n}{\xi} + k\right) \tag{8.124}$$

Thus, the energy $E(k)$ is symmetric ("even") with respect to $k = n\pi/\xi$. A plot of the continuous function $E(k)$ based on these properties is known as the *repeated-zone scheme*. (An example will be shown in Figure 8.13.)

Before solving the Schrödinger equation explicitly for an example, we demonstrate the appearance of allowed (and forbidden) energy bands by applying perturbation theory to two extreme simplified models of an electron in a one-dimensional lattice. In the first model, we approximate the degenerate unperturbed state of a single electron by assigning to every lattice point $x = n\xi$ an energy value $E^{(0)}$, and an eigenfunction $\psi_n^{(0)}(x) = \psi_0^{(0)}(x - n\xi)$ that is narrowly concentrated at the n-th lattice site. Neglecting any overlap between these sharply localized wave functions, we assume them to be mutually orthogonal and normalized. We may imagine that in this extreme ("tight-binding") approximation the electron is confined to a lattice site by an infinite potential barrier. We then relax this condition by introducing a perturbative interaction V that connects the lattice sites and allows the electron to be shared among them (such as, by tunneling through the potential barrier). In an *ad hoc* fashion, we postulate that the matrix elements of V are to be nonzero only for the interaction between nearest neighbors:

$$\langle n-1|V|n\rangle = v \quad \text{for all } n \tag{8.125}$$

From Sections 8.2 and 8.3 we know that appropriate trial eigenfunctions are of the form

$$\psi^{(0)}(x) = \sum_n c_n \psi_n^{(0)}(x) \tag{8.126}$$

and that the amplitudes c_n are to be determined from the infinite set of simultaneous linear homogeneous equations,

$$vc_{n-1} + (E^{(0)} - E)c_n + vc_{n+1} = 0 \quad \text{for all } n \tag{8.127}$$

These equations are easily solved by the substitution

$$c_n = e^{ink\xi} \tag{8.128}$$

leading to the condition:

$$\boxed{E(k) = E^{(0)} + 2v \cos k\xi} \tag{8.129}$$

This simple model shows how the interaction causes the infinitely degenerate unperturbed energy levels to be spread out into a continuous band of allowed energies, labeled by the quantum number k and ranging between $E^{(0)} + 2v \geq E(k) \geq E^{(0)} - 2v$. The energy eigenfunctions,

$$\psi^{(0)}(x) = \sum_n e^{ink\xi}\psi_n^{(0)}(x) = e^{ikx}\sum_n e^{-ik(x-n\xi)}\psi_0^{(0)}(x - n\xi) = e^{ikx}u_k(x)$$

expressible in the form (8.119), are Bloch functions. If there are other tightly bound unperturbed energy levels, there will be a separate band for each of them. Allowed bands may overlap, but if the unperturbed levels are sufficiently far apart and the interactions sufficiently weak, forbidden energy gaps occur between them.

As a second illustrative model—at the opposite extreme from the tight-binding approximation—we consider a free electron that is perturbed by a weak periodic potential $V(x)$, with period ξ. The unperturbed energy eigenstates are represented by plane waves, e^{ikx}, and the unperturbed energy is $E^{(0)}(k) = \hbar^2k^2/2m$. This model is appropriate for sufficiently high energy levels. The critical point here is that the periodic potential has nonvanishing matrix elements only between plane wave states for which $\Delta k = 2\pi n/\xi$. (In three-dimensional terms, the **k** vectors of the two plane wave states must differ by a reciprocal lattice vector.)

Exercise 8.26. Prove that for a periodic potential $V(x)$,

$$\int_{-\infty}^{+\infty} e^{-ik_2x}V(x)e^{ik_1x}\,dx = 0 \tag{8.130}$$

unless $k_1 - k_2 = 2\pi n/\xi$. Generalize this result to three dimensions (see Exercise 4.14).

The diagonal matrix elements ($k_1 = k_2$) of the perturbing potential are responsible for a uniform first-order shift of all unperturbed energies, which can be ignored if we compensate for it by redefining the zero energy level. The simple perturbation theory of Section 8.3 shows that significant corrections to the free particle spectrum can be expected when two or more unperturbed energy eigenstates that are connected by a reciprocal lattice vector are degenerate, or nearly so. In one dimension, this occurs whenever $k_1 = -k_2 = N\pi/\xi$ (N: integer). According to formula (8.40), at and near the point of degeneracy the energy levels split and produce a forbidden gap of magnitude

$$\Delta E = 2\left|\frac{1}{\xi}\int_0^{\xi} e^{2\pi iNx/\xi}V(x)\,dx\right| \tag{8.131}$$

The integral in (8.131) is simply the N-th Fourier component of the periodic potential.

Finally, to avoid the extreme assumptions of the previous two simple models, we proceed to an exact calculation of the band structure for the periodic, battlement-shaped potential of Figure 8.9. This is known as the *Kronig-Penney* potential. Its period is $\xi = 2a + 2b$, where $2b$ is the width of the square wells and $2a$ is the length of the segments between them. The matrix method introduced in Chapter 6 is especially well suited for treating this problem. The solution of the Schrödinger equa-

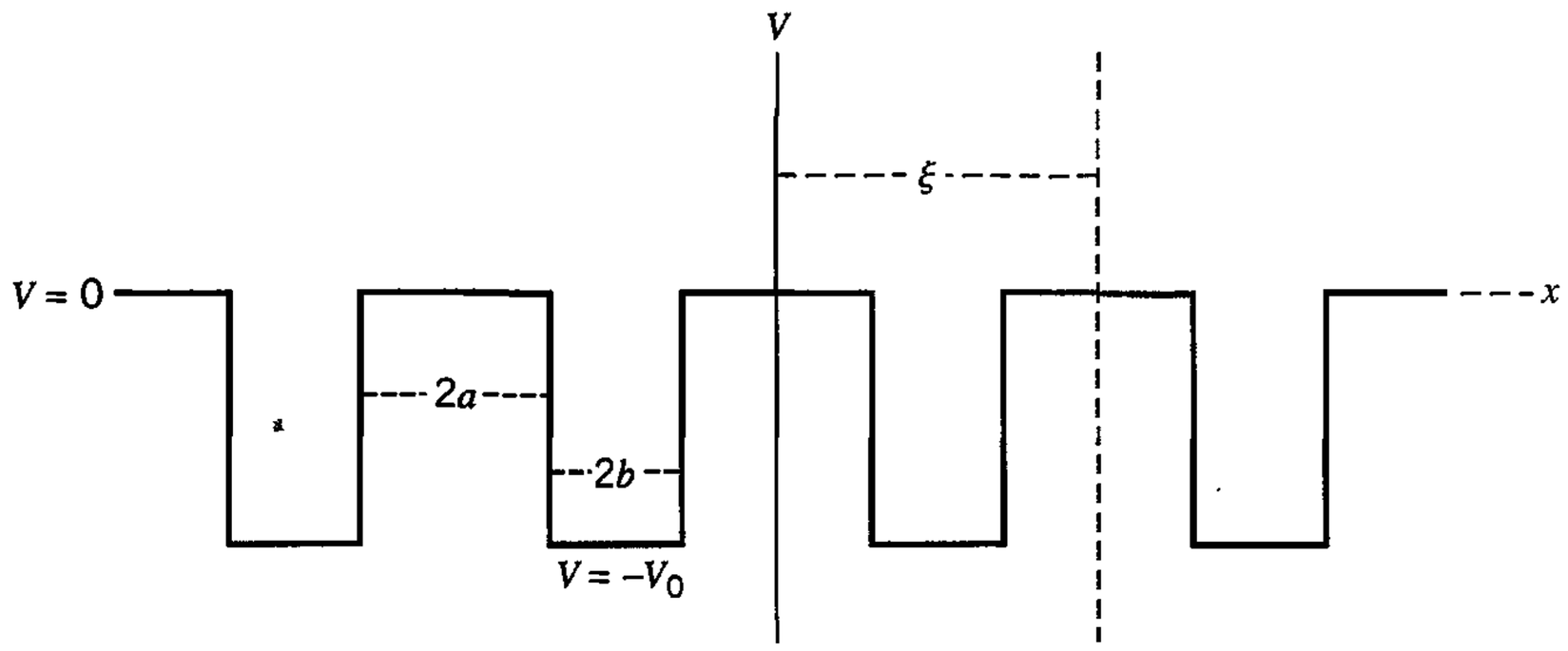


Figure 8.9. Periodic potential with rectangular sections (Kronig-Penney). The period has length $\xi = 2a + 2b$, and the well depth is V_0.

tion inside the square wells, where $V = -V_0$ and $\hbar k' = \sqrt{2m(E + V_0)}$, may be written in the form

$$\psi(x) = A_n e^{ik'(x-n\xi)} + B_n e^{-ik'(x-n\xi)} \tag{8.132}$$

for $a + (n - 1)\xi < x < n\xi - a$. Here n can be any integer, positive, negative, or zero. The coefficients belonging to successive values of n can be related by a matrix using the procedure and notation of Sections 6.2–6.4. Noting that the centers of the plateaus between the potential wells have the coordinates $x = n\xi$, we obtain

$$\begin{pmatrix} A_n \\ B_n \end{pmatrix} = \begin{pmatrix} \alpha_1 + i\beta_1 & i\beta_2 \\ -i\beta_2 & \alpha_1 - i\beta_1 \end{pmatrix} \begin{pmatrix} A_{n+1}e^{-ik'\xi} \\ B_{n+1}e^{ik'\xi} \end{pmatrix}$$

This may also be written as

$$\begin{pmatrix} A_{n+1} \\ B_{n+1} \end{pmatrix} = P \begin{pmatrix} A_n \\ B_n \end{pmatrix} \tag{8.133}$$

where the *transfer matrix* P is defined by

$$P = \begin{pmatrix} (\alpha_1 - i\beta_1)e^{ik'\xi} & -i\beta_2 e^{ik'\xi} \\ i\beta_2 e^{-ik'\xi} & (\alpha_1 + i\beta_1)e^{-ik'\xi} \end{pmatrix} \tag{8.134}$$

subject to the condition

$$\det P = \alpha_1^2 + \beta_1^2 - \beta_2^2 = 1 \tag{8.135}$$

By iteration we have

$$\begin{pmatrix} A_n \\ B_n \end{pmatrix} = P^n \begin{pmatrix} A_0 \\ B_0 \end{pmatrix} \tag{8.136}$$

Applying these considerations to an infinite periodic lattice, we must demand that as $n \to \pm\infty$ the matrix P^n should generate an acceptable finite wave function by the rule (8.136). This requirement is most conveniently discussed in terms of the eigenvalue problem of matrix P. Equation (8.133) shows that the transfer matrix represents the translation operator $D_{-\xi}$. Hence, the eigenvalues p of P may be expressed, with real-valued k, as

$$p_\pm = e^{\pm ik\xi} \tag{8.137}$$

The eigenvalues of P are roots of the characteristic equation

$$\det(P - pI) \equiv p^2 - p \operatorname{trace} P + \det P = 0$$

or

$$p^2 - 2(\alpha_1 \cos k'\xi + \beta_1 \sin k'\xi)p + 1 = 0$$

The roots are

$$p_\pm = \frac{1}{2}\left[\text{trace } P \pm \sqrt{(\text{trace } P)^2 - 4}\right] \tag{8.138}$$

An acceptable solution is obtained, and a particular energy value is allowed only if

$$\left|\frac{p_+ + p_-}{2}\right| \equiv \frac{1}{2}\left|\text{trace } P\right| = \left|\alpha_1 \cos k'\xi + \beta_1 \sin k'\xi\right| \leq 1 \tag{8.139}$$

or, using (8.137),

$$\cos k\xi = \alpha_1 \cos k'\xi + \beta_1 \sin k'\xi \tag{8.140}$$

This is the desired dispersion relationship between k and E.

If the roots (8.138) are unequal, or $k\xi \neq N\pi$, two linearly independent solutions of the Schrödinger equation are obtained by identifying the initial values $\begin{pmatrix} A_0 \\ B_0 \end{pmatrix}$ with the eigenvectors in:

$$P\begin{pmatrix} A_0^{(\pm)} \\ B_0^{(\pm)} \end{pmatrix} = p_\pm \begin{pmatrix} A_0^{(\pm)} \\ B_0^{(\pm)} \end{pmatrix} \tag{8.141}$$

For the potential shape of Figure 8.9, α_1 and β_1 can be read off Eq. (6.12), and we obtain for $E < 0$ the dispersion relationship,

$$\cos k\xi = \cosh 2\kappa a \cos 2k'b + \frac{\kappa^2 - k'^2}{2\kappa k'} \sinh 2\kappa a \sin 2k'b \tag{8.142}$$

where $\hbar\kappa = \sqrt{-2mE}$.

Since $\left|\cosh 2\kappa a\right| \geq 1$, it is readily seen from (8.142) that all energy values for which

$$2k'b = N\pi \qquad (N = \text{integer}) \tag{8.143}$$

are forbidden or are at edges of allowed bands. From the continuity of all the functions involved it follows that there must generally be forbidden ranges of energy values in the neighborhood of the discrete values determined by (8.143).

Exercise 8.27. Show that if $E > 0$, the eigenvalue condition for the periodic potential of Figure 8.9 becomes

$$\cos k\xi = \cos 2k''a \cos 2k'b - \frac{k'^2 + k''^2}{2k'k''} \sin 2k''a \sin 2k'b \tag{8.144}$$

where $\hbar k'' = \sqrt{2mE}$. Verify that the energies determined by the condition $2k''a + 2k'b = N\pi$ are forbidden.

In Figure 8.10 the right-hand side of Eqs. (8.142) and (8.144) is plotted as a function of k' for a particular choice of the constants of the potential. From this plot it is then possible to relate the values of the energy $E(k)$ to the parameter k. The condition $\left|\cos k\xi\right| \leq 1$ separates the allowed bands from the forbidden bands unambiguously in this one-dimensional potential model. Since, by (8.137), k is defined

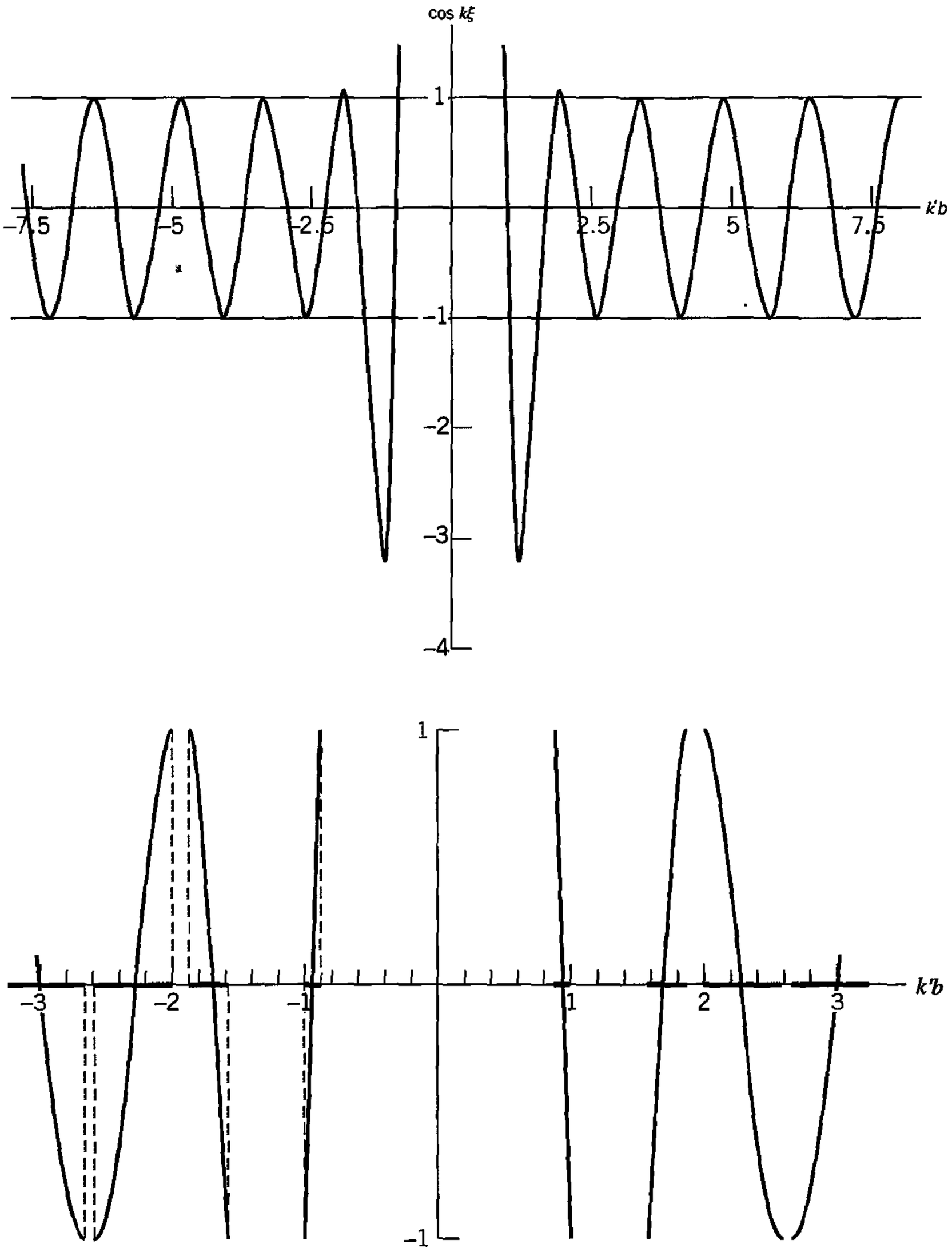


Figure 8.10. Plot of the right-hand-side of Eqs. (8.142) and (8.144) as a function of k' for a Kronig-Penney potential (Figure 8.9) with the choice of $4a = 4b = \xi$ for the linear dimensions and $\sqrt{2mb^2V_0} = \pi\hbar/2$ for the well depth. Since the plotted function equals $\cos k\xi$, only the segments of the curve between -1 and $+1$ are allowed. The enlarged detail plot shows the allowed bands (heavy black segments on the abscissa) and the forbidden energy gaps between them and illustrates the transition from tight binding to the nearly-free regime at high energies. The high peak of the curve at $k' = 0$ is suppressed.

only to within integral multiples of $2\pi/\xi$, it is possible to restrict its range so that $-\pi < k\xi \leq \pi$ and regard the energy E as a *multivalued* function of the reduced variable k. This representation of E versus k, shown in Figure 8.11, is called the *reduced-zone scheme*. It is also possible, however, to let k range from $-\infty$ to $+\infty$ and to divide the k-space into zones of length $2\pi/\xi$. In this *extended-zone scheme*, the successive energy bands are assigned to neighboring zones. Figure 8.12 illustrates the extended-zone scheme and shows that the energy spectrum consists of continuous bands separated by forbidden gaps.[5] It is natural to associate the lowest energy with the first (Brillouin) zone, $-\pi < k\xi \leq \pi$, and to let continuity and

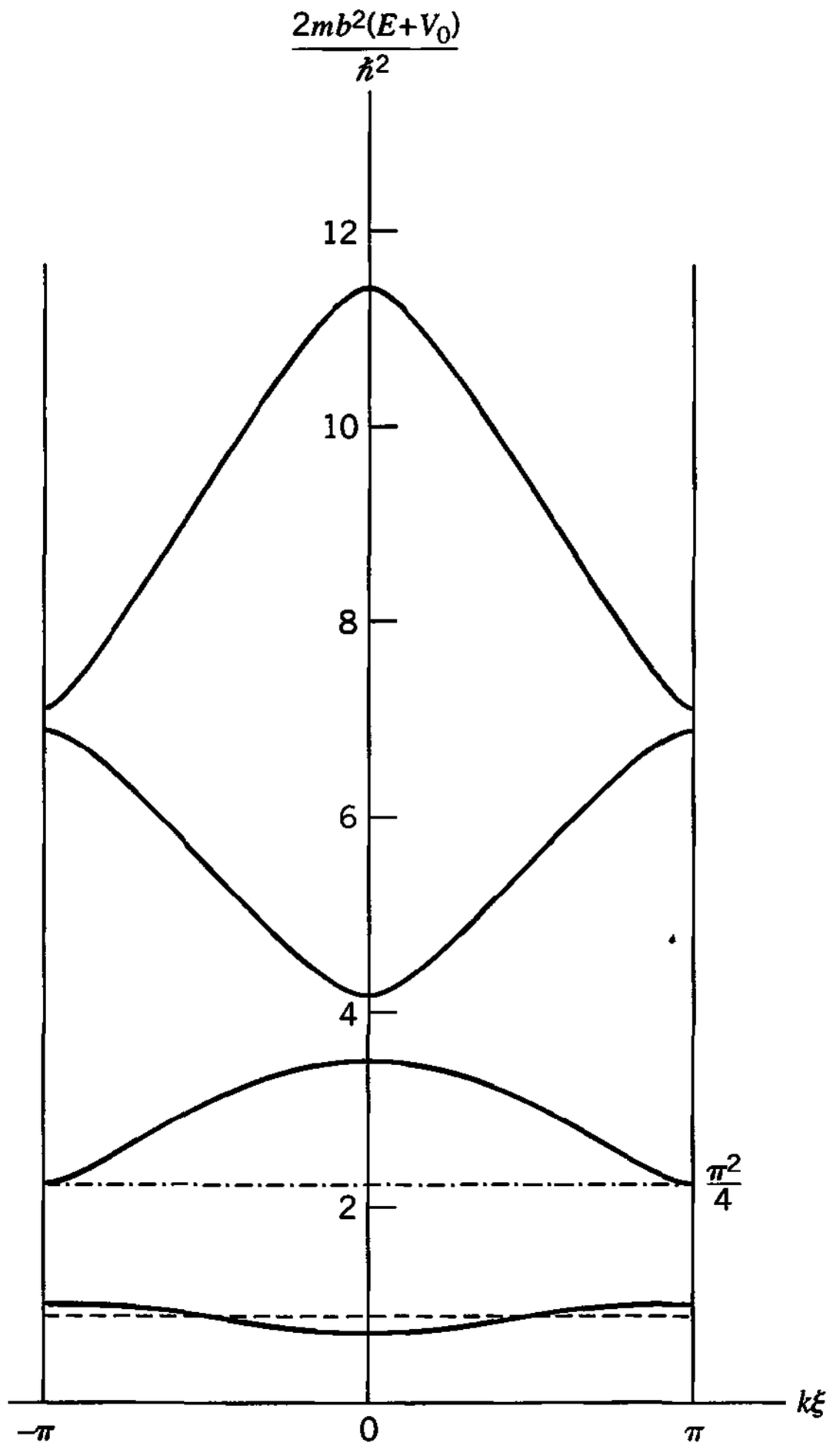


Figure 8.11. The allowed energies E versus k, in the reduced Brillouin zone scheme, for the Kronig-Penney potential defined in Figures 8.9 and 8.10, with $V_0 = \pi^2\hbar^2/8mb^2$. The dashed line indicates the energy $E = -1.60\hbar^2/2mb^2$ of the single bound level for the isolated well of width $2b$ and depth V_0 (see Figure 6.6). The dot-dashed line at $\pi^2/4$ corresponds to $E = 0$, which is the energy at the top of the barrier.

[5]Christman (1988), Ashcroft and Mermin (1976), and Dekker (1957). Insight can be gained from Weinreich (1965).

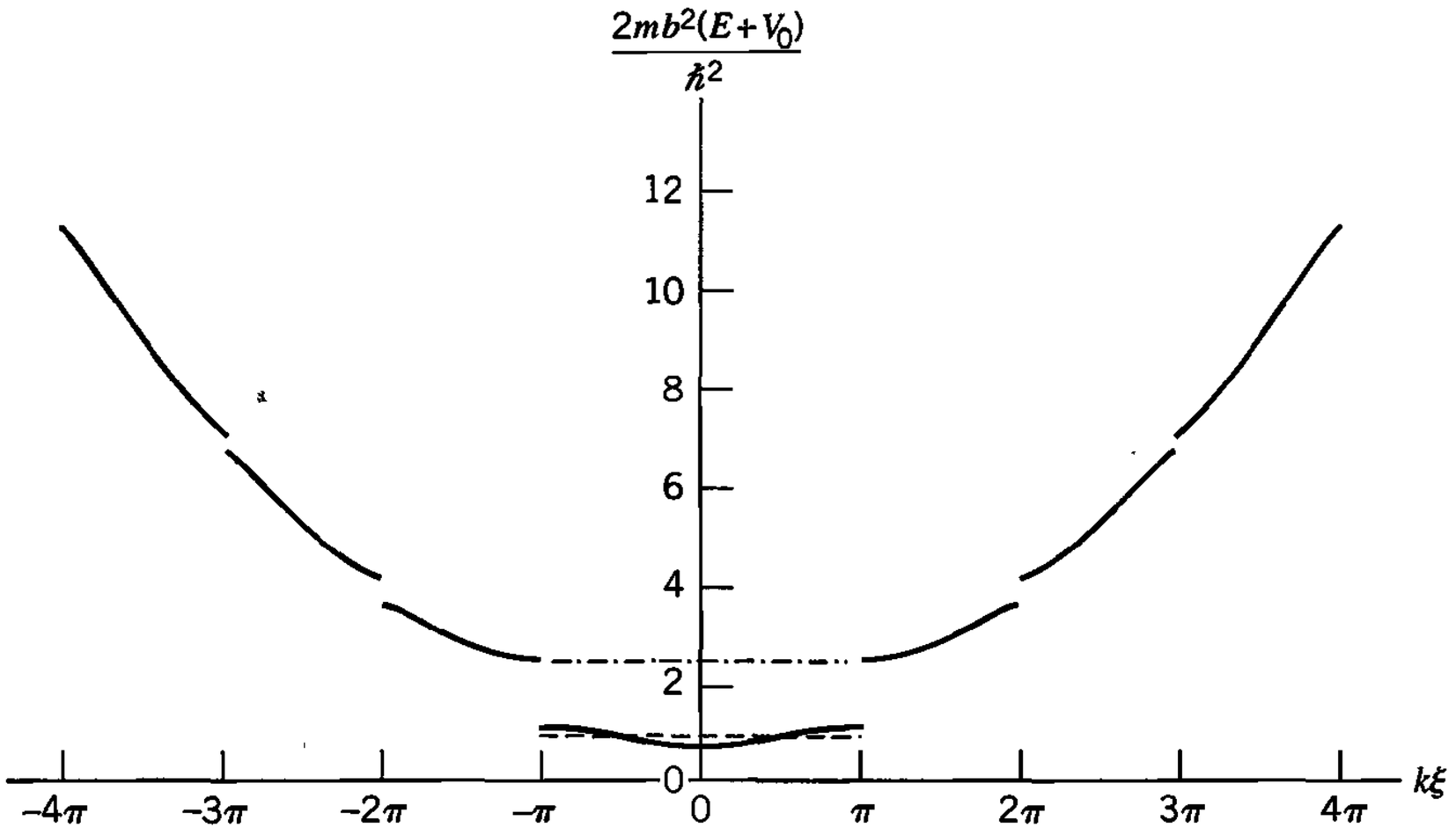


Figure 8.12. The extended-zone scheme for the same conditions as in Figure 8.11.

monotonicity, as well as physical intuition, guide the proper identification of E as a function of k. In the limit $E \to +\infty$, the perturbation method applies, k then approaches k' and becomes the wave number of the nearly-free particle. The periodicity of the energy in the reciprocal lattice space is emphasized in the *repeated-zone scheme*, which is shown in Figure 8.13 for our Kronig-Penney model. The smooth behavior at the zone boundaries is a consequence of the symmetry relation (8.124).

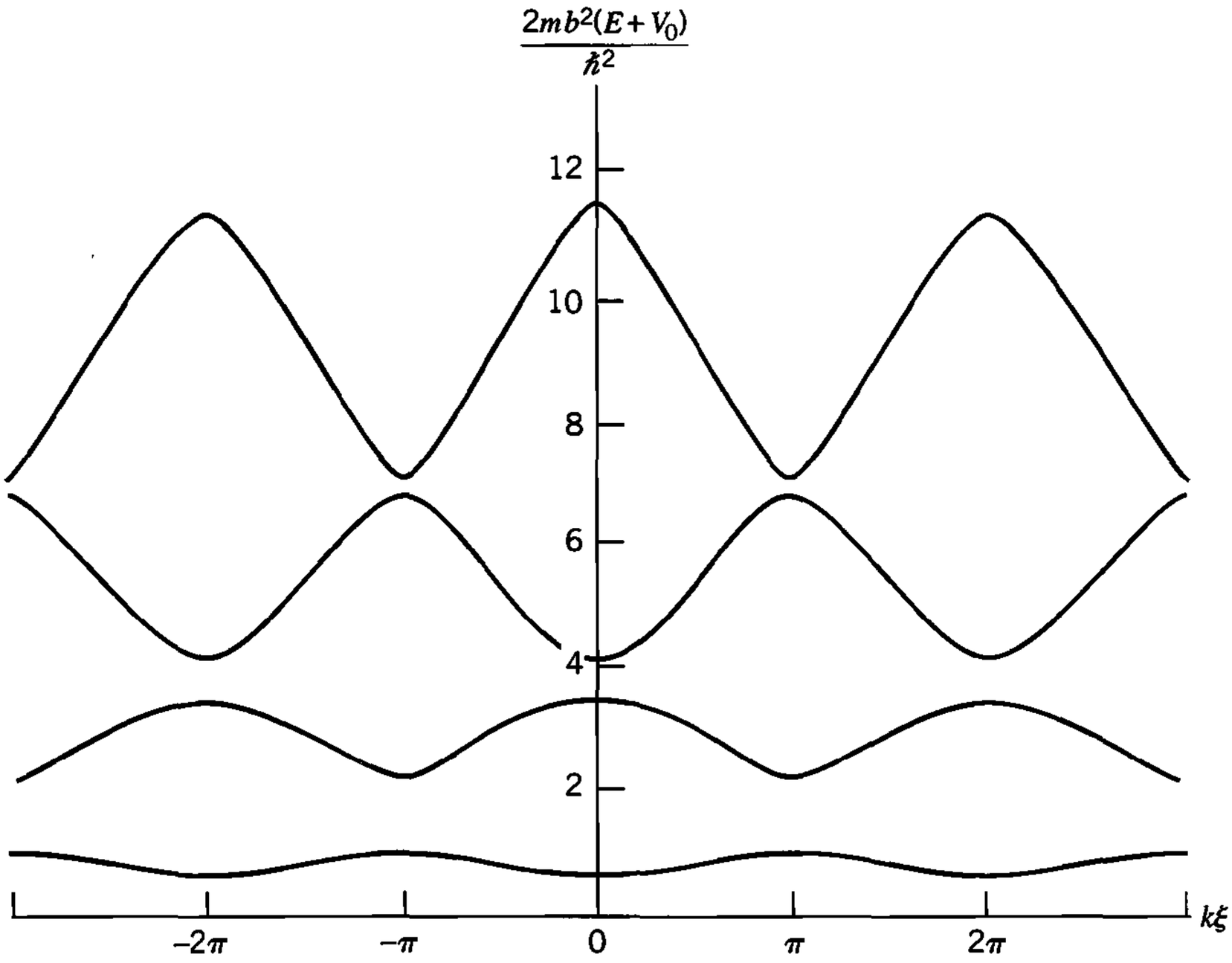


Figure 8.13. The repeated-zone scheme for the same conditions as in Figure 8.11.

Exercise 8.28. Using the results of Exercise 8.27 for the Kronig-Penney potential of Figure 8.9, prove in the limit of high energies that $k\xi = 2k''a + 2k'b$, and compare the numerical consequences of this relation with the exact dispersion curve, Figure 8.11. In this regime, compute the width of the forbidden gaps in the perturbation approximation as a function of N.

Exercise 8.29. If in the periodic Kronig-Penney potential of Figure 8.9 the square wells become delta functions in the limit $b \to 0$ and $V_0 \to \infty$ such that $V_0 b$ remains finite, show that the eigenvalue condition reduces to

$$\cos k\xi = \cosh \kappa\xi - \frac{\kappa_0}{\kappa} \sinh \kappa\xi$$

for $E < 0$, and

$$\cos k\xi = \cos k''\xi - \frac{\kappa_0}{k''} \sin k''\xi$$

for $E > 0$. Here $\hbar\kappa_0 = \dfrac{2m \lim(V_0 b)}{\hbar} = \sqrt{-2mE_0}$, with E_0 being the binding energy of the bound state in the delta-function well (see Section 6.4). Discuss the occurrence of allowed bands and forbidden energy gaps. Check the prediction for the gap width in the weak perturbation limit. Show that as $\xi \to \infty$, the allowed band for $E < 0$ degenerates into the discrete energy level E_0.

We must now look briefly at the eigenfunctions of the Schrödinger equation.[6] Inside the wells they are of the form (8.132). The coefficients of the plane waves for two fundamental solutions corresponding to the same energy are given by

$$\begin{pmatrix} A_n^{(+)} \\ B_n^{(+)} \end{pmatrix} = e^{ink\xi} \begin{pmatrix} A_0^{(+)} \\ B_0^{(+)} \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} A_n^{(-)} \\ B_n^{(-)} \end{pmatrix} = e^{-ink\xi} \begin{pmatrix} A_0^{(-)} \\ B_0^{(-)} \end{pmatrix}$$

Because of invariance under time reversal, we may assume the relation

$$\frac{A_0^{(-)}}{B_0^{(-)}} = \left(\frac{B_0^{(+)}}{A_0^{(+)}} \right)^*$$

between the two solutions. From the eigenvalue equation for the matrix P, we obtain the ratio

$$\frac{A_0^{(+)}}{B_0^{(+)}} = \frac{\beta_2 e^{ik'\xi}}{\alpha_1 \sin k'\xi - \beta_1 \cos k'\xi - \sin k\xi}$$

which can be used to construct the eigenfunctions inside the square wells:

$$\begin{aligned} \psi^{(+)}(x) &= [\psi^{(-)}(x)]^* \\ &= e^{ink\xi}\{\beta_2 e^{ik'[x-(n-1/2)\xi]} + (\alpha_1 \sin k'\xi - \beta_1 \cos k'\xi - \sin k\xi) e^{-ik'[x-(n-1/2)\xi]}\} \\ &= e^{ikx} e^{-ik(x-n\xi)}\{\cdots\} = e^{ikx} u_{k,E}(x) \end{aligned} \tag{8.145}$$

for $a + (n-1)\xi < x \leq n\xi - a$. In the plateaus where $V = 0$, the harmonic waves are, for $E < 0$, replaced by increasing and decreasing exponentials, the coefficients

[6] A more detailed discussion of quantum states in simple periodic lattices is found in Liboff (1992).

being determined by smoothly joining the eigenfunction at the discontinuities of the potential. The function $u_{k,E}(x)$ defined by Eq. (8.145) has the periodicity property

$$u_{k,E}(x + \xi) = u_{k,E}(x)$$

and the eigenfunctions (8.145) are Bloch functions, as anticipated.

Since the idealized periodic lattice is infinite and the emphasis has been on the energy spectrum, it has not been necessary to specify the boundary and normalization conditions for the Bloch-type energy eigenfunctions $\psi_E(x) = e^{ikx}u_{k,E}(x)$ of the Schrödinger equation. As for a free particle, periodic boundary conditions may be imposed in a domain $L = N\xi$, which is a large multiple of the lattice spacing. The permissible values of k are discrete for finite N but continuous in the limit $N \to \infty$.

Unlike perfect plane waves, Bloch functions are generally not momentum eigenfunctions, but it is of physical interest to evaluate the expectation value of the velocity in Bloch states. Since

$$p_x\psi_k(x) = \frac{\hbar}{i}\frac{\partial}{\partial x}e^{ikx}u_k(x) = e^{ikx}\left[\hbar k + \frac{\hbar}{i}\frac{\partial}{\partial x}\right]u_k(x)$$

we have

$$\langle p_x\rangle = m\langle v_x\rangle = C\int \psi_k^*(x)p_x\psi_k(x)\,dx = C\int u_k^*(x)\left[\hbar k + \frac{\hbar}{i}\frac{d}{dx}\right]u_k(x)\,dx \quad (8.146)$$

where C is a normalization constant. With the definition (8.122) for the Hamiltonian operator H_k, the last expression can be written as

$$\langle v_x\rangle = \frac{C}{\hbar}\int u_k^*(x)\frac{\partial H_k}{\partial k}u_k(x)\,dx = \frac{1}{\hbar}\left\langle\frac{\partial H_k}{\partial k}\right\rangle \quad (8.147)$$

With boundary conditions that ensure that H_k is Hermitian, we finally deduce the simple result:

$$\boxed{\langle v_x\rangle = \frac{1}{\hbar}\frac{d}{dk}\langle H_k\rangle = \frac{1}{\hbar}\frac{dE(k)}{dk}} \quad (8.148)$$

This relation shows that $\langle v_x\rangle$ is the group velocity of a wave packet that is a Bloch wave modulated by a broad real-valued amplitude function.

Exercise 8.30. Prove the equality

$$\frac{dE(k)}{dk} = \left\langle\frac{\partial H_k}{\partial k}\right\rangle \quad (8.149)$$

which is a special case of the Hellmann-Feynman theorem (Exercise 18.14). Calculate $\langle v_x\rangle$ for the dispersion function $E(k)$ in the tight-binding approximation, (8.129).

Exercise 8.31. Following the derivation and approximation in Section 2.3 for free particles, show that in a periodic lattice a wave packet that initially has the form of an amplitude-modulated Bloch function,

$$\psi(x, 0) = A(x)e^{ik_0x}u_{k_0}(x)$$

with a broad bell-shaped real-valued amplitude $A(x)$, propagates with constant velocity (8.148) and without change of shape, except for a phase factor.

The two solutions $\psi^{(+)}(x)$ and $\psi^{(-)}(x)$ defined in (8.145) are linearly independent and describe uniformly propagating waves, unless $k\xi = N\pi$ (N integer), in which case the two solutions become identical. Since it can be shown that for these special values of k,

$$\alpha_1 \sin k'\xi - \beta_1 \cos k'\xi = \pm\beta_2 \tag{8.150}$$

the corresponding eigenfunctions (8.145) represent standing waves, interpretable as resulting from *Bragg reflection*. These solutions correspond to the *band edges*. In the nearly-free particle approximation, the condition for the band edges coincides with the equation locating the small forbidden gaps.

Exercise 8.32. Verify (8.150) for $k\xi = N\pi$. Show that at the band edges, the group velocity is

$$\langle v_k \rangle = \frac{dE(k)}{dk} = 0 \tag{8.151}$$

In this chapter, as in the preceding ones, we have formulated quantum mechanics for systems that are best visualized in some kind of (mostly one-dimensional) coordinate space or the complementary momentum space. At almost every step, it has been evident that the theory can be greatly generalized if it is expressed in terms of the mathematical apparatus of complex vector spaces and their matrix representations. We now turn to this general formulation of quantum mechanics.

Problems

1. Apply the variational method to estimate the ground state energy of a particle confined in a one-dimensional box for which $V = 0$ for $-a < x < a$, and $\psi(\pm a) = 0$.

   (a) First, use an unnormalized trapezoidal trial function which vanishes at $\pm a$ and is symmetric with respect to the center of the well:

$$\psi_t(x) = \begin{cases} (a - |x|) & b \le |x| \le a \\ a - b & |x| \le b \end{cases}$$

   Try the choice $b = 0$ (triangular trial function) and then improve on this by optimizing the parameter b.

   (b) A more sophisticated trial function is parabolic, again vanishing at the endpoints and even in x.

   (c) Use a quartic trial function of the form

$$\psi_t(x) = (a^2 - x^2)(\alpha x^2 + \beta)$$

   where the ratio of the adjustable parameters α and β is determined variationally.

   (d) Compare the results of the different variational calculations with the exact ground state energy, and, using normalized wave functions, evaluate the mean-square deviation $\int_{-a}^{a} |\psi(x) - \psi_t(x)|^2 \, dx$ for the various cases.

   (e) Show that the variational procedure produces, in addition to the approximation to the ground state, an optimal quartic trial function with nodes between the endpoints. Interpret the corresponding stationary energy value.[7]

[7]Adapted from Cohen-Tamoudji, Diu, and Laloë (1977), Chapter 11.

2. Using scaled variables, as in Section 5.1, consider the anharmonic oscillator Hamiltonian,

$$H = \frac{1}{2} p_\xi^2 + \frac{1}{2} \xi^2 + \lambda \xi^4$$

where λ is a real-valued parameter.

(a) Estimate the ground state energy by a variational calculation, using as a trial function the ground state wave function for the harmonic oscillator

$$H_0(\omega) = \frac{1}{2} p_\xi^2 + \frac{1}{2} \omega^2 \xi^2$$

where ω is an adjustable variational parameter. Derive an equation that relates ω and λ.

(b) Compute the variational estimate of the ground state energy of H for various positive values of the strength λ.

(c) Note that the method yields answers for a discrete energy eigenstate even if λ is slightly negative. Draw the potential energy curve to judge if this result makes physical sense. Explain.

3. In first-order perturbation theory, calculate the change in the energy levels of a linear harmonic oscillator that is perturbed by a potential gx^4. For small values of the coefficient, compare the result with the variational calculation in Problem 2.

4. Using a Gaussian trial function, $e^{-\lambda x^2}$, with an adjustable parameter, make a variational estimate of the ground state energy for a particle in a Gaussian potential well, represented by the Hamiltonian

$$H = \frac{p^2}{2m} - V_0 e^{-\alpha x^2} \qquad (V_0 \text{ and } \alpha > 0)$$

5. Show that as inadequate a variational trial function as

$$\psi(x) = \begin{cases} C\left(1 - \dfrac{|x|}{a}\right) & \text{for } |x| \leq a \\ 0 & \text{for } |x| > a \end{cases}$$

yields, for the optimum value of a, an upper limit to the ground state energy of the linear harmonic oscillator, which lies within less than 10 percent of the exact value.

6. A particle of mass m moves in a potential $V(\mathbf{r})$. The n-th discrete energy eigenfunction of this system, $\psi_n(\mathbf{r})$, corresponds to the energy eigenvalue E_n. Apply the variational principle by using as a trial function,

$$\psi_t(\mathbf{r}) = \psi_n(\lambda \mathbf{r})$$

where λ is a variational (scaling) parameter, and derive the virial theorem for stationary states.

7. In Chapter 6 it was shown that every one-dimensional square well supports at least one bound state. By use of the variational principle, prove that the same is true for any one-dimensional potential that is negative for all values of x and that behaves as $V \to 0$ as $x \to \pm\infty$.

8. Work out an approximation to the energy splitting between the second and third excited levels of the double oscillator defined in Section 8.5, assuming the distance between the wells to be very large compared with the classical amplitude of the zero-point vibrations.

9. Solve the energy eigenvalue problem for a particle that is confined in a two-dimensional square box whose sides have length L and are oriented along the x- and y-coordinate axes with one corner at the origin. Find the eigenvalues and eigenfunctions, and calculate the number of eigenstates per unit energy interval for high energies.

A small perturbation $V = Cxy$ is now introduced. Find the approximate energy change of the ground state and the splitting of the first excited energy level. For the given perturbation, construct the optimal superpositions of the unperturbed wave functions in the case of the first excited state.

10. As an example of Problem 2 in Chapter 7, apply the WKB approximation to the double harmonic oscillator of Section 8.5, and contrast the energy level splitting of the two lowest levels with the results obtained in Section 8.5.

11. The energy $E_0(a)$ of the lowest eigenstate of a double harmonic oscillator with fixed ω depends on the distance a and has a minimum at $a = a_0$ (see Figure 8.5). Adapt the Hellmann-Feynman theorem for the expectation value of a parameter-dependent Hamiltonian (Exercise 8.30) to this problem, and show that if $a = a_0$, the expectation value of $|x|$ is equal to a_0.

12. Apply the WKB approximation to a periodic potential in one dimension, and derive an implicit equation for the dispersion function $E(k)$. Estimate the width of the valence band of allowed bound energy levels.

13. Assume that n unperturbed, but not necessarily degenerate, eigenstates $|k\rangle$ of an unperturbed Hamiltonian H_0 (with $k = 1, 2, \ldots, n$) all interact with one of them, say $|1\rangle$, but not otherwise so that the perturbation matrix elements $\langle k|V|k'\rangle \neq 0$ only if either $k = k'$ or $k = 1$ or $k' = 1$. Solve the eigenvalue problem in the n-dimensional vector space exactly and derive an implicit equation for the perturbed energies. Using a graphic method, discuss the solutions of the eigenvalue problem for various assumed values of the nonvanishing matrix elements of V, and exhibit the nature of the perturbed eigenstates.

CHAPTER 9

Vector Spaces in Quantum Mechanics

Generalizing the concepts of wave mechanics, in this chapter we begin to formulate the principles of quantum mechanics in terms of the mathematical structure of an abstract vector space (separable Hilbert space). Here we review and assemble in one place the tools of the state vector and operator formalism of quantum mechanics, including the commonly used notational devices. At the end, we close the circle by showing how ordinary wave mechanics reemerges from the general theory as a special representation.

1. ***Probability Amplitudes and Their Composition.*** The principles of wave mechanics were the subject of Chapter 4, and in Chapters 5 through 8 the time-independent Schrödinger equation was applied to specific simple dynamical systems, mostly in one dimension. These illustrations have provided us with several examples of Hamiltonians, or energy operators, whose admissible eigenfunctions constitute complete orthonormal sets for the expansion of an arbitrary wave function (or state) $\psi(x)$ of the one-dimensional position coordinate, x. Following Section 4.2, we summarize the standard procedure: The eigenvalue equation for the Hermitian operator A is

$$A\psi_n(x) = A'_n\psi_n(x) \tag{9.1}$$

where A'_n denotes the nth eigenvalue. The eigenfunctions satisfy the orthonormality condition

$$\int \psi_n^*(x)\psi_k(x)\,dx = \delta_{nk} \tag{9.2}$$

An arbitrary state $\psi(x)$ is assumed to be expressible as the series,

$$\psi(x) = \sum_n c_n\psi_n(x) \tag{9.3}$$

with the expansion coefficients given by

$$c_n = \int_{-\infty}^{+\infty} \psi_n^*(x)\psi(x)\,dx \tag{9.4}$$

The coefficients c_n are the probability amplitudes. If the eigenvalue A'_n is simple, measurements of A give the result A'_n with probability $|c_n|^2$. If $\psi(x)$ happens to be equal to an eigenfunction $\psi_k(x)$, the measurement of A will yield A'_k with certainty and confirm that $\psi(x)=\psi_k(x)$. Generally, $\psi(x)$ is a superposition of many eigenstates.

If an eigenvalue of A is not simple but corresponds to more than one linearly independent eigenfunction, additional labels are required to characterize the eigenfunctions, and measurements of A alone do not unambiguously identify all the individual terms in the expansion. In this very common case, the determination of which $\psi_n(x)$ represents the system can be further narrowed down if the eigenstates

$\psi_n(x)$ of A are also eigenstates of a second Hermitian operator B, with eigenvalues B'_n, such that eigenstates which correspond to the *same* eigenvalue of A are distinguished by *different* eigenvalues of B, to which they correspond. Operators A and B are said to be *compatible*. If after the introduction of B the simultaneous measurements of A and B still do not unambiguously specify all the eigenstates, the procedure is continued until a *complete* set of compatible Hermitian operators is found. Since they share a complete set of common eigenfunctions, the physical observables that these operators represent are simultaneously assignable to the system, or *simultaneously measurable*. As we will show in the next chapter, the necessary and sufficient mathematical condition for two observables A and B to be compatible, or to possess a common complete set of eigenstates, is that the operators A and B commute. For the present, we symbolize a complete set of operators by a single operator K, with simple eigenvalues K_i. In order to simplify the notation, we omit the prime on the discrete eigenvalues of a generic operator like K and rely on the subscript alone to identify K_i as denoting a set of eigenvalues.

With the experience of ordinary Schrödinger wave mechanics as our guide, we now endeavor to build up a general formalism of quantum mechanics for any dynamical system. The fundamental assumption of quantum mechanics is as follows:

> *The maximum information about the outcome of physical measurements on a system is contained in the probability amplitudes that belong to a complete set of observables for the system.*

If the state is denoted by Ψ, the probability *amplitude* for finding K_i in a measurement of K is a complex number, usually denoted by $\langle K_i|\Psi\rangle$. The *probability* of finding K_i when K is measured is given by $|\langle K_i|\Psi\rangle|^2$. Applied to the wave-mechanical formalism encapsulated in Eqs. (9.1)–(9.4), the probability amplitude c_n can be expressed as $c_n = \langle A'_n|\Psi\rangle$.

We stress again that the quantity K may stand for several compatible observables; the nature of the physical system determines how many variables K and its eigenvalues K_i represent and what they are. For example, a structureless particle moving in three dimensions is conveniently described by the three commuting observables, x, y, z, measuring the particle position. The corresponding probability amplitude $\langle x, y, z|\Psi\rangle$ is nothing other than the coordinate state function or wave function $\psi(x, y, z)$ of the particle. An alternative set of commuting observables is p_x, p_y, p_z, and the probability amplitude $\langle p_x, p_y, p_z|\Psi\rangle$ is the momentum wave function $\phi(p_x, p_y, p_z)$. A very different complete set of commuting observables for the same system consists of the three linear harmonic oscillator Hamiltonians,

$$H_j = \frac{p_j^2}{2m_j} + \frac{1}{2} m_j \omega_j^2 x_j^2 \qquad (j = 1, 2, 3)$$

with eigenvalues

$$E_j = \hbar\omega_j\left(n_j + \frac{1}{2}\right)$$

where the quantum numbers n_j are nonnegative integers. (We use the notations x, y, z and x_1, x_2, x_3 interchangeably as dictated by convenience.) The probability amplitudes, which completely specify the state Ψ, are $\langle n_1\ n_2\ n_3|\Psi\rangle$. Thus, the state may be represented equivalently by an *indenumerable continuum* of amplitudes, $\langle x, y, z|\Psi\rangle = \psi(x, y, z)$, or $\langle p_x, p_y, p_z|\Psi\rangle = \phi(p_x, p_y, p_z)$, or by a denumerably

infinite set of discrete amplitudes like $\langle n_1\ n_2\ n_3|\Psi\rangle$. In general, an amplitude may depend on both discrete and continuous variables, and some useful physical quantities, such as the energy of an atom, have eigenvalue spectra that are part discrete and part continuous.

It is convenient to be able to encompass both kinds of spectra, discrete and continuous, in a unified notation, and there are several ways of doing this. The Stieltjes integral is a mathematical tool and a generalization of the Riemann integral, allowing discrete sums to be regarded as special cases of integrals.[1] Delta functions may be employed as densities in integrations to represent discrete sums as integrals. Conversely, one can think of the continuous spectrum of an observable as the limiting case of a completely discrete spectrum of a slightly different observable. The difference between the two usually arises from a modification of the boundary conditions or from the neglect of a small "interaction" term in the operator.

In this chapter, most equations are written in a manner appropriate for *discrete* eigenvalues of the observables, but it is always easy to replace sums by integrals and Kronecker deltas by delta functions, when the spectrum is continuous (see Section 9.6).

To keep the notation uncluttered, as long as no confusion is likely to occur, operators will be denoted simply as A, B, K, etc., just like the physical quantities they represent; their measured numerical values or eigenvalues will be denoted as A'_i, B'_j, etc. (*but* K_i instead of K'_i), whenever the need to distinguish them from the operators calls for caution.

In the remainder of this section, we will infer the mathematical properties of the amplitudes from the probability interpretation and the laws of quantum physics. In every instance, these mathematical properties are satisfied by the probability amplitudes (9.4) of one-dimensional quantum mechanics (wave mechanics).

First, since the values K_i constitute all the possible results of measuring the observable K, the probabilities must add up to unity:

$$\boxed{\sum_i |\langle K_i|\Psi\rangle|^2 = 1} \tag{9.5}$$

expressing the normalization of the probability amplitudes. It is customary to denote by $\langle A'_i|B'_j\rangle$ the probability amplitude for finding A'_i upon measuring A on a system in a state for which a measurement of B yields the value B'_j with certainty. In particular, since the different measured values of K are supposed to be unambiguously distinguishable and mutually exclusive, we must require

$$\langle K_i|K_j\rangle = 0 \quad \text{if } i \neq j \tag{9.6}$$

The amplitude $\langle K_i|K_i\rangle$ must have modulus unity, and it is therefore possible to make the choice

$$\langle K_i|K_j\rangle = \delta_{ij} \tag{9.7}$$

An arbitrary state Ψ can be equivalently specified by the probability amplitudes for any complete set of observables, K, L, and so on. The formula connecting the amplitudes $\langle K_i|\Psi\rangle$ and $\langle L_j|\Psi\rangle$ is patterned after the expansion postulate, by which states describing different outcomes of a measurement are superimposed, and represents the distillation of the accumulated empirical evidence underlying quantum

[1]Ballentine (1990), p. 12.

mechanics. The link between the probability amplitudes for two representations is established by the *composition rule* for probability amplitudes,

$$\boxed{\langle L_j|\Psi\rangle = \sum_i \langle L_j|K_i\rangle\langle K_i|\Psi\rangle} \tag{9.8}$$

Exercise 9.1. Show the consistency of (9.8) with condition (9.7).

The simple relation (9.8) embraces many of the surprising and peculiar features that set quantum mechanics apart from a classical description. If L is measured on a system in state Ψ, the probability of finding L_j is

$$|\langle L_j|\Psi\rangle|^2 = \left|\sum_i \langle L_j|K_i\rangle\langle K_i|\Psi\rangle\right|^2 \tag{9.9}$$

which differs from the conventional rule for calculating probabilities:

$$\sum_i |\langle L_j|K_i\rangle|^2 \, |\langle K_i|\Psi\rangle|^2 \tag{9.10}$$

by the presence of *interference terms*. Expression (9.10) is not wrong, but it is the probability of finding L_j if an actual measurement of K is first carried out, leaving the system in one of the states in which a remeasurement of K would yield K_i with certainty. Generally, such an interim measurement of K alters the initial state dramatically and irreversibly, so that (9.10) does not represent the probability of directly finding L_j for a system that is in state Ψ.

If the composition rule (9.8) is applied to state $|\Psi\rangle = |L_n\rangle$, we get

$$\sum_i \langle L_j|K_i\rangle\langle K_i|L_n\rangle = \langle L_j|L_n\rangle = \delta_{jn} \tag{9.11}$$

For fixed j, these equations permit the evaluation of the amplitudes $\langle L_j|K_i\rangle$ if the amplitudes $\langle K_i|L_n\rangle$ are given. The solution can be inferred by comparing the set of linear equations (9.11) with the normalization condition

$$\sum_i |\langle K_i|L_j\rangle|^2 = \sum_i \langle K_i|L_j\rangle^*\langle K_i|L_j\rangle = 1$$

Putting $n = j$ in (9.11), we are led to conclude that

$$\boxed{\langle L_j|K_i\rangle = \langle K_i|L_j\rangle^*} \tag{9.12}$$

and, therefore, that there is an important *reciprocal* relationship:

$$\boxed{|\langle K_i|L_j\rangle|^2 = |\langle L_j|K_i\rangle|^2} \tag{9.13}$$

According to (9.13), the *conditional probability* of finding K_i in a state that is known to ''have'' the sharp value L_j for L is equal to the probability of finding L_j in a state that is known to have a sharp value K_i for K. Equation (9.13) is in accord with all the experimental evidence.

If conditions (9.11) and (9.12) are combined, the probability amplitudes $\langle L_j|K_i\rangle$ are seen to be subject to the conditions

$$\sum_i \langle L_j|K_i\rangle\langle L_n|K_i\rangle^* = \delta_{jn} \tag{9.14}$$

expressing the unitary character of the transformation matrix S, whose elements are defined as

$$S_{ji} = \langle L_j | K_i \rangle \tag{9.15}$$

The unitary property of S implies that

$$SS^\dagger = 1 \tag{9.16}$$

Exercise 9.2. Show that Eqs. (9.5) through (9.16) are valid if they are translated into the quantum mechanics of a particle with one position coordinate (wave mechanics in one dimension). Show that (9.14) expresses the orthonormality of eigenfunctions of observables as well as the closure property.

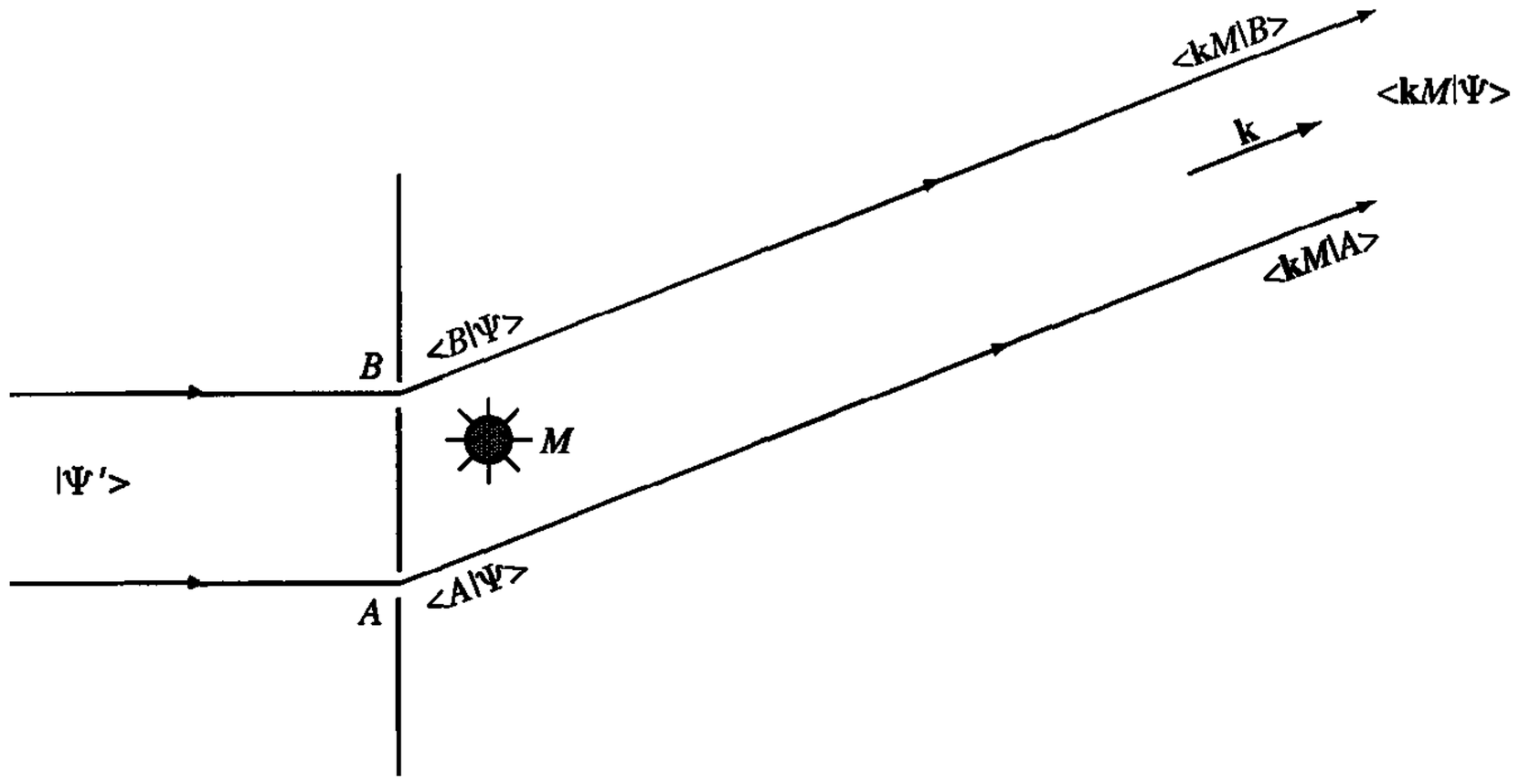
The two-slit interference experiment (Figure 1.1), interpreted in terms of a statistical distribution of particles on the detecting screen, can be used to illustrate the rules for quantum mechanical amplitudes established in this section. In a typical interference setup, the two amplitudes $\langle A|\Psi\rangle$ and $\langle B|\Psi\rangle$ for particles to appear in the separate slits or channels, A and B, completely determine the state Ψ. The same state can be described in terms of the amplitudes for the outcome of possible measurements in the region beyond the screen with the slits. If the position $\mathbf{r}$ is chosen as the observable for this description, the amplitude for finding the particle at $\mathbf{r}$ is related to the amplitudes $\langle A|\Psi\rangle$ and $\langle B|\Psi\rangle$ by the composition law:

$$\langle \mathbf{r}|\Psi\rangle = \langle \mathbf{r}|A\rangle\langle A|\Psi\rangle + \langle \mathbf{r}|B\rangle\langle B|\Psi\rangle \tag{9.17}$$

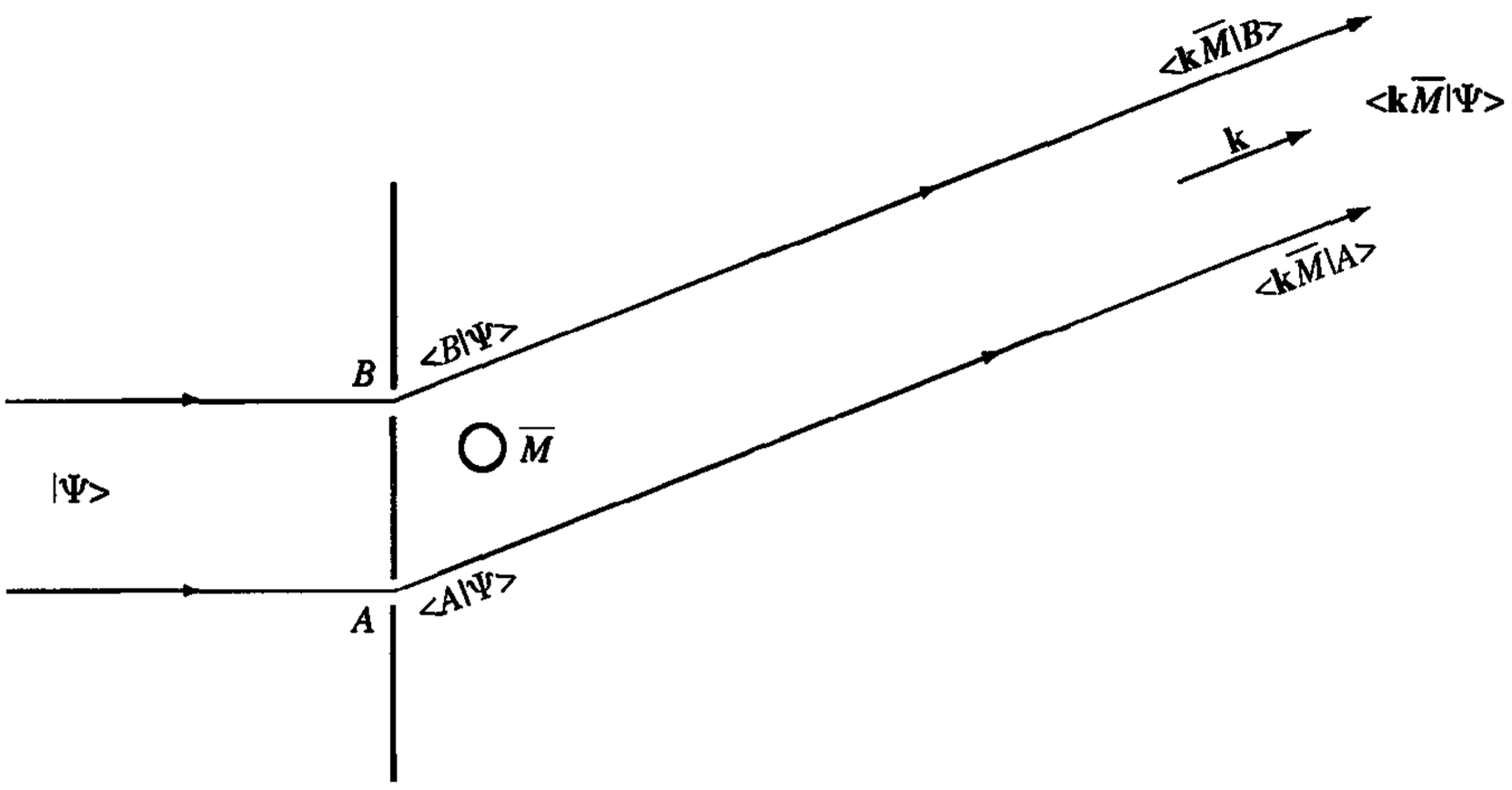
Interference phenomena arise from calculating the probability $|\langle \mathbf{r}|\Psi\rangle|^2$ with this formula.

Exercise 9.3. In a two-slit interference experiment with particles of definite wave number (energy) k, the slits A and B are located at positions $\mathbf{r}_A$ and $\mathbf{r}_B$. At large distance from the slits the amplitudes $\langle \mathbf{r}|A\rangle$ and $\langle \mathbf{r}|B\rangle$ are to reasonable accuracy represented by $\langle \mathbf{r}|A\rangle \propto e^{ik|\mathbf{r}_A-\mathbf{r}|}$ and $\langle \mathbf{r}|B\rangle \propto e^{ik|\mathbf{r}_B-\mathbf{r}|}$. Show how, to within a constant of proportionality, the probability of finding the particle at position $\mathbf{r}$ depends on the difference of the distances from A and B to the point of observation, and on the relative magnitude and phase of the amplitudes $\langle A|\Psi\rangle$ and $\langle B|\Psi\rangle$, which are determined by the experimental arrangement.

Two-slit interference is the paradigm for countless (real and thought) experiments in quantum physics. We extend the discussion to include the effect of an additional monitor that receives a signal alerting it to the passage of the particle through one of the slits. We characterize the point of observation of the particle (either at great distance r from the slits or on the focal plane of a converging lens) by the direction $\mathbf{k}$ of the diffracted wave (Figure 9.1). Under these conditions, we have an amplitude, $\langle \mathbf{k}M|A\rangle$, corresponding to the probability of finding the particle that is known to pass through A, subsequently to be diffracted in direction $\mathbf{k}$ *and* to cause the detector to fire. A second amplitude, $\langle \mathbf{k}\overline{M}|A\rangle$, corresponds to the complementary outcome that the diffracted particle is *not* accompanied by a "hit" in the detector. There are analogous amplitudes $\langle \mathbf{k}M|B\rangle$ and $\langle \mathbf{k}\overline{M}|B\rangle$ for passage of the particle through slit B.



(*a*)



(*b*)

Figure 9.1. Two-slit interference in the presence of a monitor that has the capacity to respond to the passage of a particle diffracted in direction **k** through slits A or B. In scenario (*a*) the monitor registers the passage with probability amplitudes $\langle \mathbf{k}M|A\rangle$ and $\langle \mathbf{k}M|B\rangle$, respectively. In scenario (*b*) the amplitudes $\langle \mathbf{k}\overline{M}|A\rangle$ and $\langle \mathbf{k}\overline{M}|B\rangle$ correspond to the monitor remaining in its latent initial state, while the particle is diffracted. The off-axis placement of the monitor is meant to suggest that it may be more sensitive to particles passing through one slit than the other.

Given the state Ψ, the amplitude for finding the particle to have been diffracted in direction **k** *and* the detector to have fired is, by the composition rule:

$$\langle \mathbf{k}M|\Psi\rangle = \langle \mathbf{k}M|A\rangle\langle A|\Psi\rangle + \langle \mathbf{k}M|B\rangle\langle B|\Psi\rangle \tag{9.18}$$

Similarly, the amplitude for diffraction in direction **k** and *no* response from the detector is

$$\langle \mathbf{k}\overline{M}|\Psi\rangle = \langle \mathbf{k}\overline{M}|A\rangle\langle A|\Psi\rangle + \langle \mathbf{k}\overline{M}|B\rangle\langle B|\Psi\rangle \tag{9.19}$$

Generally, the probabilities $|\langle \mathbf{k}M|\Psi\rangle|^2$ and $|\langle \mathbf{k}\overline{M}|\Psi\rangle|^2$, calculated from (9.18) and (9.19), exhibit interference.

If, by virtue of its placement near one of the slits or by some other technique, the monitor responds with high efficiency only to passage of the particle through one of the slits, say A, we have $\langle \mathbf{k}M|B\rangle \approx \langle \mathbf{k}\overline{M}|A\rangle \approx 0$. We say that A is *correlated* with M, and B with $\overline{M}$. The probability of observing the particle in direction $\mathbf{k}$ is then

$$|\langle \mathbf{k}M|\Psi\rangle|^2 + |\langle \mathbf{k}\overline{M}|\Psi\rangle|^2 \approx |\langle \mathbf{k}M|A\rangle\langle A|\Psi\rangle|^2 + |\langle \mathbf{k}\overline{M}|B\rangle\langle B|\Psi\rangle|^2 \tag{9.20}$$

without an interference term. If, on the other hand, the detector monitors the passage of all particles through either of the two slits perfectly, but without the capacity to distinguish between the originating slit (perhaps because it is located halfway between A and B), the probability of observing the particle in the direction $\mathbf{k}$ is

$$|\langle \mathbf{k}M|\Psi\rangle|^2 \approx |\langle \mathbf{k}M|A\rangle\langle A|\Psi\rangle + \langle \mathbf{k}M|B\rangle\langle B|\Psi\rangle|^2 \tag{9.21}$$

This expression, like (9.17), leads to a typical two-slit interference pattern. In general, with a monitoring device M in place, the probability of observing a particle in direction $\mathbf{k}$ is

$$\begin{aligned} |\langle \mathbf{k}M|\Psi\rangle|^2 + |\langle \mathbf{k}\overline{M}|\Psi\rangle|^2 = {} & |\langle \mathbf{k}M|A\rangle\langle A|\Psi\rangle + \langle \mathbf{k}M|B\rangle\langle B|\Psi\rangle|^2 \\ & + |\langle \mathbf{k}\overline{M}|A\rangle\langle A|\Psi\rangle + \langle \mathbf{k}\overline{M}|B\rangle\langle B|\Psi\rangle|^2 \end{aligned} \tag{9.22}$$

We see that, for particles diffracted along $\mathbf{k}$, the intensity exhibits interference effects between the amplitudes pertaining to the two slits, unless we can *unambiguously distinguish* between the two distinct paths. We emphasize that for interference to be eradicated, the monitor M, which is correlated to the passage of the particle through the slits, does not need to produce recorded signals in the detector. It may indeed be idle, but in principle it must be able to distinguish path A from path B.[2]

In Section 9.4 we will see how the physically motivated rules for probability amplitudes can be fitted into a general mathematical framework. We will consider the probability amplitudes to be *components of a vector* Ψ in a space whose "coordinate axes" are labeled by the different values K_i. We will refer to Ψ as a *state vector*, or briefly a *state* in the abstract vector space.

To prepare for the vector space formulation of quantum mechanics, we review the fundamentals of the algebra of vector spaces and operators in the next two sections. Two different but equivalent notations will be used in this chapter, because the Dirac bra-ket notation alone is not sufficiently flexible to meet all needs.

In general, the states that have so far been considered require infinitely many amplitudes for their definition, and indeed many applications require the use of infinite-dimensional vector spaces. There are, however, important physical systems or degrees of freedom of systems, such as the spin of a particle, for which a finite number of amplitudes is appropriate. In order to keep the discussion simple, most of the concepts in this chapter will be developed for complex vector spaces of a *finite* number of dimensions, n. The infinite-dimensional vector spaces that are important in quantum mechanics are analogous to finite-dimensional vector spaces and can be spanned by a countable basis. They are called *separable Hilbert spaces*. We will be interested only in those properties and theorems for $n \to \infty$ which are straightforward generalizations of the finite-dimensional theory. If, for these gen-

[2]This discussion is similar to, but more general than, that in Feynman (1965), vol. III, Chapter 3. For descriptions of many fundamental experiments relevant to understanding quantum mechanics, see Lévy-Leblond and Balibar (1990).

eralizations to hold, the vectors and operators of the space need to be subjected to certain additional restrictive conditions, we will suppose that these conditions are enforced. By confining ourselves to the complex vector space in n dimensions, we avoid questions that concern the convergence of sums over infinitely may terms, the interchangeability of several such summations, or the legitimacy of certain limiting processes. Nevertheless, the mathematical conclusions that we draw for the infinitely dimensional space by analogy with n-dimensional space can be rigorously justified.[3]

2. ***Vectors and Inner Products.*** Our abstract vector space is defined as a collection of vectors, denoted by Ψ, any two of which, say Ψ_a and Ψ_b, can be combined to define a new vector, denoted as the sum $\Psi_a + \Psi_b$, with properties

$$\Psi_a + \Psi_b = \Psi_b + \Psi_a \tag{9.23}$$

$$\Psi_a + (\Psi_b + \Psi_c) = (\Psi_a + \Psi_b) + \Psi_c \tag{9.24}$$

These rules define the addition of vectors.

We also define the multiplication of a vector by an arbitrary complex number λ. This is done by associating with any vector Ψ a vector $\lambda\Psi$, subject to the rules

$$(\lambda + \mu)\Psi = \lambda\Psi + \mu\Psi, \qquad \mu(\lambda\Psi) = (\mu\lambda)\Psi \qquad (\lambda, \mu\text{: complex numbers}) \tag{9.25}$$

$$\lambda(\Psi_a + \Psi_b) = \lambda\Psi_a + \lambda\Psi_b, \quad \text{and} \quad 1 \cdot \Psi = \Psi \tag{9.26}$$

The vector space contains the null vector, $\mathbf{0}$, such that

$$\Psi + \mathbf{0} = \Psi \quad \text{and} \quad 0 \cdot \Psi = \mathbf{0} \tag{9.27}$$

for any vector Ψ. The null vector $\mathbf{0}$ will sometimes be written just plainly as 0, as if it were a number. This is not strictly correct, but should not be misleading.

The k vectors $\Psi_1, \Psi_2, \ldots, \Psi_k$ are said to be *linearly independent* if no relation

$$\lambda_1\Psi_1 + \lambda_2\Psi_2 + \cdots + \lambda_k\Psi_k = 0$$

exists between them, except the trivial one with $\lambda_1 = \lambda_2 = \cdots = \lambda_k = 0$. The vector space is said to be n-dimensional if there exist n linearly independent vectors, but if no $n+1$ vectors are linearly independent.

In an n-dimensional space, we may choose a set of n linearly independent vectors $\Psi_1, \Psi_2, \ldots, \Psi_n$. We refer to these vectors as the members of a *basis*, or as *basis vectors*. They are said to *span* the space, or to form a *complete* set of vectors, since an arbitrary vector Ψ can be expanded in terms of them:

$$\Psi = \sum_{i=1}^{n} a_i\Psi_i \tag{9.28}$$

The coefficients a_i are complex numbers. They are called the *components* of the vector Ψ, which they are said to *represent*. The components determine the vector completely. The components of the sum of two vectors are equal to the sum of the components: If $\Psi_a = \sum a_i\Psi_i$ and $\Psi_b = \sum b_i\Psi_i$, then

$$\Psi_a + \Psi_b = \sum_i (a_i + b_i)\Psi_i \tag{9.29}$$

[3]For a compendium of mathematical results and caveats, see Jordan (1969). More detail is found in Reed and Simon (1975–80) and Riesz and Nagy (1955). See also Fano (1971).

and similarly

$$\lambda\Psi_a = \sum_i (\lambda a_i)\Psi_i \tag{9.30}$$

by the above rules for addition and multiplication.

Next we introduce an *inner* (scalar) product between two vectors, denoted by the symbol (Ψ_a, Ψ_b). This is a complex number with the following properties:

$$(\Psi_b, \Psi_a) = (\Psi_a, \Psi_b)^* \tag{9.31}$$

where the asterisk denotes complex conjugation. In the inner product (Ψ_a, Ψ_b), Ψ_a is called the prefactor, and Ψ_b is the postfactor. Their order is important.

We further require that

$$(\Psi_a, \lambda\Psi_b) = \lambda(\Psi_a, \Psi_b) \tag{9.32}$$

From (9.31) and (9.32) it follows that

$$(\lambda\Psi_a, \Psi_b) = \lambda^*(\Psi_a, \Psi_b) \tag{9.33}$$

We also postulate that

$$(\Psi_a, \Psi_b + \Psi_c) = (\Psi_a, \Psi_b) + (\Psi_a, \Psi_c) \tag{9.34}$$

and that

$$(\Psi, \Psi) \geq 0 \tag{9.35}$$

with the equality sign holding if and only if Ψ is the null vector. The value $\sqrt{(\Psi, \Psi)} \equiv \|\Psi\|$ is called the *norm* or the "length" of the vector Ψ. A vector for which $(\Psi, \Psi) = 1$ is called a *unit* vector (or a normalized vector).

Two vectors, neither of which is a null vector, are said to be *orthogonal* if their inner product vanishes.

It is possible to construct sets of n vectors that satisfy the orthogonality and normalization conditions (often briefly referred to as the *orthonormality* property)

$$(\Psi_i, \Psi_j) = \delta_{ij} \qquad (i, j = 1, \ldots, n) \tag{9.36}$$

Since mutually orthogonal vectors are linearly independent, an orthonormal set can serve as a suitable basis. Generally, we assume that the basis vectors form an orthonormal set, but occasionally, as in Section 8.4, a nonorthogonal basis is preferred.

By taking the inner product with one of the orthonormal basis vectors, we obtain from the expansion (9.28) the formula

$$a_i = (\Psi_i, \Psi) \tag{9.37}$$

for the components of Ψ.

The scalar product of two arbitrary vectors is expressible as

$$(\Psi_a, \Psi_b) = \left(\sum_i a_i\Psi_i, \sum_j b_j\Psi_j\right) = \sum_i \sum_j a_i^* b_j (\Psi_i, \Psi_j) = \sum_i a_i^* b_i \tag{9.38}$$

In particular, for the square of the norm of a vector

$$(\Psi, \Psi) = \|\Psi\|^2 = \sum_i |a_i|^2 \tag{9.39}$$

Exercise 9.4. If $f_a(\Psi)$ is a complex-valued (scalar) functional of the variable vector Ψ with the *linearity* property

$$f_a(\lambda\Psi_b + \mu\Psi_c) = \lambda f_a(\Psi_b) + \mu f_a(\Psi_c) \tag{9.40}$$

where λ and μ are arbitrary complex numbers, show that f_a can be represented as an inner product,

$$f_a(\Psi) = (\Psi_a, \Psi) \tag{9.41}$$

for every Ψ in the space and that (9.41) defines the vector Ψ_a uniquely. This is the finite-dimensional analogue of *Riesz's representation theorem*, which assures us that a vector is fully and uniquely specified by its projections on all vectors in the space.

3. ***Operators.*** We are now in a position to define operators in the vector space. An operator A is a prescription or mapping by which every vector Ψ in the space is associated with a vector Ψ' in the space. Thus, Ψ' is a function of Ψ, and the notation

$$\Psi' = A(\Psi) \tag{9.42}$$

is employed. The special class of operators that satisfy the conditions

$$A(\Psi_a + \Psi_b) = A(\Psi_a) + A(\Psi_b) \tag{9.43}$$

$$A(\lambda\Psi) = \lambda A(\Psi) \qquad (\lambda\text{: arbitrary complex number}) \tag{9.44}$$

is most important to us. Such operators are called *linear*. For linear operators the parenthesis in (9.42) can be dropped, and we may simply write

$$\Psi' = A\Psi \tag{9.45}$$

thus stressing that the application of a linear operator is in many ways similar to ordinary multiplication of a vector by a number.

Exercise 9.5. Prove that any linear operator A has the property $A\mathbf{0} = 0$, if $\mathbf{0}$ is the null vector.

On occasion we will also deal with *antilinear* operators. These share the property (9.43) with linear operators, but (9.44) is replaced by

$$A(\lambda\Psi) = \lambda^* A\Psi \tag{9.46}$$

Unless it is specifically stated that a particular operator is not linear, we will assume every operator to be linear and usually omit the adjective "linear."

Two operators, A and B, are equal if $A\Psi = B\Psi$ for every Ψ. Just as numbers can be added and multiplied, it is also sensible to define sums and products of operators by the relations:

$$(A + B)\Psi = A\Psi + B\Psi \tag{9.47}$$

$$(AB)\Psi = A(B\Psi) \tag{9.48}$$

The last equation states that the operator AB acting on Ψ produces the same vector that would be obtained if we first let B act on Ψ and then A on the result of the previous operation. But whereas with numbers $ab = ba$, there is no need for operators to yield the same result if they are applied in the reverse order. Hence, in general $AB \neq BA$, although in particular cases two operators may, of course, com-

mute. For instance, A does commute with any function $f(A)$ if $f(x)$ is an analytic function of the variable x.

A trivial example of a linear operator is the *identity* operator, denoted by I, with the property that

$$\Psi = I\Psi$$

for every Ψ. The operator λI, where λ is a number, merely multiplies each vector by the constant factor λ. Hence, this operator may be simply written as λ.

A less trivial example of a linear operator is provided by the equation

$$\Psi' = \Psi_a(\Psi_a, \Psi) \tag{9.49}$$

where Ψ_a is a given unit vector. This equation associates a vector Ψ' with every Ψ. The association is a linear one, and we write (9.49) as

$$\Psi' = P_a\Psi \tag{9.50}$$

defining the linear operator P_a.

Reasonably, P_a is termed a *projection* operator, since all Ψ' are in the "direction" of Ψ_a, and the "length" of Ψ' equals the absolute value of the component (Ψ_a, Ψ) of Ψ in that "direction."

Exercise 9.6. Prove that P_a is a linear operator.

Since

$$P_a\Psi_a = \Psi_a$$

it follows that for any vector Ψ,

$$P_a^2\Psi = P_a(P_a\Psi) = P_a\Psi_a(\Psi_a, \Psi) = \Psi_a(\Psi_a, \Psi) = P_a\Psi$$

Thus, projection operators are *idempotent*, i.e.,

$$P_a^2 = P_a \tag{9.51}$$

In particular, for the projections on the basis vectors we have

$$P_i\Psi = \Psi_i(\Psi_i, \Psi) = \Psi_i a_i \tag{9.52}$$

Hence,

$$P_iP_j\Psi = a_jP_i\Psi_j = 0$$

Consequently, the projection operators for the basis vectors have the property

$$P_iP_j = P_jP_i = 0 \qquad \text{if } i \neq j \tag{9.53}$$

Note also that for every Ψ

$$\sum_{i=1}^{n} P_i\Psi = \sum_{i=1}^{n} a_i\Psi_i = \Psi$$

Hence,

$$\sum_{i=1}^{n} P_i = I \tag{9.54}$$

When a basis is given, an operator A can be characterized by its effect on the basis vectors. Indeed, being again a vector in the space, $A\Psi_j$ can obviously be expanded as

$$A\Psi_j = \sum_{i=1}^{n} \Psi_i A_{ij} \qquad (j = 1, 2, \ldots, n) \tag{9.55}$$

where the A_{ij} are n^2 numbers which, owing to the linearity of A, completely specify the effect of A on any vector. To see this explicitly, we note

$$\Psi_b = A\Psi_a = A \sum_j a_j \Psi_j = \sum_j a_j A\Psi_j = \sum_j \sum_i a_j \Psi_i A_{ij} = \sum_i \Psi_i \left(\sum_i A_{ij} a_j \right)$$

Hence, since $\Psi_b = \sum_i b_i \Psi_i$, we have

$$\boxed{b_i = \sum_j A_{ij} a_j} \tag{9.56}$$

proving the contention that the effect of A on any vector is known if all A_{ij} are known. The order of the indices on the right-hand side of the defining equation (9.55) is deliberately chosen to produce the simple relation (9.56).

Equation (9.56) can be written conveniently in *matrix* form as

$$\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \vdots\vdots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \tag{9.57}$$

The possibility of using matrix notation here is not the result of any strange coincidence. Rather, the peculiar rule by which matrices are multiplied is naturally adapted to the theory of linear transformations and any calculation in which linear quantities play a fundamental role.

The inner product of two vectors can also be written in matrix notation. According to (9.38), we have

$$(\Psi_a, \Psi_b) = \sum_i a_i^* b_i = (a_1^* a_2^* \cdots a_n^*) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \tag{9.58}$$

The choice of a particular basis determines the matrices in (9.57) and (9.58). The column matrices

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

represent the vectors Ψ_a and Ψ_b, and the square matrix

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}$$

represents the operator A. For this reason, we say that all these matrices constitute a *representation* of the abstract vector space. The *matrix elements* of A in a given representation with an orthonormal basis can be calculated by the formula

$$\boxed{A_{ij} = (\Psi_i, A\Psi_j)} \tag{9.59}$$

which follows immediately from (9.55) and (9.36).

As an example, if Ψ_a is an arbitrary *unit* vector with components a_i, the matrix elements of the projection operator P_a are

$$(P_a)_{ij} = (\Psi_i, \Psi_a(\Psi_a, \Psi_j)) = (\Psi_a, \Psi_j)(\Psi_i, \Psi_a) = a_j^* a_i \tag{9.60}$$

Exercise 9.7. Show that for a linear operator

$$(\Psi_b, A\Psi_a) = \sum_i \sum_j b_i^* A_{ij} a_j \tag{9.61}$$

and write this equation in matrix form. If A were antilinear, how would the corresponding equation look?

Exercise 9.8. If $F_A(\Psi_a, \Psi_b)$ is a complex-valued (scalar) functional of the two variable vectors Ψ_a and Ψ_b with the linearity properties

$$F_A(\lambda\Psi_a + \mu\Psi_c, \Psi_b) = \lambda^* F_A(\Psi_a, \Psi_b) + \mu^* F_A(\Psi_c, \Psi_b)$$
$$F_A(\Psi_a, \lambda\Psi_b + \mu\Psi_c) = \lambda F_A(\Psi_a, \Psi_b) + \mu F_A(\Psi_a, \Psi_c)$$

show that F_A can be represented as an inner product,

$$F_A(\Psi_a, \Psi_b) = (\Psi_a, A\Psi_b) \tag{9.62}$$

for every Ψ_a and Ψ_b, and that (9.62) defines a linear operator A uniquely. (Compare with Exercise 9.4.)

The matrix representing the sum of two operators is obtained by adding the matrices representing the two operators:

$$(A + B)_{ij} = A_{ij} + B_{ij} \tag{9.63}$$

For the matrix of the product of two operators, we have

$$\begin{aligned} (AB)_{ij} = (\Psi_i, AB\Psi_j) &= \left(\Psi_i, A \sum_k \Psi_k B_{kj}\right) \\ &= \sum_k (\Psi_i, A\Psi_k) B_{kj} = \sum_k A_{ik} B_{kj} \end{aligned} \tag{9.64}$$

This result shows that the matrix of an operator product equals the product of the matrices representing the operators, taken in the same order.

Exercise 9.9. Prove the multiplication rule (9.64) directly from (9.55) without assuming the basis to be orthonormal.

Although there is a potential for confusion, we will denote the matrix representing an operator A by the same symbol A. We thus emphasize the parallelism between linear operators and matrices and rely on the context to establish the intended meaning. However, the reader should avoid a complete identification of the two concepts, tempting as it may be, because the *matrix* A depends on the particular choice of basis vectors, whereas the *operator* A is a geometric entity, represented by a different matrix in every representation. We will return to this point when we consider the connection between different bases.

It follows from Exercise 9.8 that corresponding to any given linear operator A we may define another linear operator, denoted by $A^\dagger$ and called the (***Hermitian***) *adjoint* of A, which has the property that for any two vectors, Ψ_a and Ψ_b,

$$(\Psi_a, A^\dagger \Psi_b) = (A\Psi_a, \Psi_b) \tag{9.65}$$

Specializing to unit vectors $\Psi_a = \Psi_i$, $\Psi_b = \Psi_j$, we see that

$$\boxed{(A^\dagger)_{ij} = (A)^*_{ji}}$$

Thus the *matrix* representing the operator $A^\dagger$ is obtained from the matrix representing the operator A by complex conjugation and transposition:

$$A^\dagger = \tilde{A}^* \tag{9.66}$$

where the symbol $\tilde{A}$ is used for the transpose of matrix A, and $\tilde{A}^*$ is called the *Hermitian conjugate* of A. Note also that

$$(\Psi_a, A\Psi_b) = (A\Psi_b, \Psi_a)^* = (\Psi_b, A^\dagger \Psi_a)^* = (A^\dagger \Psi_a, \Psi_b)$$

From this and (9.65) we see that an operator can be moved at will from its position as multiplier of the postfactor in an inner product to a new position as multiplier of the prefactor, and vice versa, provided that the operator's adjoint is taken.

An important theorem concerns the adjoint of the product

$$(AB)^\dagger = B^\dagger A^\dagger \tag{9.67}$$

for operators or the matrices representing them. The proof is left to the reader.

Exercise 9.10. If A is a linear operator, show that it is generally not possible to define a (''transpose'') *linear* operator A_t which satisfies the equation

$$(\Psi_b, A\Psi_a) = (\Psi_a, A_t \Psi_b) \tag{9.68}$$

Show, however, that Eq. (9.68) defines an *antilinear* operator A_t if A is itself antilinear.

A linear operator which is identical with its adjoint is said to be *Hermitian* (or self-adjoint). For a Hermitian operator,

$$A^\dagger = A \tag{9.69}$$

Hermitian operators thus are generalizations of real numbers (which are identical with their complex conjugates).

If A is a Hermitian operator, the corresponding matrix A satisfies the condition

$$\tilde{A}^* = A \tag{9.70}$$

That is, the matrix elements that are located symmetrically with respect to the main diagonal are complex conjugates of each other. In particular, the diagonal matrix elements of a Hermitian operator are real. Matrices that satisfy condition (9.70) are also called *Hermitian*.

For a Hermitian operator

$$(\Psi_a, A\Psi_a) = (A\Psi_a, \Psi_a) = (\Psi_a, A\Psi_a)^* = \text{real} \tag{9.71}$$

The physical interpretation makes important use of the reality of this scalar product which is brought into correspondence with the *expectation value* of a physical quantity represented by the Hermitian operator A.

An example of a Hermitian operator is afforded by the projection operator P_a. Indeed,

$$\begin{aligned}(\Psi_b, P_a\Psi_c) &= (\Psi_b, \Psi_a(\Psi_a, \Psi_c)) = (\Psi_a, \Psi_c)(\Psi_b, \Psi_a) \\ &= (\Psi_a, \Psi_b)^*(\Psi_a, \Psi_c) = (\Psi_a(\Psi_a, \Psi_b), \Psi_c) = (P_a\Psi_b, \Psi_c)\end{aligned}$$

A Hermitian operator A is said to be *positive definite* if for every vector Ψ,

$$(\Psi, A\Psi) \geq 0 \tag{9.72}$$

Exercise 9.11. Prove that projection operators are positive definite.

For positive definite operators we can derive a useful generalized *Schwarz inequality* by substituting in (9.72) the vector

$$\Psi = \Psi_a + \lambda\Psi_b \tag{9.73}$$

where Ψ_a and Ψ_b are arbitrary vectors, and λ may be any complex number. We obtain

$$(\Psi, A\Psi) = (\Psi_a, A\Psi_a) + \lambda(\Psi_a, A\Psi_b) + \lambda^*(\Psi_b, A\Psi_a) + \lambda^*\lambda(\Psi_b, A\Psi_b) \geq 0$$

The "best" inequality is obtained if λ is chosen so as to minimize the left-hand side. By differentiation, the value of λ that accomplishes this is found to be

$$\lambda = -\frac{(\Psi_b, A\Psi_a)}{(\Psi_b, A\Psi_b)} = -\frac{(\Psi_a, A\Psi_b)^*}{(\Psi_b, A\Psi_b)}$$

Substituting this value of λ into the above inequality, we obtain the important result:

$$\boxed{(\Psi_a, A\Psi_a)(\Psi_b, A\Psi_b) \geq |(\Psi_a, A\Psi_b)|^2} \tag{9.74}$$

For the choice $A = 1$, the *Schwarz inequality*

$$\boxed{(\Psi_a, \Psi_a)(\Psi_b, \Psi_b) \geq |(\Psi_a, \Psi_b)|^2} \tag{9.75}$$

is derived.

It follows from the Schwarz inequality that the inner product of two vectors (Ψ_a, Ψ_b) is finite if the norms of Ψ_a and Ψ_b are finite. The Schwarz inequality may

be interpreted geometrically as stating that the cosine of the "angle" α between two vectors

$$\cos \alpha = \frac{|(\Psi_a, \Psi_b)|}{\sqrt{(\Psi_a, \Psi_a)(\Psi_b, \Psi_b)}} = \frac{|(\Psi_a, \Psi_b)|}{\|\Psi_a\| \ \|\Psi_b\|}$$

has modulus less than or equal to unity.

We note that in (9.75) the equality sign holds if and only if $(\Psi, \Psi) = 0$, i.e. if $\Psi_a + \lambda\Psi_b = 0$, or if and only if Ψ_a and Ψ_b are multiples of each other, which may be interpreted to imply that they are "parallel" vectors.

A linear operator A, which is defined by

$$\Psi' = A\Psi \tag{9.45}$$

may or may not have an *inverse*. An operator B which reverses the action of A, such that

$$\Psi = B\Psi' \tag{9.76}$$

exists only if (9.45) associates different vectors Ψ'_a and Ψ'_b, with any two different vectors Ψ_a and Ψ_b or in other words if the operator A preserves linear independence. Hence, for such an operator, as Ψ runs through the entire n-dimensional vector space, Ψ' does the same. We may substitute (9.45) in (9.76) or vice versa, and conclude that

$$AB = 1 \quad \text{and} \quad BA = 1 \tag{9.77}$$

The operator B is unique, for if there were another operator B' with the property $AB' = 1$, we would have $A(B' - B) = 0$, or $BA(B' - B) = 0$; according to the second of the equations (9.77), this implies $B' - B = 0$. It is therefore legitimate to speak of *the inverse* of A and use the notation $B = A^{-1}$. Evidently,

$$(A_1 A_2)^{-1} = A_2^{-1} A_1^{-1} \tag{9.78}$$

If an operator A has an inverse, there can be no vector Ψ (other than the null vector) such that

$$A\Psi = 0 \tag{9.79}$$

Conversely, if there is no nontrivial vector which satisfies (9.79), then A has an inverse.

Exercise 9.12. Show that a projection operator has no inverse (if the vector space has more than one dimension).

The matrix that represents A^{-1} is the inverse of the matrix A. A necessary and sufficient condition for the existence of the inverse matrix is that $\det A \neq 0$.

A linear operator whose inverse and adjoint are identical is called *unitary*. Such operators are generalizations of complex numbers of absolute value 1, that is, e^{ia}. For a unitary operator U:

$$U^{-1} = U^\dagger \quad \text{or} \quad UU^\dagger = U^\dagger U = 1 \tag{9.80}$$

Evidently,

$$\boxed{(\Psi_a, \Psi_b) = (U\Psi_a, U\Psi_b)} \tag{9.81}$$

Hence, a unitary operator preserves the norms or "lengths" of vectors and the "angles" between any two of them. In this sense U can be regarded as defining a "rotation" in the abstract vector space. In fact, the matrix representing U satisfies the condition of unitarity,

$$U\tilde{U}^* = 1 \tag{9.82}$$

If U is a real matrix, condition (9.82) becomes identical with the orthogonality relation in Euclidean space, emphasizing the formal analogy between unitary operators in the complex vector space and rotations in ordinary space.

Exercise 9.13. Prove that products of unitary operators are also unitary.

Hermitian and unitary operators are special cases of the class of *normal* operators. A linear operator A is said to be normal if it commutes with its adjoint:

$$[A, A^\dagger] = 0 \tag{9.83}$$

An arbitrary operator A can be written according to (4.6) as

$$A = \frac{A + A^\dagger}{2} + i\frac{A - A^\dagger}{2i} = H_1 + iH_2 \tag{9.84}$$

where both H_1 and H_2 are Hermitian operators. The operator A is then seen to be normal if and only if

$$[H_1, H_2] = 0 \tag{9.85}$$

In dealing with expressions containing operators and their commutators, it is helpful to have recourse to some algebraic rules and identities. Since these were already developed and compiled in Section 3.4, there is no need to rederive them here. They are as valid in abstract linear vector space as in the coordinate or momentum representation.

4. ***The Vector Space of Quantum Mechanics and the Bra-Ket Notation.*** We can now proceed to express the physical principles of Section 9.1 in terms of the mathematics of vector spaces summarized in Sections 9.2 and 9.3.

The physical state Ψ of a system is fully characterized by the probability amplitudes $\langle K_i|\Psi\rangle$ relative to a complete set of observables. These probability amplitudes may now be taken to be the components of a *state vector*, which we shall also denote by Ψ. The basis vectors Ψ_i are the states in which the observable K may, in the interest of brevity, be said to have the sharp value K_i. By this we mean only that the particular value K_i is certain to be found if K is measured in state Ψ_i. An equivalent notation is $\langle K_i|\Psi\rangle \equiv \langle \Psi_i|\Psi\rangle$. The state Ψ is thus expressed as

$$\Psi = \sum_i \Psi_i \langle K_i|\Psi\rangle = \sum_i \Psi_i \langle \Psi_i|\Psi\rangle \tag{9.86}$$

from which we see that, on account of the orthonormality relation (9.36),

$$\boxed{\langle K_i|\Psi\rangle = \langle \Psi_i|\Psi\rangle = (\Psi_i, \Psi)} \tag{9.87}$$

demonstrating the identification of the *probability amplitude* $\langle K_i|\Psi\rangle$ or $\langle \Psi_i|\Psi\rangle$ with the inner product (Ψ_i, Ψ). From now on the three expressions (9.87) will therefore be used interchangeably with no misgivings. By taking the inner product of (9.86)

with the state in which the observable L has the sharp value L_j, Eq. (9.86) is seen to be consistent with the composition rule (9.8) of probability amplitudes. This circumstance, above all, testifies to the appropriateness of using vector spaces as the formal structure in which to conduct quantum mechanics.

Instead of thinking of the inner product (9.87) as a product of two vectors Ψ and Ψ_i, properly ordered, we can also think of the two factors as vectors in two *different* spaces. The postfactor Ψ is said to be a vector in *ket-space* and is denoted by $|\Psi\rangle$. The prefactor Ψ_i is a vector in *bra-space* and is denoted by $\langle\Psi_i|$. Their product $\langle\Psi_i\|\Psi\rangle$ is defined to be the inner product $\langle\Psi_i|\Psi\rangle$. The distinction between the two vector spaces arises because the inner product is not commutative, since generally

$$\langle\Psi_a|\Psi_b\rangle = \langle\Psi_b|\Psi_a\rangle^* \neq \langle\Psi_b|\Psi_a\rangle$$

The same state Ψ_a can be expressed by either a ket vector $|\Psi_a\rangle$ or a bra vector $\langle\Psi_a|$. Dirac has stylized this notation into $|a\rangle$ and $\langle a|$ for the two kinds of vectors. The inner product is written as

$$\boxed{\langle\Psi_a|\Psi_b\rangle = \langle a|b\rangle = (\Psi_a, \Psi_b)} \tag{9.88}$$

This notation has led to the colorful designation of the $\langle a|$ vector as a *bra* and the $|a\rangle$ vector as a *ket*. To every ket $|a\rangle$ there corresponds a bra $\langle a|$, and vice versa, subject to the conditions

$$|a\rangle + |b\rangle \leftrightarrow \langle a| + \langle b| \tag{9.89}$$

and

$$\lambda|a\rangle \leftrightarrow \lambda^*\langle a| \tag{9.90}$$

where the arrows indicate the correspondence between the two spaces. Taken by itself, each one of the two spaces is a vector space satisfying postulates (9.23)–(9.27). The two spaces are said to be *dual* to each other. Kets are analogues of what in differential geometry are called vectors (or contravariant vectors), and we use the notation $|\Psi\rangle$ or Ψ for them interchangeably. Bras $\langle\Psi|$ in this context are *one-forms* (or covariant vectors); and inner products are then also referred to as *contractions*.

Equation (9.86) can be written in ket form as

$$|\Psi\rangle = \sum_i |K_i\rangle\langle K_i|\Psi\rangle \tag{9.91}$$

The corresponding bra equation is

$$\langle\Psi| = \sum \langle K_i|\Psi\rangle^*\langle K_i| = \sum \langle\Psi|K_i\rangle\langle K_i| \tag{9.92}$$
$$i \qquad\qquad i$$

The components of a bra are complex conjugates of the components of the corresponding ket, since

$$\langle b|a\rangle = \langle a|b\rangle^* \tag{9.93}$$

The state $|\Psi\rangle$ is normalized to unity if

$$\langle\Psi|\Psi\rangle = \sum_i \sum_j \langle K_i|K_j\rangle\langle K_i|\Psi\rangle^*\langle K_j|\Psi\rangle = \sum_i |\langle K_i|\Psi\rangle|^2 = 1 \tag{9.94}$$

consistent with the normalization condition (9.5).

The bra-ket notation also extends to operators. A linear operator A associates with every ket $|a\rangle$ another ket,

$$|b\rangle = A|a\rangle$$

such that

$$A(|a_1\rangle + |a_2\rangle) = A|a_1\rangle + A|a_2\rangle$$

and

$$A\lambda|a\rangle = \lambda(A|a\rangle)$$

Some unnecessary vertical bars on the left of certain kets have been omitted for reasons of economy in notation.

By letting the equation

$$\langle c|\{A|a\rangle\} = \{\langle c|A\}|a\rangle = \langle c|A|a\rangle \tag{9.95}$$

define a bra $\langle c|A$, we may allow a linear operator also to act from right to left, or "backwards." The rules for multiplication to the left are

$$(\langle c_1| + \langle c_2|)A = \langle c_1|A + \langle c_2|A$$

and

$$(\langle c|\lambda)A = (\langle c|A)\lambda$$

We thus see that A is a linear operator in bra-space as well as in ket-space, and we are justified in writing the expression (9.95) unambiguously as $\langle c|A|a\rangle$. We can consider the operator as acting either to the right or to the left, whichever is convenient, emphasizing the symmetry of the dual spaces.

The definition (9.65) of the Hermitian adjoint operator becomes, in bra-ket notation,

$$\langle a|A^\dagger|c\rangle = \langle c|A|a\rangle^* \tag{9.96}$$

Hence, the general correspondence

$$|b\rangle = A|a\rangle \leftrightarrow \langle b| = \langle a|A^\dagger \tag{9.97}$$

is established.

A Hermitian operator A is characterized in this notation by

$$\langle a|A|b\rangle = \langle b|A|a\rangle^* \tag{9.98}$$

An example of a Hermitian operator is the projection operator P_a, defined by (9.49) and (9.50). In bra-ket notation, it is written as

$$P_a = |a\rangle\langle a| \tag{9.99}$$

Using the projection operators for the basis vectors

$$P_i = |K_i\rangle\langle K_i| \tag{9.100}$$

a general state, as given by (9.91) and (9.92), can be expressed in terms of projection operators as

$$|\Psi\rangle = \sum_i |K_i\rangle\langle K_i|\Psi\rangle = \sum_i P_i|\Psi\rangle \tag{9.101}$$

and

$$\langle\Psi| = \sum_i \langle\Psi|K_i\rangle\langle K_i| = \sum_i \langle\Psi|P_i \tag{9.102}$$

Since $|\Psi\rangle$ is arbitrary, it follows that the projection operators P_i satisfy the *closure* relation

$$\boxed{\sum_i P_i = \sum_i |K_i\rangle\langle K_i| = 1} \tag{9.103}$$

The probability doctrine of quantum mechanics leads to the definition of the *expectation value* $\langle K\rangle$ of an observable K in the state $|\Psi\rangle$. If $|\Psi\rangle$ is normalized to unity:

$$\langle K\rangle = \sum_i K_i|\langle K_i|\Psi\rangle|^2 = \sum_i \langle\Psi|K_i\rangle K_i\langle K_i|\Psi\rangle \tag{9.104}$$

where K_i are the possible results of measuring K. From this expression we infer that the observable K is represented by the Hermitian operator

$$\boxed{K = \sum_i |K_i\rangle K_i\langle K_i|} \tag{9.105}$$

No confusion is likely to arise from the use of the same symbol K for both the observable and the operator that represents it.

With (9.105), the expectation value (9.104) of K becomes

$$\boxed{\langle K\rangle = \langle\Psi|K|\Psi\rangle} \tag{9.106}$$

which is one of the fundamental formulas of quantum mechanics.

Exercise 9.14. Prove that K defined by (9.105) is a Hermitian operator if all numbers K_i are real.

If $|\Psi\rangle = |K_j\rangle$, all probabilities $|\langle K_i|\Psi\rangle|^2$ for $i \neq j$ are zero, and $|\langle K_j|\Psi\rangle|^2 = 1$. Thus, in the state $|K_j\rangle$ the observable K assumes the *sharp* and definite value K_j. The state $|K_j\rangle$ is an *eigenvector* of the operator K, as

$$K|K_j\rangle = \sum_i |K_i\rangle K_i\langle K_i|K_j\rangle = \sum_i |K_i\rangle K_i\delta_{ij} = K_j|K_j\rangle \tag{9.107}$$

and the sharp value K_j is an *eigenvalue* of K. The eigenvalue problem of Hermitian operators in quantum mechanics will receive detailed consideration in Chapter 10.

Just as a state is specified by its components in a given basis, an operator is specified by its *matrix elements*. As in (9.55), for an operator A these are defined as the expansion coefficients in the equation:

$$A|K_j\rangle = \sum_i |K_i\rangle\langle K_i|A|K_j\rangle = \sum_i |K_i\rangle A_{ij} \tag{9.108}$$

which follows from the identity (9.103). By comparison with (9.59), we deduce that

$$A_{ij} = (\Psi_i, A\Psi_j) = \langle K_i|A|K_j\rangle \tag{9.109}$$

The matrix representing the special Hermitian operator K, whose eigenvectors according to (9.107) make up the chosen basis, has a particularly simple structure, since

$$\langle K_i|K|K_j\rangle = K_i\delta_{ij} \tag{9.110}$$

This matrix is "diagonal," i.e. all off-diagonal elements ($i \neq j$) are zero, and the eigenvalues of K are the diagonal matrix elements.

The deceptive ease with which it is possible to derive useful relations by employing the identity (9.103) is exemplified by a calculation of the matrix elements of the product of two operators A and B:

$$\begin{aligned}\langle K_i|AB|K_j\rangle &= \langle K_i|A1B|K_j\rangle = \langle K_i|A\sum_\ell |K_\ell\rangle\langle K_\ell|B|K_j\rangle \\ &= \sum_\ell \langle K_i|A|K_\ell\rangle\langle K_\ell|B|K_j\rangle\end{aligned} \tag{9.111}$$

which is the matrix element of the product of the two matrices representing A and B, as in Eq. (9.64).

Exercise 9.15. For an arbitrary normalized state $|a\rangle$ and an operator A, calculate the sum $\sum_i |\langle a|A|K_i\rangle|^2$ over the entire basis $|K_i\rangle$. What value is obtained if A is unitary?

5. ***Change of Basis.*** In introducing the notion of the complex vector space for quantum mechanical states, we are guided by the similarity between the geometry of this abstract vector space and geometry in ordinary Euclidean space.

A *representation* in our vector space corresponds to the introduction of a *coordinate system* in Euclidean space. Just as we study rotations of coordinate systems in analytic geometry, we must now consider the transformation from one representation to another in the general vector space.

Along with an old basis Ψ_i, used in (9.28), we consider a new basis, $\overline{\Psi}_j$. The new basis vectors may be expressed in terms of the old ones:

$$\boxed{\overline{\Psi}_k = \sum_i \Psi_i S_{ik}} \tag{9.112}$$

The nonsingular matrix of the transformation coefficients

$$S = \begin{pmatrix} S_{11} & S_{12} & \cdots & S_{1n} \\ S_{21} & S_{22} & \cdots & S_{2n} \\ \vdots & \vdots & \cdots & \cdots \\ \vdots & \vdots & \cdots & \cdots \\ S_{n1} & S_{n2} & \cdots & S_{nn} \end{pmatrix}$$

defines the change of basis.

A succession of two such basis changes, S and R, performed in this order, is equivalent to a single one whose matrix is simply the product matrix SR.

To obtain the new components $\overline{a}_k$ of an arbitrary vector, we write

$$\Psi_a = \sum_i a_i\Psi_i = \sum_k \overline{a}_k\overline{\Psi}_k$$

Substituting $\overline{\Psi}_k$ from (9.112), we get

$$a_i = \sum_k S_{ik}\overline{a}_k \tag{9.113}$$

or

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = S\begin{pmatrix} \overline{a}_1 \\ \overline{a}_2 \\ \vdots \\ \overline{a}_n \end{pmatrix}$$

The inverse of this relation is

$$\begin{pmatrix} \overline{a}_1 \\ \overline{a}_2 \\ \vdots \\ \overline{a}_n \end{pmatrix} = S^{-1} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \tag{9.114}$$

We must also determine the connection between the matrices A and $\overline{A}$ representing the operator A in the old and new representations. Evidently, the new matrix elements are defined by

$$A\overline{\Psi}_j = \sum_i \overline{\Psi}_i \overline{A}_{ij} = \sum_i \sum_k \Psi_k S_{ki} \overline{A}_{ij}$$

But on the other hand

$$A\overline{\Psi}_j = A \sum_\ell \Psi_\ell S_{\ell j} = \sum_\ell \sum_k \Psi_k A_{k\ell} S_{\ell j}$$

Comparing the right-hand sides of these equations, we obtain in matrix notation

$$S\overline{A} = AS$$

or

$$\boxed{\overline{A} = S^{-1}AS} \tag{9.115}$$

We say that $\overline{A}$ is obtained from A by a *similarity transformation.*

Exercise 9.16. If $f(A, B, C, \ldots)$ is any function that is obtained from the matrices, $A, B, C, \ldots$ by algebraic processes involving numbers (but no other, constant, matrices), show that

$$f(S^{-1}AS, S^{-1}BS, S^{-1}CS, \ldots) = S^{-1}f(A, B, C, \ldots)S \tag{9.116}$$

Give examples.

So far in this section it has not yet been assumed that either the old or the new basis is orthonormal. If nonorthogonal basis vectors are used, the transformation coefficients can generally *not* be calculated from the formula

$$S_{ik} = (\Psi_i, \overline{\Psi}_k) \tag{9.117}$$

but this relation does hold if the old basis is orthonormal. Nonorthogonal basis vectors were allowed in the Rayleigh-Ritz method of variation-perturbation theory in Section 8.4.

If, as is usually the case, both the old and the new basis are orthonormal, the bra-ket notation is convenient. We identify Ψ_i with the unit ket $|K_i\rangle$ and $\overline{\Psi}_k$ with $|L_k\rangle$. The orthonormality of both bases is expressed as

$$\langle K_i | K_j \rangle = \delta_{ij}, \qquad \langle L_n | L_k \rangle = \delta_{nk} \tag{9.118}$$

and their closure as

$$\sum_i |K_i\rangle\langle K_i| = \sum_j |L_j\rangle\langle L_j| = 1 \tag{9.119}$$

Multiplying (9.119) from the right by $|L_k\rangle$, we get

$$|L_k\rangle = \sum_i |K_i\rangle\langle K_i | L_k\rangle \tag{9.120}$$

linking the two representations just as Eq. (9.112) does. Hence, for *orthonormal bases*, (9.117) holds and can be written in the form

$$S_{ik} = \langle K_i | L_k \rangle \tag{9.121}$$

From (9.120), by multiplication by $\langle \Psi |$ from the left and subsequent complex conjugation, we obtain

$$\langle L_j | \Psi \rangle = \sum_i \langle L_j | K_i \rangle \langle K_i | \Psi \rangle \tag{9.122}$$

expressing the new components of the state Ψ in terms of the old components, similar to (9.114). This last equation is nothing more than the composition rule (9.8) for probability amplitudes.

For two orthonormal bases, the transformation coefficients must satisfy the conditions

$$\sum_i S^*_{i\,k} S_{i\ell} = \sum_i \langle L_k | K_i \rangle \langle K_i | L_\ell \rangle = \langle L_k | L_\ell \rangle = \delta_{k\ell} \tag{9.123}$$

and

$$\sum_k S_{ik} S^*_{\ell\,k} = \sum_k \langle K_i | L_k \rangle \langle L_k | K_\ell \rangle = \langle K_i | K_\ell \rangle = \delta_{i\ell} \tag{9.124}$$

Hence, S must be a unitary matrix

$$S\,S^\dagger = S^\dagger\,S = 1 \tag{9.125}$$

where S denotes the transformation matrix with rows labeled by the eigenvalues of L and columns labeled by the eigenvalues of K.

Using the unitarity condition, we may rewrite the similarity transformation equation (9.115) for a matrix representing the operator A in the form

$$\overline{A} = S^\dagger\,AS \tag{9.126}$$

or more explicitly,

$$\langle L_j | A | L_k \rangle = \sum_{i\ell} \langle L_j | K_i \rangle \langle K_i | A | K_\ell \rangle \langle K_\ell | L_k \rangle \tag{9.127}$$

An alternative interpretation of a unitary transformation consists of regarding (9.114) not as a relation between different representations of the components of the same vector (''passive'' transformation), but as defining the components of a new vector in a fixed representation (''active'' transformation). Comparing (9.114) with (9.57), we see that if the matrix $S^{-1} = S^\dagger$ can be considered as connecting the components of two vectors in the same basis, it represents a unitary operator U that changes every vector Ψ into a vector $U\Psi$. The operator $\overline{A}$ which maps the unitary transform $U\Psi$ of Ψ into the unitary transform $U\,A\Psi$ of $A\Psi$ is defined by the equation

$$\overline{A}(U\Psi) = U(A\Psi)$$

Hence, we have the operator equation

$$\overline{A} = U\,AU^\dagger \tag{9.128}$$

which agrees with the matrix equation (9.126). The two ''rotations,'' one (passive) affecting only the basis, and the other (active) keeping the basis fixed while rotating all vectors and operators, are equivalent but they are performed in opposite ''directions''; that is, one is the inverse of the other.

Exercise 9.17. Show that under an active unitary transformation a Hermitian operator remains Hermitian, a unitary operator remains unitary, and a normal operator remains normal. Also show that a symmetric matrix does *not* in general remain symmetric under such a transformation.

Exercise 9.18. Show by the use of the bra-ket notation that

$$\text{trace } A = \sum_i \langle K_i|A|K_i\rangle \tag{9.129}$$

is independent of the choice of the basis $|K_i\rangle$ and that

$$\text{trace } AB = \text{trace } BA \tag{9.130}$$

Exercise 9.19. Show that

$$\sum_i \sum_j |\langle K_i|A|L_j\rangle|^2 = \text{trace } AA^\dagger = \text{trace } A^\dagger A \tag{9.131}$$

and that this expression is independent of the bases $|K_i\rangle$ and $|L_j\rangle$.

Dirac introduced the bra-ket notation in order to make the formal expressions of quantum mechanics more transparent and easier to manipulate. We will see in Section 9.6 that it also makes it easy to unify the formalism for finite- and infinite-dimensional, but separable, vector spaces ($n \to \infty$) with which quantum mechanics works. The bra-ket notation is not particularly convenient when, as in Section 8.4, nonorthogonal bases are employed, because such bases are not self-reciprocal. Generally, components of vectors are then not simple inner products, the matrix representing a Hermitian operator may not be Hermitian, and unitary operators are not necessarily represented by unitary matrices.

6. ***Hilbert Space and the Coordinate Representation.*** As explained at the end of Section 9.1, we have so far assumed the vector space to be of a finite number of dimensions, n, and thus any operator to have at most n distinct eigenvalues. Yet, in ordinary wave mechanics, which motivated our interest in the mathematics of linear vector spaces in the first place, most observables have infinitely many possible values (their eigenvalues), and many of those, instead of being discrete, are *indenumerable* and form a continuum. If n is allowed to grow to infinity, a number of important and difficult mathematical questions arise, to which we have alluded on several occasions.

If the limit $n \to \infty$ is to have meaning for quantum mechanics with its probability amplitudes and expectation values, we must work only with a mathematical structure that allows us to expand state vectors in terms of complete sets of orthogonal vectors. This strategy was already employed in Sections 4.3 and 4.4 for discussing the continuous spectrum of such commonplace observables as the position or momentum of a particle. In this section, we will recover the fundamentals of wave mechanics in the framework of a vector space that supports operators with eigenvalue spectra consisting of continuous portions in addition to discrete points. The bra-ket notation helps to express the unified theory of finite- and infinite-dimensional vector spaces in compact form.

The principles spelled out in Sections 9.1 and 9.4, demand that if the eigenvalue equation for the observable K,

$$K|K'\rangle = K'|K'\rangle \tag{9.132}$$

has a continuous spectrum, only those eigenvectors are admissible for which the eigenvalue K' is real and which can be normalized in terms of delta functions rather than Kronecker deltas:

$$\langle K' | K'' \rangle = \delta(K' - K'') \tag{9.133}$$

With this normalization, all the formulas for the discrete and continuous cases are similar, except that integrals in the continuous case replace sums in the discrete case. For continuously variable eigenvalues, the prime notation to distinguish eigenvalues (K') from operators (K) has been reinstated.

Thus, an arbitrary state vector can be expanded as

$$|\Psi\rangle = \sum_{K_i} |K_i\rangle\langle K_i|\Psi\rangle + \int |K'\rangle dK' \langle K'|\Psi\rangle \tag{9.134}$$

the sum being over the discrete and the integral over the continuous eigenvalues of the observable K. For simplicity, we assume that there are no repeated eigenvalues of K. It is easy to generalize the formalism to accommodate repeated eigenvalues, and this will be done in Section 10.4. It is even possible for a particular eigenvalue to belong to both the discrete and continuous spectrum. The corresponding eigenkets must be orthogonal. For an example from atomic physics, see Section 19.7.

The expansion (9.134) gives the expectation value of K as

$$\langle K \rangle = \langle \Psi | K | \Psi \rangle = \sum_{K_i} K_i |\langle K_i|\Psi\rangle|^2 + \int K' |\langle K'|\Psi\rangle|^2 \, dK' \tag{9.135}$$

showing that $|\langle K'|\Psi\rangle|^2 \, dK'$ is the probability of finding for the observable K a value between K' and $K' + dK'$ when K' lies in the continuous part of the spectrum. Thus, $|\langle K'|\Psi\rangle|^2$ is the probability per unit K'-interval, or a *probability density*.

The application of the formalism to wave mechanics for a point particle in one dimension is straightforward. The three-dimensional theory is worked out in Section 15.1. Since we can measure the particle's position along the x axis, there is a Hermitian operator x for this observable. The results of a measurement can be any real number between $-\infty$ and $+\infty$, and the eigenvalues of x, denoted here by x', form a continuum. The corresponding eigenvectors are written as $|x'\rangle$

$$x|x'\rangle = x'|x'\rangle \tag{9.136}$$

with the assumed normalization. The closure relation is now

$$\int |x'\rangle dx' \langle x'| = 1 \tag{9.137}$$

and the state is expanded as

$$|\Psi\rangle = \int |x'\rangle dx' \langle x'|\Psi\rangle \tag{9.138}$$

The components $\langle x'|\Psi\rangle$ constitute a complex-valued function of the real variable x'. By identifying it with the wave function,

$$\boxed{\psi(x') = \langle x'|\Psi\rangle} \tag{9.139}$$

we establish the connection between the *state vector* $|\Psi\rangle$ and the *wave function* $\psi(x')$. The coordinate x' serves as the continuously variable label of the component

$\psi(x')$ of the state vector $|\Psi\rangle$ in an infinite-dimensional abstract vector space. From this point of view, $\psi(x')$ is merely one of many possible ways of representing the state vector. The representation used here is spanned by the eigenvectors of the position operator x and is called the *coordinate representation*.

The inner product of two states is represented by

$$\begin{aligned}\langle\Psi_2|\Psi_1\rangle &= \iint \langle\Psi_2|x''\rangle dx''\langle x''|x'\rangle dx'\langle x'|\Psi_1\rangle \\ &= \iint \langle\Psi_2|x''\rangle\, dx''\, \delta(x'' - x')\, dx'\langle x'|\Psi_1\rangle \qquad (9.140) \\ &= \int_{-\infty}^{+\infty} \psi_2^*(x')\psi_1(x')\, dx'\end{aligned}$$

The inner product of two orthogonal states is zero, in agreement with the earlier conventions (Section 4.1).

In the coordinate representation, the matrix that represents an operator A is now a matrix whose indices, instead of being denumerable, take on a continuum of values. Although such a matrix can no longer be written down as a square array of numbers, it is nevertheless appropriate to use the term *matrix element* for the quantity $\langle x''|A|x'\rangle$. A function $f(x)$ is represented by the matrix element

$$\langle x''|f(x)|x'\rangle = f(x')\delta(x' - x'') \qquad (9.141)$$

This is said to be diagonal, since it vanishes for $x'' \neq x'$.

Linear momentum in the x direction is another important operator for the system. We know that in the coordinate representation it acts on wave functions as $(\hbar/i)\,\partial/\partial x'$, and we can use this information to conjecture that in the abstract vector space it is a Hermitian operator, p, which satisfies the commutation relation

$$xp - px = i\hbar\, 1 \qquad (9.142)$$

For brevity we omit the subscript x on the momentum p, because we are dealing with a one-dimensional system. The fundamental relationship between linear momentum and linear displacements, which is at the root of (9.142), was already discussed in Section 4.5 and will again be taken up in Section 15.1. Here we merely assume the commutation relation (9.142) and deduce its consequences.

Exercise 9.20. Taking the trace on the two sides of the commutation relation (9.142), show that a contradiction arises from the application of (9.130). Resolve the paradox.

The matrix element of the operator equation (9.142), taken between the bra $\langle x''|$ and the ket $|x'\rangle$, is

$$\langle x''|xp - px|x'\rangle = (x'' - x')\langle x''|p|x'\rangle = i\hbar\delta(x'' - x')$$

from which we infer that

$$\langle x''|p|x'\rangle = i\hbar\frac{\delta(x'' - x')}{x'' - x'} = i\hbar\frac{\partial}{\partial x'}\delta(x'' - x') = \frac{\hbar}{i}\frac{\partial}{\partial x''}\delta(x'' - x') \qquad (9.143)$$

which, though highly singular, is a well-defined quantity.

Exercise 9.21. Show that for any function $f(p)$ that can be expressed as a power series of its argument,

$$\langle x''|f(p)|x'\rangle = f\left(\frac{\hbar}{i}\frac{\partial}{\partial x''}\right)\delta(x'' - x') \tag{9.144}$$

It follows that the action of an operator $f(p)$ on a state is represented as

$$\begin{aligned} f(p)|\Psi\rangle &= \int\int |x'\rangle dx'\langle x'|f(p)|x''\rangle dx''\langle x''|\Psi\rangle \\ &= \int\int |x'\rangle dx' \; f\left(\frac{\hbar}{i}\frac{\partial}{\partial x'}\right)\delta(x' - x'')\psi(x'')dx'' \\ &= \int |x'\rangle dx' \; f\left(\frac{\hbar}{i}\frac{\partial}{\partial x'}\right)\psi(x')dx' \end{aligned} \tag{9.145}$$

and that

$$\langle\Psi_2|f(p)|\Psi_1\rangle = \int \psi_2^*(x') \; f\left(\frac{\hbar}{i}\frac{\partial}{\partial x'}\right)\psi_1(x')dx' \tag{9.146}$$

All of these results confirm that the identification of the operator p which satisfies the commutation relation (9.142) as the momentum, is correct. More generally, they reassure us that the abstract vector space (bra-ket) formulation of quantum mechanics contains wave mechanics as a representation.

Exercise 9.22. Change from the coordinate basis to the momentum basis, showing that the transformation coefficients are

$$\langle x'|p'\rangle = \frac{1}{\sqrt{2\pi\hbar}}\exp\left(\frac{i}{\hbar}p'x'\right)$$

Represent states and operators in the momentum basis. Derive the equations connecting the expressions for components of states and matrix elements of operators in the coordinate and momentum representations.

As an illustration, we reconsider the linear harmonic oscillator with its Hamiltonian or energy operator,

$$H = \frac{p^2}{2m} + \frac{m\omega^2 x^2}{2} \tag{9.147}$$

Denoting the eigenvalues of H by E as is customary (instead of H'), we see that the eigenvalue equation for the energy operator,

$$H|E\rangle = E|E\rangle \tag{9.148}$$

appears in the coordinate representation as

$$\begin{aligned} \frac{1}{2m}(p^2 + m^2\omega^2x^2)|E\rangle &= \frac{1}{2m}\int |x'\rangle\left(-\hbar^2\frac{\partial^2}{\partial x'^2}\right)\psi_E(x')dx' + m^2\omega^2\int |x'\rangle x'^2\psi_E(x')dx \\ &= E\int |x'\rangle dx'\psi_E(x') \end{aligned}$$

where $\psi_E(x') = \langle x'|E\rangle$. Since all the vectors $|x'\rangle$ are linearly independent, this equation requires that for the components of $|E\rangle$, labeled by x',

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x'^2} + \frac{m\omega^2}{2}x'^2 - E\right)\psi_E(x') = 0 \tag{9.149}$$

which is nothing but the Schrödinger equation for the harmonic oscillator. Only those solutions are admissible for which E is real and the normalization

$$\langle E_2|E_1\rangle = \int \psi^*_{E_2}(x')\psi_{E_1}(x')dx' = \delta_{12} \quad \text{or} \quad \delta(E_1 - E_2)$$

can be achieved. For an equation like (9.149), this condition is equivalent to

$$\psi_{E'}(+\infty) = \psi_{E'}(-\infty) = 0 \tag{9.150}$$

From Chapter 5 we know that only discrete eigenvalues,

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) \qquad (n = 0, 1, 2, \ldots)$$

exist for this Schrödinger equation and that, by (5.39):

$$\langle x'|E_n\rangle = \psi_{E_n}(x') = \frac{1}{2^{n/2}}\frac{1}{n!^{1/2}}\left(\frac{m\omega}{\hbar\pi}\right)^{1/4}\exp\left(-\frac{m\omega x'^2}{2\hbar}\right)H_n\left(\sqrt{\frac{m\omega}{\hbar}}\,x'\right) \tag{9.151}$$

Exploiting the freedom to choose a representation, we may use the eigenvectors of H as the basis of a representation, which we designate as the (harmonic oscillator) *energy representation*. It is spanned by a denumerable infinity of basis vectors, labeled by the quantum number n. Although it is qualitatively different from either the coordinate or the momentum representation, any of these three representations can be equivalently used to expand an arbitrary vector of the same Hilbert space. The transformation coefficients $\langle x'|E_n\rangle$ are subject to the unitarity conditions

$$\int_{-\infty}^{+\infty}\langle E_k|x'\rangle dx'\langle x'|E_n\rangle = \delta_{kn} \tag{9.152}$$

$$\sum_{n=0}^{\infty}\langle x'|E_n\rangle\langle E_n|x''\rangle = \delta(x' - x'') \tag{9.153}$$

These conditions are satisfied by virtue of the orthonormality and completeness of the eigenfunctions (9.151).

Problems

1. If $\psi_n(\mathbf{r})$ is the normalized eigenfunction of the time-independent Schrödinger equation, corresponding to energy eigenvalue E_n, show that $|\psi_n(\mathbf{r})|^2$ is not only the probability density for the coordinate vector $\mathbf{r}$, if the system has energy E_n, but also conversely the probability of finding the energy to be E_n, if the system is known to be at position $\mathbf{r}$.

2. Using the momentum representation, calculate the bound-state energy eigenvalue and the corresponding eigenfunction for the potential $V(x) = -g\delta(x)$ (for $g > 0$). Compare with the results in Section 6.4.

CHAPTER 10

Eigenvalues and Eigenvectors of Operators, the Uncertainty Relations, and the Harmonic Oscillator

A thorough understanding of the eigenvalue problem of physically relevant operators and of the corresponding eigenvectors (eigenstates) is essential in quantum mechanics. In this chapter, we examine some further ramifications of this problem. The physical significance of commutation relations will be discussed and illustrated by the Heisenberg uncertainty relations. The chapter concludes with a return to the harmonic oscillator, now in terms of raising (creation) and lowering (annihilation) operators, preparing for applications in many-particle theory. Coherent oscillator states shed light on the connection with classical mechanics and are central to the interpretation of processes in quantum optics. Squeezed (or stretched) oscillator states make it possible to give an introduction to the concept of quasiparticles.

1. ***The Eigenvalue Problem for Normal Operators.*** A ket $|A_i'\rangle$ is called an *eigenvector*, or *eigenket*, of the operator A if

$$A|A_i'\rangle = A_i'|A_i'\rangle \tag{10.1}$$

The number A_i' which characterizes the eigenvector is called an *eigenvalue*. The effect of A on $|A_i'\rangle$ is merely multiplication by a number. We first give our attention to the eigenvalue problem for *normal* operators, which include Hermitian and unitary operators. In Section 10.7 the discussion will be extended to an important nonnormal operator, the annihilation operator. Different eigenvalues will be distinguished by their subscripts. An eigenvalue enclosed in a ket $|\ \rangle$, as in $|A_i'\rangle$, names the eigenket belonging to the eigenvalue A_i'.

Our main objective will be to prove that, at least in a finite-dimensional vector space, every normal operator has a *complete* set of orthonormal eigenvectors, which may be used as basis vectors spanning the space. The normal operator A may have *repeated* eigenvalues. By this we mean the occurrence of more than one linearly independent eigenvector belonging to the same eigenvalue. (When this happens for the Hamiltonian operator of a system, we speak of *degenerate* energy eigenvalues.)

Since any linear combination of eigenvectors belonging to the same eigenvalue is again an eigenvector belonging to the same eigenvalue, these eigenvectors form a subspace of the original vector space. The original linearly independent eigenvectors in this subspace may be replaced by an equal number of *orthogonal* eigenvectors, all of which correspond to the same eigenvalue. This is accomplished by a successive orthogonalization algorithm (*Schmidt* orthogonalization method), the essence of which was already described in Section 4.1 and illustrated in Figure 4.1.

If Ψ_1 and Ψ_2 are two nonorthogonal eigenvectors with repeated eigenvalues, we construct as a replacement for Ψ_2 the eigenvector

$$\Psi_2' = \Psi_2 - \frac{\Psi_1(\Psi_1, \Psi_2)}{(\Psi_1, \Psi_1)} = \Psi_2 - P_{(\Psi_1)}\Psi_2 \tag{10.2}$$

where $P_{(\Psi_1)}$ is the projection operator in the "direction" of Ψ_1. The new eigenvector Ψ_2' belongs to the same eigenvalue as Ψ_1 and Ψ_2, but it is orthogonal to Ψ_1:

$$(\Psi_1, \Psi_2') = 0$$

If there is a third linearly independent eigenvector, Ψ_3, belonging to the same repeated eigenvalue, we replace it by

$$\Psi_3' = [1 - P_{(\Psi_1)} - P_{(\Psi_2')}]\Psi_3 \tag{10.3}$$

which is orthogonal to both Ψ_1 and Ψ_2'. This procedure is continued until a complete orthonormal basis has been constructed for the subspace of eigenvectors belonging to the repeated eigenvalue. In the bra-ket notation, the occurrence of repeated eigenvalues requires that the name of an eigenket be specified by additional labels to supplement the information conveyed by the common eigenvalue A_i'.

Exercise 10.1. Show that Ψ_3' in (10.3) is orthogonal to both Ψ_1 and Ψ_2'.

For a normal operator, for which by definition

$$\boxed{[A, A^\dagger] = 0} \tag{10.4}$$

we infer from (10.1) that

$$A^\dagger A|A_i'\rangle = AA^\dagger|A_i'\rangle = A_i' A^\dagger|A_i'\rangle \tag{10.5}$$

Hence, $A^\dagger|A_i'\rangle$ is an eigenket of A, belonging to the eigenvalue A_i' and may be written as

$$A^\dagger|A_i'\rangle = \alpha|A_i'\rangle + |\beta\rangle \quad . \tag{10.6}$$

where $|\beta\rangle$ is also an eigenvector of A with eigenvalue A_i'. We may assume that $|\beta\rangle$ is orthogonal to $|A_i'\rangle$; i.e., $\langle A_i'|\beta\rangle = 0$. From (10.1) we have also

$$\langle A_i'|A^\dagger = A_i'^*\langle A_i'| \tag{10.7}$$

as well as

$$\langle\beta|A^\dagger = A_i'^*\langle\beta| \tag{10.8}$$

If we multiply (10.6) on the left by $\langle A_i'|$ we obtain, using (10.7),

$$A_i'^*\langle A_i'|A_i'\rangle = \alpha\langle A_i'|A_i'\rangle \quad \text{or} \quad \alpha = A_i'^*$$

If we substitute this result in (10.6) and then multiply on the left by $\langle\beta|$, we find, using (10.8), that $\langle\beta|\beta\rangle = 0$ and that thus for all eigenvectors,

$$A^\dagger|A_i'\rangle = A_i'^*|A_i'\rangle \tag{10.9}$$

as well as

$$\langle A_i'|A = A_i'\langle A_i'| \tag{10.10}$$

From Eqs. (10.1) and (10.10) we obtain

$$(A_i' - A_j')\langle A_i'|A_j'\rangle = 0 \tag{10.11}$$

showing that any two eigenkets belonging to different eigenvalues of a normal operator are *orthogonal.*

We thus conclude that *all* linearly independent eigenvectors of a normal operator may be assumed to be orthogonal. It is convenient to take these eigenvectors also as being normalized to unity.

It may be useful to recapitulate here that the eigenvalue problem for a *normal* operator can be expressed in any one of four equivalent ways:

$$\boxed{\begin{aligned} A|A_i'\rangle &= A_i'|A_i'\rangle \\ \langle A_i'|A^\dagger &= A_i'^*\langle A_i'| \end{aligned}}$$

$$\boxed{\begin{aligned} A^\dagger|A_i'\rangle &= A_i'^*|A_i'\rangle \\ \langle A_i'|A &= A_i'\langle A_i'| \end{aligned}}$$

Note that the last two relations follow from the first two, or vice versa, only for normal operators.

2. ***The Calculation of Eigenvalues and the Construction of Eigenvectors.*** Although we have demonstrated the orthogonality of the eigenvectors of normal operators, we have not established that any solutions to the eigenvalue problem (10.1) actually exist, nor have we yet found a method for calculating the eigenvalues of A and for determining the corresponding eigenvectors. Our hope is not only that solutions exist, but that in an n-dimensional vector space there are n orthogonal eigenvectors so that a *complete* basis can be constructed from them. We simultaneously attack both problems—the existence of solutions of (10.1) and the program for obtaining them.

If a basis is introduced in the space, the representation of the eigenvalue problem (10.1) takes the form of a matrix equation:

$$\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{10.12}$$

The matrix elements $x_1, x_2, \ldots, x_n$ are the components of the eigenvector which belongs to the eigenvalue λ. Equation (10.12) is a set of n linear homogeneous equations that possess nontrivial solutions only if

$$\begin{vmatrix} A_{11} - \lambda & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} - \lambda & \cdots & A_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} - \lambda \end{vmatrix} = 0$$

or

$$\det(A_{ij} - \lambda\delta_{ij}) = 0 \tag{10.13}$$

This equation of the nth degree in the unknown λ is called the *secular* or *characteristic* equation. The roots of (10.13), $\lambda = A_i'$, are the eigenvalues of A.

According to the theorems of linear algebra, the existence of at least one solution of the set of homogeneous equations is assured. We may thus substitute $\lambda = A_1'$ in (10.12) and solve the set of n homogeneous equations to obtain an eigenvector $|A_1'\rangle$. We then change to a basis that includes $|A_1'\rangle$ as one of its elements. Because of the properties,

$$A|A_1'\rangle = A_1'|A_1'\rangle \quad \text{and} \quad \langle A_1'|A = A_1'\langle A_1'|$$

valid for a normal operator, the normal matrix representing the operator A in this new representation takes the partially diagonalized form:

$$\begin{pmatrix} A_1' & 0 & 0 & \cdots & 0 \\ 0 & \bar{A}_{22} & \bar{A}_{23} & \cdots & \bar{A}_{2n} \\ 0 & \bar{A}_{32} & \bar{A}_{33} & \cdots & \bar{A}_{3n} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & \bar{A}_{n2} & \cdot & \cdot & \bar{A}_{nn} \end{pmatrix}$$

The $n-1$ dimensional matrix of the matrix elements $\bar{A}_{ij}$ is again a normal matrix. Its eigenvalue problem has the same solutions as the original problem except for an eigenvector belonging to the eigenvalue A_1'. The same procedure as before can then be continued in the $n-1$ dimensional subspace, which is orthonormal to $|A_1'\rangle$. After $n-1$ such steps, the matrix representing A will result in the completely diagonal form

$$\begin{pmatrix} A_1' & 0 & 0 & \cdots & 0 \\ 0 & A_2' & 0 & \cdots & 0 \\ 0 & 0 & A_3' & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdot & A_n' \end{pmatrix}$$

and the ultimately obtained basis vectors are the eigenvectors of A. This procedure, which is equally applicable whether or not there are repeated eigenvalues, proves that for a *normal* operator n orthonormal eigenvectors can always be found. If the operator is Hermitian, all expectation values and eigenvalues are real. If the operator is unitary, the eigenvalues have modulus unity.

Exercise 10.2. Prove the converse proposition that an operator A whose eigenvectors span a complete orthonormal basis must commute with its adjoint, i.e., be normal.

Exercise 10.3. Prove that the eigenvalues U_i' of a unitary operator have the property

$$|U_i'|^2 = 1 \tag{10.14}$$

In the bra-ket notation, the eigenvalue problem for the normal operator A appears in the form

$$\sum_j \langle K_i|A|K_j\rangle\langle K_j|A_\ell'\rangle = A_\ell'\langle K_i|A_\ell'\rangle \tag{10.15}$$

As we have seen, the eigenvectors may all be chosen to be orthonormal:

$$\langle A'_i | A'_\ell \rangle = \delta_\ell \tag{10.16}$$

and they form a complete set:

$$\sum_{i=1}^{n} |A'_i\rangle\langle A'_i| = \sum_{i=1}^{n} P_i = 1 \tag{10.17}$$

where $P_i = |A'_i\rangle\langle A'^*_i|$ is the projection operator for the eigenvector $|A'_i\rangle$. (*A reminder:* If an eigenvalue is repeated, an additional label is required to characterize the eigenvectors. More about this is detailed in Section 10.4.)

The transformation matrix S with matrix elements

$$S_{i\ell} = \langle K_i | A'_\ell \rangle \tag{10.18}$$

changes the original basis $|K_i\rangle$ into one spanned by the eigenvectors of A, and it is *unitary*. The resulting form of the matrix representing A is diagonal, as is seen explicitly in the relation obtained from (10.15):

$$\sum_{ij} \langle A'_n | K_i\rangle\langle K_i | A | K_j\rangle\langle K_j | A'_\ell\rangle = A'_\ell \delta_{\ell n} \tag{10.19}$$

Either (10.15) or (10.19) may be used to determine the components of the eigenvectors of A. Quite different algebraic methods for determining eigenvalues and eigenvectors of certain privileged operators will be discussed later in this chapter.

Although a particular representation was introduced to calculate the eigenvalues of A as roots of (10.13), the eigenvalues are properties of the operator and cannot depend on the choice of basis. Indeed, if we choose some other basis, linked to the previous one by the similarity transformation (9.115),

$$\bar{A} = S^{-1}AS$$

the new characteristic equation is

$$\det(\bar{A} - \lambda 1) = \det[S^{-1}(A - \lambda 1)S] = \det(A - \lambda 1) = 0$$

In this proof the property of determinants

$$\det AB = \det A \det B \tag{10.20}$$

has been used. Hence, the eigenvalues of A as defined by (10.13) are independent of the representation. Consequently, if we expand the characteristic equation explicitly,

$$(-\lambda)^n + (\text{trace } A)(-\lambda)^{n-1} + \cdots + \det A = 0 \tag{10.21}$$

the coefficient of each power of λ must be independent of the choice of representation.

It is easy to prove these properties for the trace and the determinant of A directly. Since for finite dimensional matrices we know from (9.130),

$$\text{trace } AB = \text{trace } BA \tag{10.22}$$

it follows that

$$\text{trace } (ABC) = \text{trace}(CAB) = \text{trace}(BCA) \tag{10.23}$$

hence,

$$\text{trace } \bar{A} = \text{trace}(S^{-1}AS) = \text{trace } A \tag{10.24}$$

Similarly, using (10.20),

$$\det \bar{A} = \det(S^{-1}AS) = \det A \tag{10.25}$$

It is therefore legitimate to consider trace A and det A to be properties of the operator A itself and to attach a representation-independent meaning to them.

Furthermore, from the diagonalized form of the normal operator A we see that

$$\text{trace } A = A'_1 + A'_2 + \cdots + A'_n = \text{sum of the eigenvalues of } A \tag{10.26}$$

and

$$\det A = A'_1 \times A'_2 \times \cdots \times A'_n = \text{product of the eigenvalues of } A \tag{10.27}$$

As an application, consider the matrix e^A defined as

$$e^A = 1 + \frac{1}{1!}A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \cdots = \lim_{N\to\infty}\left(1 + \frac{A}{N}\right)^N \tag{10.28}$$

The eigenvalues of e^A are $e^{A'_i}$. Hence, we have the useful relation

$$\det e^A = \prod_i e^{A'_i} = \exp\left(\sum_i A'_i\right) = \exp(\text{trace } A) \tag{10.29}$$

Exercise 10.4. Prove (10.29) directly from (10.28) without recourse to the eigenvalues.

If $f(z)$ is an analytic function, the function $B = f(A)$ of the normal operator A can be expanded in finite powers of A very simply. From the completeness relation (10.17) we see that

$$f(A) = \sum_{i=1}^{n} f(A'_i)P_i \tag{10.30}$$

provided that the singularities of $f(z)$ do not coincide with any eigenvalue of A. If n' of the n eigenvalues of A are distinct, we may label these nonrepeating eigenvalues by the subscripts $i = 1, \ldots, n'$ and express (10.30) in the form

$$f(A) = \sum_{j=1}^{n'} f(A'_j) \prod_{\substack{i=1 \\ (i\neq j)}}^{n'} \frac{A'_i 1 - A}{A'_i - A'_j} \tag{10.31}$$

showing that any function $f(A)$ can be written as a polynomial in A of degree less than n.

Exercise 10.5. Prove that a normal n-dimensional matrix A satisfies its own characteristic equation (10.21), and show that A^n can be expanded as a polynomial in A of order less than n. (The *Cayley-Hamilton* theorem states that this is true for any matrix.)

3. *Variational Formulation of the Eigenvalue Problem for a Bounded Hermitian Operator.*

In the last section, we treated the eigenvalue problem for a normal operator as a problem in linear algebra. The characteristic equation provides a means of calculating eigenvalues to any desired approximation, but the task can be prohibitively complicated if the dimensionality n is very large. In this section, we take a

different tack and assume that A, rather than being a general normal operator, is Hermitian. This assumption covers all observables, especially Hamiltonians. In Section 8.1 the variational method was already introduced as a useful tool for estimating the low-lying eigenvalues of the Hamiltonian operator. The Rayleigh-Ritz method described in Chapter 8 links the two approaches and takes advantage of the variational principle to justify the use of an approximate characteristic equation of lower dimensionality.

We employ a variational principle by defining a *real*-valued functional

$$\lambda[\Psi] = \frac{(\Psi, A\Psi)}{(\Psi, \Psi)} \tag{10.32}$$

and look for the particular Ψ which minimizes (or maximizes) $\lambda[\Psi]$. By dividing $(\Psi, A\Psi)$ by (Ψ, Ψ) we have made λ independent of the "length" of Ψ and dependent only on its "direction." We note immediately that if Ψ is an eigenvector of A, such that

$$A\Psi = A_i'\Psi$$

then $\lambda = A_i'$. Suppose that λ has a greatest lower bound λ_0, which it assumes for the vector Ψ_0:

$$\lambda[\Psi_0] = \lambda_0 = \frac{(\Psi_0, A\Psi_0)}{(\Psi_0, \Psi_0)} \qquad (\lambda \geq \lambda_0)$$

Let us calculate λ for $\Psi = \Psi_0 \pm \varepsilon\Phi$, where ε is a small positive number and Φ is an arbitrary vector. Since λ_0 is the greatest lower bound, we have

$$\lambda[\Psi_0 \pm \varepsilon\Phi] \geq \lambda_0$$

Upon substitution, we obtain the result

$$\pm\varepsilon[(\Phi, (A - \lambda_0)\Psi_0) + (\Psi_0, (A - \lambda_0)\Phi) \pm\varepsilon(\Phi, (A - \lambda_0)\Phi)] \geq 0$$

Since $Q \equiv (\Phi, (A - \lambda_0)\Phi) \geq 0$, we find by applying the above inequality that

$$-\varepsilon Q \leq (\Phi, (A - \lambda_0)\Psi_0) + (\Psi, (A - \lambda_0)\Phi_0) \leq \varepsilon Q$$

Now let $\varepsilon \to 0$. Then

$$(\Phi, (A - \lambda_0)\Psi_0) + (\Phi, (A - \lambda_0)\Psi_0)^* = 0$$

owing to the Hermitian nature of A. Since Φ is arbitrary, we may choose it equal to $(A - \lambda_0)\Psi_0$ and thus conclude

$$((A - \lambda_0)\Psi_0, (A - \lambda_0)\Psi_0) = 0$$

which implies that

$$A\Psi_0 = \lambda_0\Psi_0$$

Thus, a vector Ψ_0 that makes λ of (10.32) a minimum is an eigenvector of A, and λ_0 is the corresponding eigenvalue. Evidently, it must be the least of all eigenvalues, allowing the identification $\lambda_0 = A_0'$ if $A_0' \leq A_1' \leq A_2' \ldots$.

We now consider a new variational problem and construct

$$\mu[\overline{\Psi}] = \frac{(\overline{\Psi}, A\overline{\Psi})}{(\overline{\Psi}, \overline{\Psi})}$$

where $\overline{\Psi}$ is the orthocomplement of Ψ_0:

$$\overline{\Psi} = \Psi - \frac{\Psi_0(\Psi_0, \Psi)}{(\Psi_0, \Psi_0)}$$

As Ψ ranges through the n-dimensional space, $\overline{\Psi}$ runs through the entire subspace of $n-1$ dimensions, orthogonal to Ψ_0. The same argument as before gives for the minimum μ_0 of μ:

$$A\overline{\Psi} = \mu_0\overline{\Psi}_0$$

and $\overline{\Psi}$ must be the eigenvector belonging to the second lowest eigenvalue, $\mu_0 = A_1'$. In this manner we may continue, and we will eventually exhaust the entire n-dimensional space after n steps. Hence, there are n orthogonal eigenvectors.

While the variational proof of the existence of eigenvalues as given here is limited to Hermitian operators, it has the merit of avoiding the explicit use of a representation. Also, since it makes no essential use of the assumption that n is finite, it can be generalized to the case $n \to \infty$. The generalization requires only that the operator A be bounded at least from one side. The operator A is said to be bounded if λ as defined in (10.32) is bounded: $+\infty > \lambda > -\infty$. From the Schwarz inequality, it follows that boundedness of A is assured if, for a given A, there exists a positive number C, independent of Ψ, such that

$$(A\Psi, A\Psi) \leq C(\Psi, \Psi)$$

for every Ψ. Many operators common in quantum mechanics, such as the energy, have only a lower bound.

4. ***Commuting Observables and Simultaneous Measurements.*** The physical meaning of Hermitian operators as candidates for representing observables motivates us to use basis vectors that are eigenvectors of Hermitian operators. Ideally, we would like to identify all basis vectors by the *eigenvalues* of the observable that supports a particular basis, and the bra-ket notation was designed with that objective in mind.

Because of the occurrence of repeated eigenvalues, the eigenvalues of a *single* observable A are usually not enough to characterize a basis unambiguously, and we must resort to additional labels or "quantum numbers" to distinguish from each other the different orthonormal basis vectors that correspond to a particular eigenvalue A_i'. The presence of repeated eigenvalues for a physical observable, selected because of its relevance to the system under study, can usually be attributed to some symmetry properties of the system. The example of the Hamiltonian of the free particle in three-dimensional space in Section 4.4 is a case in point. Owing to the translational symmetry of the Hamiltonian, the energy eigenvalue $E \geq 0$ alone does not suffice to identify an energy eigenstate.

We look for additional observables that share *common* or *simultaneous* eigenvectors with the operator A, but that are sufficiently different from A so that their eigenvalues can serve as distinguishing indices for the eigenvectors belonging to repeated eigenvalues of A. For the free particle Hamiltonian, the direction of linear momentum can serve as an observable that complements the characterization of the degenerate energy eigenstates, but other choices are possible. In the next chapter we

will see that for any system with rotational symmetry, of which the free particle is a special case, the angular momentum is the additional observable of choice.

Confining our attention first to a single operator B in addition to A, we ask under what conditions two observables A and B possess a complete set of common eigenvectors. Such eigenvectors would then represent states in which definite sharp values A'_i and B'_j can be assigned *simultaneously* to two observables. In an often-used but opaque terminology, A and B are said to be *simultaneously measurable* or *compatible* observables.

Mathematically, we require that there be a complete set of states $|A_iB_j\rangle$ such that

$$A|A'_iB'_j\rangle = A'_i|A'_iB'_j\rangle \tag{10.33}$$

and

$$B|A'_iB'_j\rangle = B'_j|A'_iB'_j\rangle \tag{10.34}$$

both hold. If (10.33) is multiplied by B and (10.34) by A, we obtain by subtraction

$$(AB - BA)|A'_iB'_j\rangle = 0$$

If this is to be true for all members of the complete set, $AB - BA$ must be equal to the null operator, or

$$\boxed{[A, B] = AB - BA = 0} \tag{10.35}$$

Hence, a necessary condition for two observables to be *simultaneously measurable* is that they *commute*.

We emphasize that the commutivity is a necessary condition only if *all* the eigenvectors of A are also to be eigenvectors of B. A limited number of eigenvectors may be common to both A and B even if the two operators do not commute. (An example is the state of zero angular momentum, which in Chapter 11 will be shown to be a common eigenstate of the noncommuting operators L_x, L_y, and L_z.)

The commutation relation (10.35) for A and B is not only *necessary*, but it is also *sufficient* for the existence of a common set of eigenvectors for the two operators. To show this, we consider a particular repeated eigenvalue A'_i of A and its associated eigenvectors, which we assume to be r in number. We denote the corresponding eigenvectors by $|A'_i, \kappa\rangle$, letting κ serve as an index $\kappa = 1 \ldots r$, which differentiates between the r orthogonal eigenvectors belonging to the same repeated eigenvalue A'_i,

$$A|A'_i, \kappa\rangle = A'_i|A'_i, \kappa\rangle \tag{10.36}$$

If B commutes with A, it follows that

$$BA|A'_i, \kappa\rangle = AB|A'_i, \kappa\rangle = A'_iB|A'_i, \kappa\rangle$$

This equation shows that $B|A'_i, \kappa\rangle$ is also an eigenvector of A with eigenvalue A'_i; hence,

$$B|A'_i, \kappa\rangle = \sum_{\lambda=1}^{r} |A'_i, \lambda\rangle\langle A'_i, \lambda|B|A'_i, \kappa\rangle \tag{10.37}$$

Exercise 10.6. Show that if A and B commute, B has no nonvanishing matrix element between eigenstates corresponding to different eigenvalues of A.

In the r-dimensional subspace of the eigenvectors belonging to A_i' a change of basis may now be effected in order to construct a new set of eigenvectors of A, which are simultaneously also eigenvectors of B. They are the kets designated as

$$|A_i'B_j'\rangle = \sum_{\kappa=1}^{r} |A_i', \kappa\rangle S_{\kappa j} \tag{10.38}$$

Here the coefficients $S_{\kappa j}$ are defined by the conditions

$$\begin{aligned} B|A_i'B_j'\rangle &= \sum_{\kappa=1}^{r}\sum_{\lambda=1}^{r} |A_i', \lambda\rangle\langle A_i', \lambda|B|A_i', \kappa\rangle S_{\kappa j} \\ &= B_j'|A_i'B_j'\rangle = B_j' \sum_{\kappa=1}^{r}\sum_{\lambda=1}^{r} S_{\kappa j}\delta_{\kappa\lambda}|A_i', \lambda\rangle \end{aligned} \tag{10.39}$$

From (10.39) we obtain r linear homogeneous equations for the r unknown transformation coefficients $S_{\kappa j}$:

$$\sum_{\lambda=1}^{r} (\langle A_i', \lambda|B|A_i', \kappa\rangle - B_j'\delta_{\lambda\kappa})S_{kj} = 0 \qquad (j \text{ fixed}, \kappa = 1, \ldots, r) \tag{10.40}$$

This system of equations possesses nontrivial solutions only if the determinant of the coefficients vanishes:

$$\det(\langle A_i', \lambda|B|A_i', \kappa\rangle - B_j'\delta_{\lambda\kappa}) = 0 \tag{10.41}$$

As in Section 10.2, the r roots of this characteristic equation give us the eigenvalues $B_1', B_2', \ldots, B_r'$. Equations (10.40) can then be used to calculate the transformation coefficients $S_{\kappa j}$. The new vectors $|A_i'B_j'\rangle$ are the required simultaneous eigenvectors of both operators A and B.

If among the r eigenvalues B_j' ($j = 1, \ldots, r$) there remain any that are repeated, then a further index may be used to distinguish those eigenvectors that have the same values A_i' and B_j' in common. One then continues the procedure of choosing additional Hermitian operators $C, D, \ldots$, all commuting with A and B as well as with each other. If the choice is made intelligently, it will eventually be possible to characterize all n basis vectors in the space by addresses composed of sets of eigenvalues $A_i'\ B_j', C_k', \ldots$.

If we can find a set of commuting Hermitian operators $A, B, C, \ldots$, whose n common eigenvectors can be characterized completely by the eigenvalues $A_i', B_j', C_k', \ldots$, such that no two eigenvectors have exactly identical eigenvalue addresses, this set of operators is said to be *complete*. We assume that

$$\langle A_i'\ B_j'\ C_k' \ldots|A_p'\ B_q'\ C_s' \ldots\rangle = \delta_{ip}\delta_{jq}\delta_{ks} \ldots \tag{10.42}$$

but often we write for this simply the compact orthonormality condition

$$\langle K_\ell|K_m\rangle = \delta_{\ell m} \tag{10.43}$$

The operator K here symbolizes the complete set $A, B, C, \ldots$, and K_ℓ (omitting the prime for simplicity) is a symbol for the set of eigenvalues $A_i', B_j', C_k', \ldots$. In particular, $K_\ell = K_m$ means that in (10.42) each pair of eigenvalues satisfies the equalities: $A_i' = A_p', B_j' = B_q', C_k' = C_s', \ldots$. In the rare case that all eigenvalues of A are simple (no repeated eigenvalues), A alone constitutes the complete set of operators $K = A$. This simple situation was implicitly assumed to hold when we introduced bras and kets in Section 9.4.

In letting K symbolize the entire complete set of commuting operators, care must be taken to interpret sums over all sets of eigenvalues of K properly. For example, the completeness of the eigenvectors is expressed as the closure relation

$$\sum_{\ell} P_{\ell} = \sum_{\ell} |K_{\ell}\rangle\langle K_{\ell}| = 1 \tag{10.44}$$

where

$$P_{\ell} = |K_{\ell}\rangle\langle K_{\ell}| = |A'_i\ B'_j\ C'_k \ldots\rangle\langle A'_i\ B'_j\ C'_k \ldots| \tag{10.45}$$

is the projection operator for K_{ℓ}. These equations can be used to reformulate the eigenvalue problem for one of the operators in the set symbolized by K. For example, we have

$$A|K_{\ell}\rangle = A'_{\ell}|K_{\ell}\rangle$$

and hence, using the closure relation:

$$A = \sum_{\ell} A|K_{\ell}\rangle\langle K_{\ell}| = \sum_{\ell} A'_i|K_{\ell}\rangle\langle K_{\ell}| \tag{10.46}$$

The eigenvalue problem for A can thus be expressed as follows:

> *Given a Hermitian (or more generally, a normal) operator A, decompose the space into a complete set of orthonormal vectors* $|K_{\ell}\rangle$ *such that A is a linear combination of the projection operators* $|K_{\ell}\rangle\langle K_{\ell}|$. *The coefficients in this expansion are the eigenvalues of A.*

If the partial sum of all projection operators, which correspond to the same eigenvalue A'_i, is denoted by

$$P_{A'_i} = \sum_{K_{\ell} \in A'_i} |K_{\ell}\rangle\langle K_{\ell}| \tag{10.47}$$

we may write

$$A = \sum_i A'_i P_{A'_i}, \qquad \sum_i P_{A'_i} = 1 \tag{10.48}$$

The sums in (10.48) extend over all *distinct* eigenvalues of A. Equations (10.48) define what is called the *spectral decomposition* of the Hermitian operator A. This form of the problem is convenient because the operators $P_{A'_i}$ are unique, whereas the eigenvectors belonging to a repeated eigenvalue A'_i contain an element of arbitrariness. A one-dimensional projection operator like (10.45) is said to have "rank one" to distinguish it from higher rank projection operators like (10.47).

For a function $f(A)$ we may write

$$f(A) = \sum_{\ell} f(A'_i)|K_{\ell}\rangle\langle K_{\ell}| = \sum_i f(A'_i)P_{A'_i} \tag{10.49}$$

Since this sum extends only over distinct eigenvalues of A, this is the same equation as (10.31).[1]

5. ***The Heisenberg Uncertainty Relations.*** We have seen that only commuting observables can in principle be measured and specified with perfect precision simultaneously. If A and B are two Hermitian operators that do *not* commute, the physical quantities A and B cannot both be sharply defined simultaneously. This

[1]For a discussion of functions of operators and matrices, see Merzbacher (1968).

suggests that the degree to which the *commutator* $[A, B] \equiv AB - BA$ of A and B is different from zero may give us information about the inevitable lack of precision in simultaneously specifying both of these observables. We define

$$\boxed{AB - BA = iC} \tag{10.50}$$

The imaginary unit has been introduced to ensure that C is a Hermitian operator.

Exercise 10.7. Prove that C is an Hermitian operator.

The uncertainty ΔA in an observable A was first introduced qualitatively and approximately in Section 2.2. To make it precise, we now define the *uncertainty* ΔA in A as the positive square root of the variance,

$$(\Delta A)^2 = \langle (A - \langle A \rangle)^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2 \tag{10.51}$$

and similarly, for B,

$$(\Delta B)^2 = \langle (B - \langle B \rangle)^2 \rangle = \langle B^2 \rangle - \langle B \rangle^2 \tag{10.52}$$

For these quantities we will prove that

$$(\Delta A)^2 (\Delta B)^2 \geq \left(\frac{1}{2} \langle C \rangle \right)^2 \tag{10.53}$$

or

$$\boxed{\Delta A \Delta B \geq \frac{1}{2} |\langle C \rangle|} \tag{10.54}$$

Proof. Since A is Hermitian, we can write

$$(\Delta A)^2 = ((A - \langle A \rangle)\Psi, (A - \langle A \rangle)\Psi)$$

if Ψ is the state vector of the system. Similarly,

$$(\Delta B)^2 = ((B - \langle B \rangle)\Psi, (B - \langle B \rangle)\Psi)$$

We now apply the Schwarz inequality (9.75) by making the identifications

$$\Psi_a = (A - \langle A \rangle)\Psi \quad \text{and} \quad \Psi_b = (B - \langle B \rangle)\Psi$$

and get

$$(\Delta A)^2 (\Delta B)^2 \geq |(\Psi, (A - \langle A \rangle)(B - \langle B \rangle)\Psi)|^2 \tag{10.55}$$

The equality sign holds if and only if

$$(B - \langle B \rangle)\Psi = \lambda (A - \langle A \rangle)\Psi \tag{10.56}$$

Now we use the simple identity, based on the decomposition (9.84),

$$\begin{aligned} (A - \langle A \rangle)(B - \langle B \rangle) &= \frac{(A - \langle A \rangle)(B - \langle B \rangle) + (B - \langle B \rangle)(A - \langle A \rangle)}{2} + \frac{i}{2} C \\ &= F + \frac{i}{2} C \end{aligned} \tag{10.57}$$

by which the operator on the left is written as a linear combination of two *Hermitian* operators F and C. Since their expectation values are real, we can write (10.55) as

$$(\Delta A)^2(\Delta B)^2 \geq \langle F\rangle^2 + \frac{1}{4}\langle C\rangle^2 \geq \frac{1}{4}\langle C\rangle^2$$

which proves the theorem (10.53). The last equality holds if and only if

$$\langle F\rangle = 0 \tag{10.58}$$

It is of interest to study the particular state Ψ for which (10.53) becomes an equality:

$$\boxed{\Delta A \Delta B = \frac{1}{2}|\langle C\rangle|}$$

Such a state obeys the conditions (10.56) and (10.58). From (10.56) we can obtain two simple relations:

$$(\Psi, (A-\langle A\rangle)(B-\langle B\rangle)\Psi) = \lambda(\Delta A)^2$$
$$(\Psi, (B-\langle B\rangle)(A-\langle A\rangle)\Psi) = \frac{1}{\lambda}(\Delta B)^2$$

Adding the left-hand sides yields $2\langle F\rangle$; hence by (10.58)

$$\lambda(\Delta A)^2 + \frac{1}{\lambda}(\Delta B)^2 = 0 \tag{10.59}$$

Subtracting the left-hand sides gives $i\langle C\rangle$; hence

$$\lambda(\Delta A)^2 - \frac{1}{\lambda}(\Delta B)^2 = i\langle C\rangle \tag{10.60}$$

From (10.59) and (10.60) we obtain

$$\lambda = \frac{i\langle C\rangle}{2(\Delta A)^2} \tag{10.61}$$

As a special case, let $A = x$, $B = p_x$. Then, as in (3.47),

$$xp_x - p_x x = i\hbar 1 \tag{10.62}$$

and hence $C=\hbar 1$. The right-hand side of (10.54) is independent of the state Ψ in this case, and we conclude that for *all* states

$$\boxed{\Delta x \Delta p_x \geq \frac{\hbar}{2}} \tag{10.63}$$

making specific and quantitative the somewhat vague statements of Chapter 2 and elsewhere. The Heisenberg uncertainty relation is thus seen to be a direct consequence of the noncommutivity of the position and momentum operators.

Applied to $A = x$, $B = p_x$, in the coordinate representation the equation (10.56) becomes a differential equation for the wave function ψ representing the special states Ψ that make (10.63) an equality:

$$\Delta x \Delta p_x = \frac{1}{2}\hbar \tag{10.64}$$

The equation for ψ is

$$\left(\frac{\hbar}{i}\frac{d}{dx} - \langle p_x\rangle\right)\psi = \frac{i\hbar}{2(\Delta x)^2}(x - \langle x\rangle)\psi \tag{10.65}$$

and has the normalized solution

$$\psi(x) = [2\pi(\Delta x)^2]^{-1/4}\exp\left[-\frac{(x - \langle x\rangle)^2}{4(\Delta x)^2} + \frac{i\langle p_x\rangle x}{\hbar}\right] \tag{10.66}$$

Known somewhat imprecisely as a *minimum uncertainty wave packet*, this state represents a plane wave that is modulated by a Gaussian amplitude function. Since λ is imaginary, the expression (10.66) is, according to (10.56), an eigenfunction of the non-Hermitian operator $p_x - \lambda x$. As such, the state represented by (10.66) is known as a *coherent state*, with properties that will be discussed in Section 10.7.

Exercise 10.8. Relate Δx to the mass and frequency of the harmonic oscillator, of which (10.66), with particular values of $\langle x\rangle$ and $\langle p_x\rangle$, is the ground state wave function.

6. ***The Harmonic Oscillator.*** Although the harmonic oscillator has been discussed in detail in Chapter 5, it is instructive to return to it and to treat it here as an application of algebraic operator techniques. Instead of using the more traditional tools of differential equations, special functions, and integral representations, these methods exploit the commutation relations among operators and operator identities shown in Chapter 4.

The generalized linear harmonic oscillator is a system whose Hamiltonian, expressed in terms of two canonical observables p and q, is given by

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2 = \frac{m\omega^2}{2}\left(q + i\frac{p}{m\omega}\right)\left(q - i\frac{p}{m\omega}\right) - \frac{\hbar\omega}{2} \tag{10.67}$$

where the Hermitian operators p and q satisfy the canonical commutation relation,

$$qp - pq = i\hbar 1 \tag{10.68}$$

Both p and q have continuous spectra extending from $-\infty$ to $+\infty$.

We first consider the eigenvalue problem of H, because it will give us the important stationary states of the system. It is convenient to introduce a new operator

$$\boxed{a = \sqrt{\frac{m\omega}{2\hbar}}\left(q + i\frac{p}{m\omega}\right)} \tag{10.69}$$

which is not Hermitian. By use of the commutation relation, we prove easily that

$$H = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right) = \hbar\omega\left(aa^\dagger - \frac{1}{2}\right) \tag{10.70}$$

where $a^\dagger$ is the Hermitian adjoint of a:

$$a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\left(q - i\frac{p}{m\omega}\right) \tag{10.71}$$

The commutator of a and $a^\dagger$ is

$$\boxed{aa^\dagger - a^\dagger a = 1} \tag{10.72}$$

which shows that a is not even normal. Since by (10.70) H is a linear function of $a^\dagger a$, the eigenvectors of H and of $a^\dagger a$ are the same, and it is sufficient to solve the eigenvalue problem for $a^\dagger a$. Expressing the eigenvalues as λ_n ($n = 0, 1, 2, \ldots$) and the corresponding eigenvectors by

$$\Psi_n = |n\rangle \tag{10.73}$$

we have

$$a^\dagger a \Psi_n = \lambda_n \Psi_n \tag{10.74}$$

or

$$a^\dagger a|n\rangle = \lambda_n |n\rangle \tag{10.75}$$

This is the equation that we must solve.

First, we prove that $\lambda_n \geq 0$. From (10.74) we get

$$(\Psi_n, a^\dagger a \Psi_n) = \lambda_n (\Psi_n, \Psi_n) \quad \text{or} \quad (a\Psi_n, a\Psi_n) = \lambda_n (\Psi_n, \Psi_n)$$

Since (9.35) holds for all vectors, we conclude that

$$\lambda_n \geq 0 \tag{10.76}$$

If $|n\rangle$ is an eigenvector of $a^\dagger a$, then $a^\dagger|n\rangle$ is also an eigenvector, as can be seen from the equation

$$(a^\dagger a)a^\dagger|n\rangle = a^\dagger(a^\dagger a + 1)|n\rangle = (\lambda_n + 1)a^\dagger|n\rangle$$

where the commutation relation (10.72) has been used. Hence, $a^\dagger|n\rangle$ is an eigenvector of $a^\dagger a$, with eigenvalue $\lambda_n + 1$. Similarly, we can show that $a|n\rangle$ is also an eigenvector of $a^\dagger a$ with eigenvalue $\lambda_n - 1$. These properties justify the designation of $a^\dagger$ as the *raising operator* and a as the *lowering operator*. By applying these operators repeatedly, we can generate from any given eigenvector $|n\rangle$ new eigenvectors with different eigenvalues by what is graphically called a *ladder method*. However, condition (10.76) limits the number of times a lowering operator can be applied. When by successive downward steps an eigenvalue between 0 and 1 has been reached, by applying a again we do not obtain a new eigenvector, because that would be an eigenvector whose eigenvalue violates the restriction (10.76). Since we have arbitrarily (but conveniently) labeled the lowest step in the ladder by setting $n = 0$, we obtain

$$a^\dagger a|0\rangle = \lambda_0|0\rangle, \qquad 1 > \lambda_0 \geq 0$$

and

$$a|0\rangle = 0 \tag{10.77}$$

Consequently,

$$\lambda_0 = 0 \tag{10.78}$$

and this is the only eigenvalue below unity.

Starting from $|0\rangle$, we obtain all other eigenvectors and eigenvalues by repeated application of the raising operator $a^\dagger$. The eigenvalues increase in unit steps. Hence,

$$|n\rangle = N_n(a^\dagger)^n|0\rangle \qquad (n = 0, 1, 2, \ldots) \tag{10.79}$$

and

$$\lambda_n = n \tag{10.80}$$

The normalization constant N_n must yet be determined. There is no degeneracy as long as no dynamical variables other than p and q appear to characterize the system. The set of eigenvectors obtained is complete. Combining (10.70), (10.75), and (10.80), we obtain

$$H|n\rangle = \hbar\omega\left(n + \frac{1}{2}\right)|n\rangle \tag{10.81}$$

Hence, the eigenvalues of the Hamiltonian are

$$\boxed{E_n = \hbar\omega\left(n + \frac{1}{2}\right)} \tag{10.82}$$

in agreement with the discrete energy eigenvalues found in Chapter 5.

Exercise 10.9. What are the eigenvalues of the kinetic and the potential energy operators of the harmonic oscillator? Explain why these don't add up to the (discrete) eigenvalues of the total energy.

Since its eigenvalues are all the nonnegative integers, the operator $N = a^\dagger a$ plays a central role when the number of (identical) particles is an observable of a system, and it is then called the *number operator*. The notion of a particle as an excitation of an oscillator-like dynamical system (such as an electromagnetic field or an elastically vibrating body) has been at the core of quantum physics from the beginning. An excited energy level of the harmonic oscillator with quantum number n is interpreted as corresponding to the presence of n particles or quasiparticles, each carrying energy $\hbar\omega$. These particles or quanta are named *phonons*, *excitons*, *photons*, and so on, depending on the physical context in which the system is represented by a harmonic oscillator. The eigenstate $|0\rangle$, which must not be confused with the null vector $|\mathbf{0}\rangle$, is variously known as the *ground state*, the *vacuum state*, or the *no-particle state* of the system. For more detail about the quantum mechanics of identical particles, see Chapter 21.

The ladder property of the lowering and raising operators a and $a^\dagger$, and the orthonormality of the states $|n\rangle$, leads us to conclude that the matrix elements of a and $a^\dagger$ connect only neighboring basis states:

$$\langle n - 1|a|n\rangle = C_n \tag{10.83}$$

$$\langle n + 1|a^\dagger|n\rangle = \langle n|a|n + 1\rangle^* = C^*_{n+1} \tag{10.84}$$

To evaluate C_n we may use the closure relation

$$\sum_{n'=0}^{\infty} |n'\rangle\langle n'| = 1 \tag{10.85}$$

as follows:

$$|C_n|^2 = \langle n-1|a|n\rangle^*\langle n-1|a|n\rangle = \langle n|a^\dagger|n-1\rangle\langle n-1|a|n\rangle$$
$$= \sum_{n'} \langle n|a^\dagger|n'\rangle\langle n'|a|n\rangle = \langle n|a^\dagger a|n\rangle = n$$

Thus,

$$C_n = \sqrt{n}\, e^{i\alpha_n} \tag{10.86}$$

Since there is no other restriction on the matrix elements, $\alpha_n = 0$ for all n is a possible and consistent choice for the phase. We may therefore write

$$a|n\rangle = |n-1\rangle\langle n-1|a|n\rangle = \sqrt{n}|n-1\rangle \tag{10.87a}$$

and

$$a^\dagger|n\rangle = |n+1\rangle\langle n+1|a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \tag{10.87b}$$

From here it follows that the normalized eigenkets of $a^\dagger a$ are

$$\boxed{\Psi_n = |n\rangle = (n!)^{-1/2}(a^\dagger)^n\Psi_0 = (n!)^{-1/2}(a^\dagger)^n|0\rangle} \tag{10.88}$$

In the representation spanned by the basis vectors $|n\rangle$, the matrix H is diagonal and given by

$$H = \frac{\hbar\omega}{2}\begin{pmatrix} 1 & 0 & 0 & \cdot & \cdot \\ 0 & 3 & 0 & \cdot & \cdot \\ 0 & 0 & 5 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & & & \cdot & \cdot \end{pmatrix} \tag{10.89}$$

The matrices representing a and $a^\dagger$ are

$$a = \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \cdot & \cdot \\ 0 & 0 & \sqrt{2} & 0 & \cdot & \cdot \\ 0 & 0 & 0 & \sqrt{3} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \tag{10.90a}$$

$$a^\dagger = \begin{pmatrix} 0 & 0 & 0 & \cdot & \cdot \\ \sqrt{1} & 0 & 0 & \cdot & \cdot \\ 0 & \sqrt{2} & 0 & \cdot & \cdot \\ 0 & 0 & \sqrt{3} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \tag{10.90b}$$

Exercise 10.10. Using the matrices (10.90a) and (10.90b), verify the commutation relation $aa^\dagger - a^\dagger a = 1$.

The coordinate q of the oscillator can be expressed as

$$q = \sqrt{\frac{\hbar}{2m\omega}}\,(a + a^\dagger) \tag{10.91}$$

and its matrix in the $|n\rangle$ basis is

$$q = \sqrt{\frac{\hbar}{2m\omega}} \begin{pmatrix} 0 & \sqrt{1} & 0 & \cdot & \cdot \\ \sqrt{1} & 0 & \sqrt{2} & \cdot & \cdot \\ 0 & \sqrt{2} & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \tag{10.92}$$

An eigenstate of the coordinate q with eigenvalue q' is represented by a column matrix, and the eigenvalue equation for q appears in the form

$$\sqrt{\frac{\hbar}{2m\omega}} \begin{pmatrix} 0 & \sqrt{1} & 0 & \cdot & \cdot \\ \sqrt{1} & 0 & \sqrt{2} & \cdot & \cdot \\ 0 & \sqrt{2} & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ \cdot \\ \cdot \end{pmatrix} = q' \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ \cdot \\ \cdot \end{pmatrix} \tag{10.93}$$

where the components of the eigenvector of q are the transformation coefficients

$$c_n = \langle n|q'\rangle \tag{10.94}$$

Equation (10.93) leads to a set of simultaneous linear equations:

$$\begin{aligned} \sqrt{1}\, c_1 &= \sqrt{\frac{2m\omega}{\hbar}}\, q' c_0 \\ \sqrt{2}\, c_2 + \sqrt{1}\, c_0 &= \sqrt{\frac{2m\omega}{\hbar}}\, q' c_1 \\ \cdots\cdots\cdots &= \cdots\cdots\cdots \\ \sqrt{n+1}\, c_{n+1} + \sqrt{n}\, c_{n-1} &= \sqrt{\frac{2m\omega}{\hbar}}\, q' c_n \\ \cdots\cdots\cdots &= \cdots\cdots\cdots \end{aligned}$$

These simultaneous equations are solved by

$$c_n(q') = 2^{-(n/2)}(n!)^{-1/2} H_n\left(\sqrt{\frac{m\omega}{\hbar}}\, q'\right) c_0 \tag{10.95}$$

by virtue of the recurrence relation,

$$H_{n+1}(x) - 2xH_n(x) + 2nH_{n-1}(x) = 0$$

for Hermite polynomials.

The closure condition (10.85), represented in the form

$$\sum_{n=0}^{\infty} \langle q'|n\rangle\langle n|q''\rangle = \delta(q' - q'')$$

finally determines the constant factor in (10.95). The result is

$$\langle n|q'\rangle = 2^{-(n/2)}(n!)^{-1/2}\left(\frac{m\omega}{\hbar\pi}\right)^{1/4} \exp\left(-\frac{m\omega}{2\hbar}\, q'^2\right) H_n\left(\sqrt{\frac{m\omega}{\hbar}}\, q'\right) \tag{10.96}$$

in agreement with Eq. (5.39)

Exercise 10.11. Verify the recurrence relation for Hermite polynomials from the formula (5.35).

Exercise 10.12. Transcribe Eqs. (10.77) and (10.88) in the coordinate (q) representation and calculate $\langle q' | n \rangle$ from these differential relations. Using the mathematical tools of Section 5.3, verify Eq. (10.96).

7. *Coherent and Squeezed States.* The general state of an harmonic oscillator can be expressed as a superposition of the energy eigenstates $|n\rangle$. A class of states that is of particular importance consists of the eigenstates of the non-Hermitian lowering operator a, with eigenvalue α

$$\boxed{a|\alpha\rangle = \alpha|\alpha\rangle} \tag{10.97}$$

A trivial solution of this equation is the ground state $|0\rangle$ for $\alpha = 0$, as seen from (10.77).

The unitary shifting or *displacement* operator

$$\boxed{D_\alpha = e^{\alpha a^\dagger - \alpha^* a}} \tag{10.98}$$

causes a shift of the operator a, since from Eq. (3.58) we see that

$$D_\alpha^\dagger a D_\alpha = e^{-\alpha a^\dagger + \alpha^* a} a e^{\alpha a^\dagger - \alpha^* a} = a + \alpha \tag{10.99}$$

for an arbitrary complex number α. We deduce from (10.98) and (10.99) that D_α has the properties

$$D_\alpha^\dagger = D_\alpha^{-1} = D_{-\alpha} \tag{10.100}$$

and

$$a D_\alpha |0\rangle = \alpha D_\alpha |0\rangle$$

This result shows that the solution of the eigenvalue problem (10.97) may be taken to be

$$|\alpha\rangle = D_\alpha |0\rangle \tag{10.101}$$

and that *all complex numbers* are eigenvalues of the operator a. Since D_α is unitary, the eigenket $|\alpha\rangle$ in (10.101) is normalized,

$$\langle \alpha | \alpha \rangle = \langle 0 | 0 \rangle = 1 \tag{10.102}$$

Using (3.61) and (10.77), this eigenket can be expressed as

$$\boxed{|\alpha\rangle = e^{\alpha a^\dagger - \alpha^* a}|0\rangle = e^{-|\alpha|^2/2} e^{\alpha a^\dagger}|0\rangle} \tag{10.103}$$

These eigenstates of the lowering (annihilation) operator a are known as *coherent* states. Their relation to the minimum uncertainty wave packets (10.66) will be brought out shortly. For some purposes it is instructive to depict the eigenvalue α of a coherent state as a vector in the two-dimensional complex plane (see Argand diagram, Figure 10.1). It is interesting to note that the coherent states are normalized to unity, even though the eigenvalue spectrum of the operator a is continuous.[2]

[2]The first comprehensive treatment of coherent states was given by Glauber (1965).

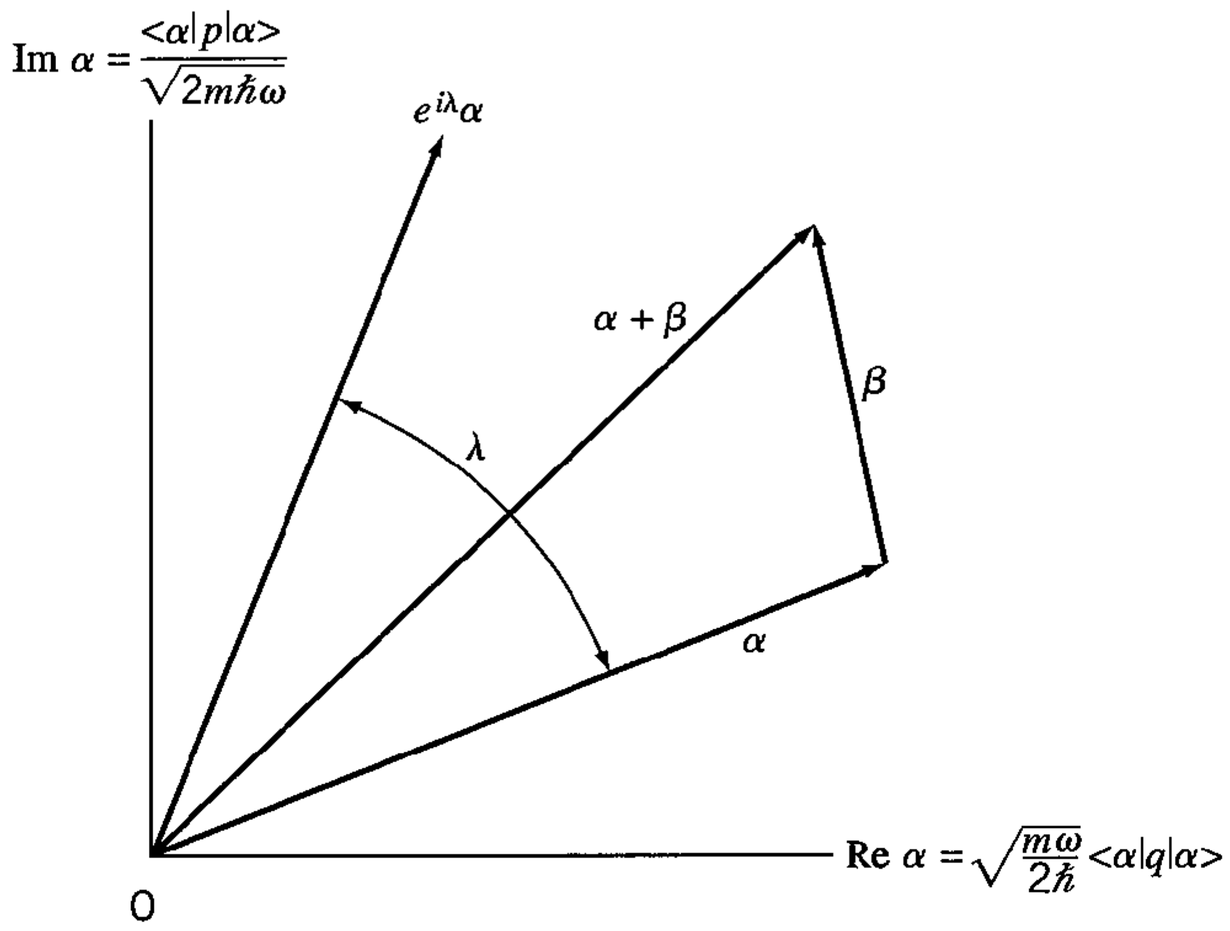


Figure 10.1. Two-dimensional phase space or Argand diagram representing a coherent state $|\alpha\rangle$ in terms of the eigenvalue α of the lowering or annihilation operator a. Also shown is the effect of a "displacement" D_β and a "rotation" R_λ on the coherent state $|\alpha\rangle$.

Exercise 10.13. Using the property (10.99), show that for any coherent state $|\alpha\rangle$,

$$D_\beta|\alpha\rangle = C|\alpha + \beta\rangle \tag{10.104}$$

where $|\alpha + \beta\rangle$ is again a coherent state and C is a phase factor. Interpret this result in terms of the complex eigenvalue plane (Figure 10.1).

A second useful unitary operator is

$$R_\lambda = e^{i\lambda a^\dagger a} \tag{10.105}$$

with a real-valued parameter λ. Since $[a^\dagger a, a] = -a$, the identity (3.59) gives

$$\boxed{R_\lambda^\dagger\, a\, R_\lambda = e^{-i\lambda a^\dagger a}\, a\, e^{i\lambda a^\dagger a} = e^{i\lambda} a} \tag{10.106}$$

Exercise 10.14. Show that for any coherent state $|\alpha\rangle$,

$$R_\lambda|\alpha\rangle = C'|e^{i\lambda}\alpha\rangle$$

where $|e^{i\lambda}\alpha\rangle$ is again a coherent state and C' is a phase factor. Interpret the meaning of this result in the complex eigenvalue plane (Figure 10.1).

There is an eigenstate $|\alpha\rangle$ of a for any complex number α, but the coherent states do not form an orthogonal set. The inner product of two coherent states $|\alpha\rangle$ and $|\beta\rangle$ is

$$\begin{aligned}\langle\beta|\alpha\rangle &= e^{-|\alpha|^2/2-|\beta|^2/2}\langle 0|e^{\beta^* a}e^{\alpha a^\dagger}|0\rangle = e^{-|\alpha|^2/2-|\beta|^2/2}\langle 0|e^{\alpha a^+}e^{\beta^* a}e^{\alpha\beta^*}|0\rangle \\ &= e^{-|\alpha|^2/2-|\beta|^2/2+\alpha\beta^*}\end{aligned} \tag{10.107}$$

and

$$\boxed{|\langle\beta|\alpha\rangle|^2 = e^{-|\alpha\,-\,\beta|^2}} \tag{10.108}$$

Hence, the distance $|\alpha - \beta|$ in the complex eigenvalue plane (Figure 10.1) measures the degree to which the two eigenstates are approximately orthogonal.

To expand the coherent state $|\alpha\rangle$ in terms of the energy or the number-operator eigenstates $|n\rangle$, we calculate

$$\begin{aligned} |\alpha\rangle &= \sum_n |n\rangle\langle n|\alpha\rangle = \sum_n |n\rangle\langle n|e^{-|\alpha|^2/2} e^{\alpha a^\dagger}|0\rangle \\ &= \sum_{n=0}^{\infty} |n\rangle e^{-|\alpha|^2/2}\langle n| \sum_{k=0}^{\infty} \frac{(\alpha a^\dagger)^k}{k!}|0\rangle = \sum_{n=0}^{\infty} |n\rangle e^{-|\alpha|^2/2} \frac{\alpha^n}{\sqrt{n!}} \end{aligned} \tag{10.109}$$

In the last step, Eq. (10.88) and the orthonormality of the energy eigenstates have been used. The probability $P_n(\alpha)$ of finding the coherent state $|\alpha\rangle$ to have the value n when the operator $a^\dagger a$ is measured is thus given by the *Poisson distribution*:

$$\boxed{P_n(\alpha) = e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!}} \tag{10.110}$$

The mean (expectation) value of n for this distribution is $|\alpha|^2$.

Exercise 10.15. Evaluate the integral

$$\int |\langle\beta|\alpha\rangle|^2 \, d^2\beta$$

over the entire complex β plane, and interpret the result. How can this be reconciled with the probability doctrine of quantum mechanics? (See Section 15.5.)

Exercise 10.16. By requiring that $|\alpha\rangle = \sum_n |n\rangle\langle n|\alpha\rangle$ is an eigenket of the operator a, with eigenvalue α, obtain a recurrence relation for $\langle n|\alpha\rangle$. Verify (10.109).

As we might expect from the lack of restrictions imposed on the eigenvalues and eigenstates of a, the latter form an *overcomplete* set. An arbitrary state can be expanded in terms of them in infinitely many different ways. Even so, an identity bearing a remarkable similarity to a closure relation can be proved:

$$\boxed{\frac{1}{\pi}\int |\alpha\rangle \, d^2\alpha \, \langle\alpha| = \frac{1}{\pi}\int |\alpha\rangle \, d\mathrm{Re}(\alpha) \, d\mathrm{Im}(\alpha) \, \langle\alpha| = 1} \tag{10.111}$$

Here, the integration is extended over the entire α plane with a real element of area.

Exercise 10.17. Prove Eq. (10.111). This is most easily done by expanding the coherent states in terms of the harmonic oscillator eigenstates, using (10.109) and plane polar coordinates in the complex α plane.

Exercise 10.18. Prove that the raising operator $a^\dagger$ has no normalizable eigenvectors, and explain the reason.

An arbitrary state of a system, which has the coordinate q as its complete set of dynamical variables, can be written, on account of (10.88), as

$$|\Psi\rangle = \sum_n |n\rangle\langle n|\Psi\rangle = \sum_n \frac{c_n}{\sqrt{n!}} (a^\dagger)^n |0\rangle = F(a^\dagger)|0\rangle$$

where $c_n = \langle n|\Psi\rangle$ and $F(a^\dagger)$ is a convergent power series (entire function). Hence

$$\langle\alpha|\Psi\rangle = \langle\alpha|F(a^\dagger)|0\rangle = F(\alpha^*)\langle\alpha|0\rangle = e^{-|\alpha|^2/2}F(\alpha^*) \tag{10.112}$$

The entire function $F(\alpha^*)$ thus represents the state Ψ.

Exercise 10.19. What function $F(\alpha^*)$ represents the coherent state $|\beta\rangle$?

The action of a normally ordered operator on an arbitrary state can be expressed conveniently in terms of this representation. An operator is *normally ordered* if, by use of the commutation relations, all lowering operators have been brought to the right of all raising operators. For example:

$$A = a^2a^\dagger = a(1 + a^\dagger a) = a + (1 + a^\dagger a)a = 2a + a^\dagger a^2$$

shows how normal ordering is achieved. The expectation value $\langle\alpha|A(a^\dagger, a)|\alpha\rangle$ of a normally ordered operator is

$$\langle\alpha|A(a^\dagger, a)|\alpha\rangle = A(\alpha^*, \alpha) = A(\langle a^\dagger\rangle, \langle a\rangle) \tag{10.113}$$

For example, in a coherent state $|\alpha\rangle$ the expectation value of a product of normally ordered operators, like $(a^\dagger)^n a^m$, can be factored:

$$\langle\alpha|(a^\dagger)^n a^m|\alpha\rangle = \alpha^{*n}\alpha^m = \langle\alpha|a^\dagger|\alpha\rangle^n\langle\alpha|a|\alpha\rangle^m \tag{10.114}$$

and written as a product of expectation values of $a^\dagger$ and a. In general, such factorizations are not permissible in quantum mechanics, but coherent states enjoy the unusual property of minimizing certain quantum correlations (or *quantum fluctuations*). This has led to their designation as *quasiclassical* or *semiclassical* states. The term *coherent* reflects their important role in optics and quantum electronics (Section 23.4).

Exercise 10.20. For a coherent state $|\alpha\rangle$, evaluate the expectation value of the number operator $a^\dagger a$, its square and its variance, using the commutation relation (10.72). Check the results by computing the expectation values of n, n^2, and $(\Delta n)^2$ directly from the Poisson distribution (10.110).

Consider an operator $A(a^\dagger, a)$ which is normally ordered, and let it act on an arbitrary ket

$$A(a^\dagger, a)|\Psi\rangle = A(a^\dagger, a)F(a^\dagger)|0\rangle = A\left(a^\dagger, \frac{\partial}{\partial a^\dagger}\right)F(a^\dagger)|0\rangle \tag{10.115}$$

where $\partial/\partial a^\dagger$ denotes formal differentiation. This last equation follows from the commutation relation

$$a(a^\dagger)^n - (a^\dagger)^n a = n(a^\dagger)^{n-1} = \frac{\partial}{\partial a^\dagger}(a^\dagger)^n$$

and the property $a|0\rangle = 0$. Hence, we infer that the entire function

$$A\left(\alpha^*, \frac{\partial}{\partial \alpha^*}\right)F(\alpha^*)$$

represents the state $A|\Psi\rangle$ in the same sense as, according to (10.112) $F(\alpha^*)$ represents the state $|\Psi\rangle$.

Exercise 10.21. Rederive the function $F(\alpha^*)$ which represents the coherent state $|\beta\rangle$ by letting $A = a$ and requiring $a|\beta\rangle = \beta|\beta\rangle$.

Exercise 10.22. Choose $A = a^\dagger a$, the number operator, and obtain the entire function $F(\alpha^*)$ which represents its eigenkets. Verify that the eigenvalues must be nonnegative integers.

By definition (10.69), the non-Hermitian operators a and $a^\dagger$ are related to Hermitian canonical coordinate and momentum operators as

$$q = \sqrt{\frac{\hbar}{2m\omega}}\,(a^\dagger + a), \qquad p = i\sqrt{\frac{m\hbar\omega}{2}}\,(a^\dagger - a) \tag{10.116}$$

The expectation values of q and p in a coherent state $|\alpha\rangle$ are

$$\langle\alpha|q|\alpha\rangle = \sqrt{\frac{\hbar}{2m\omega}}\,(\alpha + \alpha^*) = \sqrt{\frac{2\hbar}{m\omega}}\,\mathrm{Re}\;\alpha$$

and

$$\langle\alpha|p|\alpha\rangle = i\sqrt{\frac{\hbar m\omega}{2}}\,(\alpha^* - \alpha) = \sqrt{2\hbar m\omega}\;\mathrm{Im}\;\alpha \tag{10.117}$$

Furthermore, taking advantage of normal ordering, we can calculate

$$\langle\alpha|q^2|\alpha\rangle = \frac{\hbar}{2m\omega}\langle\alpha|a^2 + a^{\dagger 2} + a^\dagger a + 1|\alpha\rangle$$

$$= \frac{\hbar}{2m\omega}\,[(\alpha + \alpha^*)^2 + 1] = \langle\alpha|q|\alpha\rangle^2 + \frac{\hbar}{2m\omega}$$

and similarly

$$\langle\alpha|p^2|\alpha\rangle = \langle\alpha|p|\alpha\rangle^2 + \frac{\hbar m\omega}{2}$$

The last terms on the right-hand side are proportional to $\hbar$ and exhibit the quantum fluctuations in q and p. The variances are

$$(\Delta q)^2 = \langle\alpha|q^2|\alpha\rangle - \langle\alpha|q|\alpha\rangle^2 = \frac{\hbar}{2m\omega} \tag{10.118}$$

$$(\Delta p)^2 = \langle\alpha|p^2|\alpha\rangle - \langle\alpha|p|\alpha\rangle^2 = \frac{\hbar m\omega}{2} \tag{10.119}$$

so that

$$\boxed{\Delta q \cdot \Delta p = \frac{\hbar}{2}} \tag{10.120}$$

showing that in the coordinate and momentum language the coherent states are minimum uncertainty (product) states.

As discussed in Section 9.6, in the coordinate representation the eigenvalue condition

$$a|\alpha\rangle = \alpha|\alpha\rangle$$

is transcribed as

$$\left(q + \frac{\hbar}{2m\omega}\frac{\partial}{\partial q}\right)\langle q|\alpha\rangle = \sqrt{\frac{2\hbar}{m\omega}}\,\alpha\langle q|\alpha\rangle \tag{10.121}$$

This differential equation has the same form as (10.65) and the solution

$$\begin{aligned}\langle q|\alpha\rangle &= C' \exp\left[-\frac{m\omega}{2\hbar}\left(q - \sqrt{\frac{2\hbar}{m\omega}}\,\alpha\right)^2\right] \\ &= C \exp\left[-\frac{m\omega}{2\hbar}\left(q - \sqrt{\frac{2\hbar}{m\omega}}\,\mathrm{Re}\,\alpha\right)^2 + \frac{i}{\hbar}\,q\sqrt{2\hbar m\omega}\,\mathrm{Im}\,\alpha\right] \\ &= C \exp\left[-\frac{(q - \langle q\rangle^2)}{4(\Delta q)^2} + \frac{i}{\hbar}\langle p\rangle q\right]\end{aligned} \tag{10.122}$$

In this form, known to us from (10.66), coherent states have been familiar since the early days of quantum mechanics.

Exercise 10.23. Compute the normalization factors C' and C in (10.122) and show how they are related.

For a fixed oscillator mode, specified by a given value of $m\omega$, the coherent states are the manifold of those minimum uncertainty states that have definite values for Δq and Δp, given in Eqs. (10.118) and (10.119). (If $m\omega = 1$, the uncertainties in q and p are both equal to $\sqrt{\hbar/2}$.) We can construct other minimum uncertainty states with narrower width Δq, so-called *squeezed states*, for the same oscillator by defining a new set of raising and lowering operators

$$b = \sqrt{\frac{m\omega'}{2\hbar}}\left(q + i\frac{p}{m\omega'}\right) \tag{10.123}$$

$$b^\dagger = \sqrt{\frac{m\omega'}{2\hbar}}\left(q - i\frac{p}{m\omega'}\right) \tag{10.124}$$

using an arbitrarily chosen positive parameter ω'. Obviously

$$[b, b^\dagger] = I \tag{10.125}$$

The operators b, $b^\dagger$ can be expressed in terms of a and $a^\dagger$ by substituting (10.116) into (10.123) and (10.124):

$$b = \frac{1}{2}\left(\sqrt{\frac{\omega'}{\omega}} + \sqrt{\frac{\omega}{\omega'}}\right)a + \frac{1}{2}\left(\sqrt{\frac{\omega'}{\omega}} - \sqrt{\frac{\omega}{\omega'}}\right)a^\dagger \tag{10.126}$$

$$b^\dagger = \frac{1}{2}\left(\sqrt{\frac{\omega'}{\omega}} + \sqrt{\frac{\omega}{\omega'}}\right)a^\dagger + \frac{1}{2}\left(\sqrt{\frac{\omega'}{\omega}} - \sqrt{\frac{\omega}{\omega'}}\right)a$$

or

$$\begin{aligned} b &= \lambda a + \nu a^\dagger \\ b^\dagger &= \lambda a^\dagger + \nu a \end{aligned} \tag{10.127}$$

where λ and ν are two real numbers that are related by the condition

$$\lambda^2 - \nu^2 = 1 \tag{10.128}$$

(with $\lambda > 1$).

Exercise 10.24. Verify (10.128).

In the language of bosons (Chapter 21), the transformation (10.127), which generally does not preserve the number of particles (since $b^\dagger b \neq a^\dagger a$), is referred to as a *quasiparticle transformation*. The operator $b^\dagger$ creates a quasiparticle, which is a superposition of particle states, and b annihilates a quasiparticle.

If (10.127) is inverted, subject to the restriction (10.128), we have

$$\begin{aligned} a &= \lambda b - \nu b^\dagger \\ a^\dagger &= \lambda b^\dagger - \nu b \end{aligned} \tag{10.129}$$

The eigenstates of the lowering operator b are defined by

$$b|\beta\rangle = \beta|\beta\rangle \tag{10.130}$$

From the relations (10.123) and (10.124) it is apparent that these states are minimum uncertainty states for p and q, but the uncertainties of these quantities are determined by ω', and not by ω:

$$\begin{aligned} (\Delta q)^2 &= \frac{\hbar}{2m\omega'} = \frac{\hbar}{2m\omega}(\lambda - \nu)^2 \\ (\Delta p)^2 &= \frac{\hbar m\omega'}{2} = \frac{\hbar m\omega}{2}(\lambda + \nu)^2 \end{aligned} \tag{10.131}$$

so that

$$(\Delta q)^2(\Delta p)^2 = \frac{\hbar^2}{4}$$

as it should be for a minimum uncertainty state. Although ω is fixed, the uncertainty in either q or p can be controlled in these states by changes in the parameter ω'. Since it is possible, for instance, by choosing very large values $\omega' \gg \omega$ to reduce Δq arbitrarily at the expense of letting Δp grow correspondingly, these states have been named *squeezed states* in quantum optics.

Exercise 10.25. For a squeezed state $|\beta\rangle$ verify the values of Δq and Δp given in (10.131).

Exercise 10.26. Prove that the operators a and b are related by a unitary transformation

$$b = UaU^\dagger$$

where

$$U = e^{\zeta(a^2 - a^{\dagger 2})/2} \tag{10.132}$$

and $e^\zeta = \lambda + \nu$. Show that U transforms a coherent state into a squeezed state. [*Hint:* Use identity (3.54).]

Problems

1. Carry out numerical integrations to test the uncertainty relation $\Delta x \Delta k_x \geq 1/2$ for the wave packets defined by Figures 2.1 and 2.2.

2. Assuming a particle to be in one of the stationary states of an infinitely high one-dimensional box, calculate the uncertainties in position and momentum, and show that they agree with the Heisenberg uncertainty relation. Also show that in the limit of very large quantum numbers the uncertainty in x equals the root-mean-square deviation of the position of a particle moving in the enclosure classically with the same energy.
3. Calculate the value of $\Delta x \Delta p$ for a linear harmonic oscillator in its nth energy eigenstate.
4. Using the uncertainty relation, but not the explicit solutions of the eigenvalue problem, show that the expectation value of the energy of a harmonic oscillator can never be less than the zero-point energy.
5. Rederive the one-dimensional minimum uncertainty wave packet by using the variational calculus to minimize the expression $I = (\Delta x)^2(\Delta p)^2$ subject to the condition

$$\int |\psi|^2 \, dx = 1$$

Show that the solution ψ of this problem satisfies a differential equation which is equivalent to the Schrödinger equation for the harmonic oscillator, and calculate the minimum value of $\Delta x \Delta p$.

6. The Hamiltonian representing an oscillating LC circuit can be expressed as

$$H = \frac{Q^2}{2C} + \frac{\Phi^2}{2L}$$

Establish that Hamilton's equations are the correct dynamical equations for this system, and show that the charge Q and the magnetic flux Φ can be regarded as canonically conjugate variables, q, p (or the dual pair p, $-q$). Work out the Heisenberg relation for the product of the uncertainties in the current I and the voltage V. If a mesoscopic LC circuit has an effective inductance of $L = 1$ μH and an effective capacitance $C = 1$ pF, how low must the temperature of the device be before quantum fluctuations become comparable to thermal energies? Are the corresponding current-voltage tolerances in the realm of observability?

7. If a coherent state $|\alpha\rangle$ (eigenstate of a) of an oscillator is transformed into a squeezed state by the unitary operator

$$U = \exp\left[\frac{\zeta}{2}(a^2 - a^{\dagger 2})\right]$$

calculate the value of ζ that will reduce the width of the Hermitian observable $(a + a^\dagger)/\sqrt{2}$ to 1 percent of its original coherent-state value. What happens to the width of the conjugate observable $(a - a^\dagger)/\sqrt{2}i$ in this transformation?

CHAPTER 11

Angular Momentum in Quantum Mechanics

We now turn to the motion of a particle in ordinary three-dimensional space. Bohr found the key to the theory of electronic motion in the Coulomb field of the nucleus in the quantization of angular momentum (in units of Planck's constant divided by 2π). Beyond its relevance to the classification of energy levels in central-force systems, the study of orbital angular momentum brings us one step closer to a detailed exposition of symmetry in quantum mechanics (Chapter 17).

1. ***Orbital Angular Momentum.*** *Central forces* are derivable from a potential that depends only on the distance r of the moving particle from a fixed point, usually the coordinate origin. The Hamiltonian operator is

$$\boxed{H = \frac{p^2}{2m} + V(r) = -\frac{\hbar^2}{2m}\nabla^2 + V(r)} \tag{11.1}$$

Since central forces produce no torque about the origin, the orbital *angular momentum*

$$\boxed{\mathbf{L} = \mathbf{r} \times \mathbf{p}} \tag{11.2}$$

is conserved. In classical mechanics this is the statement of *Kepler's second law.*

According to the correspondence principle, we must expect angular momentum to play an equally essential role in quantum mechanics. The operator that represents angular momentum in the coordinate representation is obtained from (11.2) by replacing $\mathbf{p}$ by $(\hbar/i)\nabla$:

$$\boxed{\mathbf{L} = \mathbf{r} \times \frac{\hbar}{i}\nabla} \tag{11.3}$$

No difficulty arises here with operators that fail to commute, because only products like xp_y, yp_z appear.

In view of the great importance of angular momentum as a physical quantity, it is well to derive some of the properties of the operator $\mathbf{L}$, using the basic commutation relations between the components of $\mathbf{r}$ and $\mathbf{p}$ and the algebraic rules of Section 3.4 for commutators. For example

$$\begin{aligned} [L_x, y] &= [yp_z - zp_y, y] = -z[p_y, y] = i\hbar z \\ [L_x, p_y] &= [yp_z - zp_y, p_y] = [y, p_y]p_z = i\hbar p_z \\ [L_z, x] &= 0, \qquad [L_x, p_x] = 0 \end{aligned} \tag{11.4}$$

Similar relations hold for all other commutators between **L** and **r** and between **L** and **p**. From these relations we can further deduce the commutation relations between the various components of **L**:

$$\begin{aligned}[L_x, L_y] &= [L_x, zp_x - xp_z] = [L_x, z]p_x - x[L_x, p_z] \\ &= -i\hbar yp_x + i\hbar xp_y = i\hbar L_z\end{aligned}$$

and by cyclic permutation ($x \to y \to z \to x$) of this result,

$$\boxed{[L_x, L_y] = i\hbar L_z, \qquad [L_y, L_z] = i\hbar L_x, \qquad [L_z, L_x] = i\hbar L_y} \tag{11.5}$$

Since the components of **L** do not commute, the system cannot in general be assigned definite values for all angular momentum components simultaneously.

Insight into the nature of the angular momentum operator is gained by noting its connection with (rigid) rotations. Suppose $f(\mathbf{r})$ is an arbitrary differentiable function in space. If the value $f(\mathbf{r})$ of the function at point **r** is displaced to the new point $\mathbf{r} + \mathbf{a}$, where the displacement vector **a** may itself depend on **r**, a new function $F(\mathbf{r})$ is obtained by the mapping

$$f(\mathbf{r}) \mapsto F(\mathbf{r} + \mathbf{a}) = f(\mathbf{r})$$

For an *infinitesimal* displacement $\boldsymbol{\varepsilon}$,

$$F(\mathbf{r} + \boldsymbol{\varepsilon}) = F(\mathbf{r}) + \boldsymbol{\varepsilon} \cdot \nabla F(\mathbf{r}) = f(\mathbf{r}) \tag{11.6}$$

and the change of the function f is, to the first order,

$$\delta f(\mathbf{r}) = F(\mathbf{r}) - f(\mathbf{r}) = -\boldsymbol{\varepsilon} \cdot \nabla F(\mathbf{r}) = -\boldsymbol{\varepsilon} \cdot \nabla f(\mathbf{r}) \tag{11.7}$$

A *finite rotation* R by an angle ϕ about an axis that points in the direction of the unit vector $\hat{\mathbf{n}}$ through the origin is characterized by the displacement vector (Figure 11.1)

$$\mathbf{a}_R = \hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{r})(1 - \cos\phi) + \hat{\mathbf{n}} \times \mathbf{r} \sin\phi \tag{11.8}$$

Exercise 11.1. Verify (11.8) and show that it gives the expected answer for a rotation about the z axis.

Exercise 11.2. For an infinitesimal displacement $\boldsymbol{\varepsilon}$, applied to the vector function $f(\mathbf{r}) = \mathbf{r}$, show that $\delta f(\mathbf{r}) = \delta\mathbf{r} = -\boldsymbol{\varepsilon}$.

The *inverse rotation* R^{-1} by an angle $-\phi$ about the same axis (or, equivalently by an angle ϕ about the unit vector $-\hat{\mathbf{n}}$) is described by the displacement vector

$$\mathbf{a}_{R^{-1}} = \hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{r})(1 - \cos\phi) - \hat{\mathbf{n}} \times \mathbf{r} \sin\phi \tag{11.9}$$

Because of the **r**-dependence of the displacement, for *finite* rotations generally $\mathbf{a}_{R^{-1}} \neq -\mathbf{a}_R$. The rotation R causes the value of the function f at position $\mathbf{r} + \mathbf{a}_{R^{-1}}$ to be displaced to the original position **r**; hence,

$$F(\mathbf{r}) = f(\mathbf{r} + \mathbf{a}_{R^{-1}})$$

Exercise 11.3. If the **r** dependence of the displacement vector is explicitly indicated as $\mathbf{a}_R(\mathbf{r})$, prove that

$$\mathbf{a}_R(\mathbf{r}) + \mathbf{a}_{R^{-1}}(\mathbf{r} + \mathbf{a}_R(\mathbf{r})) = 0$$

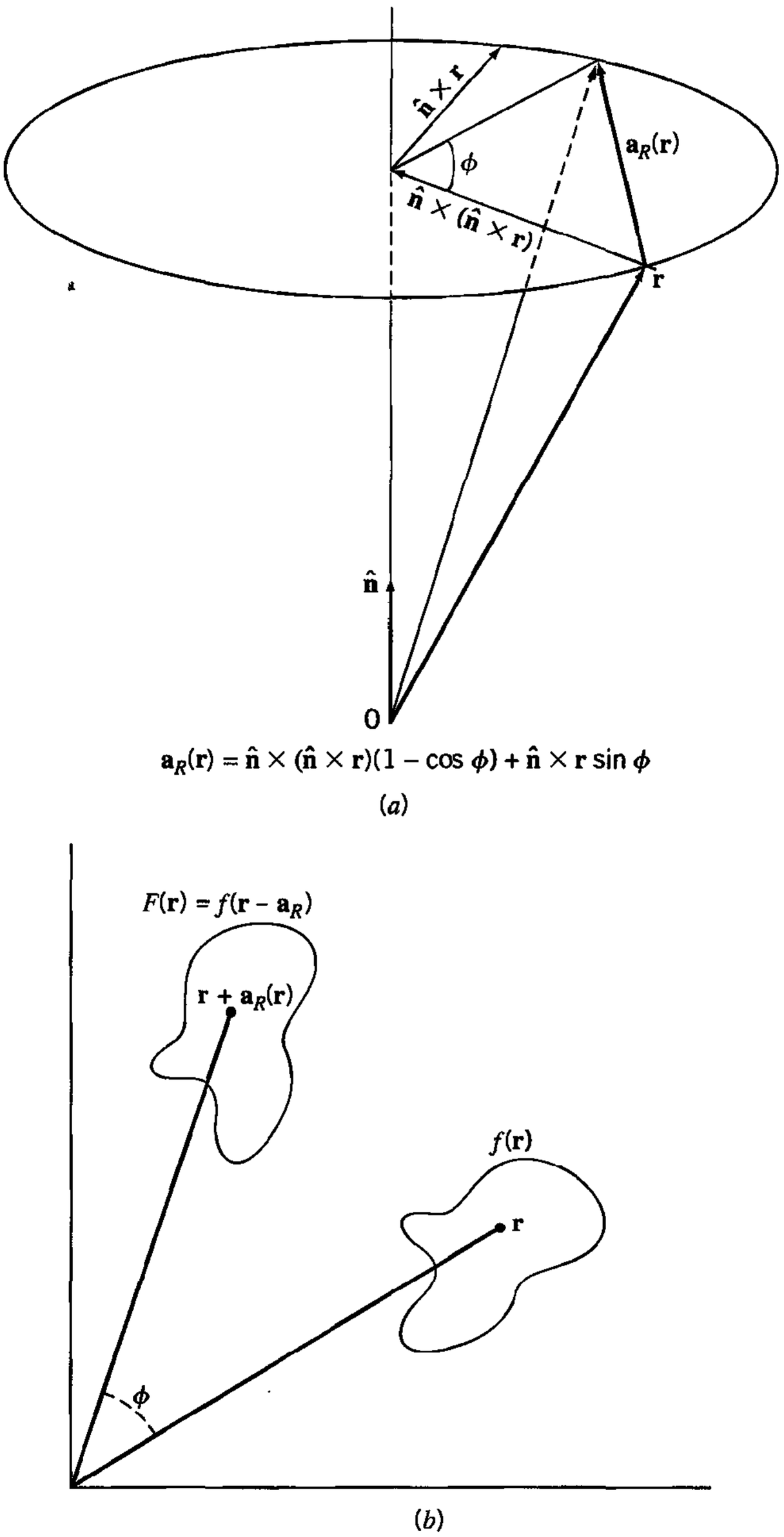


Figure 11.1. Rotation about an axis defined by the unit vector $\hat{\mathbf{n}}$ and the rotation angle ϕ. (*a*) shows the displacement $\mathbf{a}_R(\mathbf{r})$ of the point whose position vector is $\mathbf{r}$. (*b*) illustrates the active rotation of a function or state $f(\mathbf{r})$ about an axis ($\hat{\mathbf{n}}$) perpendicular to the plane of the figure: $f(\mathbf{r}) \mapsto F(\mathbf{r}) = f(\mathbf{r} - \mathbf{a}_R)$.

If the rotation angle $\delta\phi$ is infinitesimal, (11.8) shows that the displacement $\boldsymbol{\varepsilon}$ can, to second order in $\delta\phi$, be expressed as

$$\boldsymbol{\varepsilon} = \delta\phi\hat{\mathbf{n}} \times \mathbf{r} + \frac{1}{2}(\delta\phi)^2\, \mathbf{n} \times (\mathbf{n} \times \mathbf{r}) = \boldsymbol{\delta\phi} \times \mathbf{r} + \frac{1}{2}\boldsymbol{\delta\phi} \times (\boldsymbol{\delta\phi} \times \mathbf{r}) \quad (11.10)$$

where $\boldsymbol{\delta\phi} = \delta\phi\hat{\mathbf{n}}$ is a vector of length $\delta\phi$ pointing in the direction of the axis of rotation with an orientation defined by a right-handed screw. If the inverse rotation

is *infinitesimal*, the displacement is simply $-\boldsymbol{\varepsilon}$. For an infinitesimal rotation, the change of the function f is then, to first order in $\delta\phi$,

$$\delta f = -\delta\phi\, \hat{\mathbf{n}} \times \mathbf{r} \cdot \nabla f = -\delta\phi\, \hat{\mathbf{n}} \cdot \mathbf{r} \times \nabla f \tag{11.11}$$

or

$$\boxed{\delta f = -\frac{i}{\hbar}\, \boldsymbol{\delta\phi} \cdot \mathbf{L}\, f} \tag{11.12}$$

The operator $\mathbf{L}/\hbar$ is called the *generator of infinitesimal rotations*.

Equation (11.12) can be integrated for a finite rotation R about the fixed axis $\hat{\mathbf{n}}$. The result is straightforward:

$$F(\mathbf{r}) = e^{-(i/\hbar)\phi\hat{\mathbf{n}}\cdot\mathbf{L}}\, f(\mathbf{r}) \tag{11.13}$$

and defines the unitary *rotation operator*

$$U_R = e^{-(i/\hbar)\phi\hat{\mathbf{n}}\cdot\mathbf{L}} \tag{11.14}$$

The rotation operator U_R rotates a state represented by the wave function $\psi(\mathbf{r})$ into a new state represented by $\psi'(\mathbf{r}) = U_R\psi(\mathbf{r})$. For any operator A, we define a rotationally transformed operator A' such that $A'\psi'(\mathbf{r}) = U_R A\psi(\mathbf{r})$, which implies that

$$A' = U_R A U_R^\dagger = e^{-(i/\hbar)\phi\hat{\mathbf{n}}\cdot\mathbf{L}} A e^{(i/\hbar)\phi\hat{\mathbf{n}}\cdot\mathbf{L}} \tag{11.15}$$

For infinitesimal rotations this becomes

$$\delta A = A' - A = [A, \boldsymbol{\delta\phi} \times \mathbf{r} \cdot \nabla] = \frac{i}{\hbar}\,[A, \boldsymbol{\delta\phi} \cdot \mathbf{L}] \tag{11.16}$$

A *vector operator* $\mathbf{A}$ is a set of three component operators whose expectation value in the rotated state $\psi'(\mathbf{r})$ is obtained by rotating the expectation value of $\mathbf{A}$ in the original state $\psi(\mathbf{r})$. For an infinitesimal rotation:

$$\langle\Psi'|\mathbf{A}|\Psi'\rangle - \langle\Psi|\mathbf{A}|\Psi\rangle = \boldsymbol{\delta\phi} \times \langle\Psi|\mathbf{A}|\Psi\rangle \tag{11.17}$$

Keeping only terms up to first order in the rotation angle, we find that this relation leads for a vector operator $\mathbf{A}$ to the condition

$$U_R \mathbf{A} U_R^\dagger - \mathbf{A} = \delta\mathbf{A} = -\boldsymbol{\delta\phi} \times \mathbf{A} \tag{11.18}$$

The operators $\mathbf{r}$, $\mathbf{p}$, and $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ are examples of vector operators. For these, substitution of (11.18) into (11.16) yields the commutation relation

$$\boxed{[\mathbf{A}, \hat{\mathbf{n}} \cdot \mathbf{L}] = i\hbar\hat{\mathbf{n}} \times \mathbf{A}} \tag{11.19}$$

Exercise 11.4. By letting $\mathbf{A} = \mathbf{r}$, $\mathbf{p}$, and $\mathbf{L}$ in (11.19), verify (11.4) and (11.5). Also check that if $\mathbf{A} = \mathbf{r}$ in (11.18), the correct expression for $\delta\mathbf{r}$ is obtained.

Exercise 11.5. Apply to infinitesimal *translations* the reasoning that led to (11.19) for rotations, and rederive the fundamental commutation relations of $\mathbf{r}$ and $\mathbf{p}$.

A *scalar operator* S is an operator whose expectation value is invariant under rotation and which therefore transforms according to the rule

$$\delta S = 0 \tag{11.20}$$

It follows from (11.16) that for a scalar operator S:

$$S\mathbf{L} - \mathbf{L}S = [S, \mathbf{L}] = 0 \tag{11.21}$$

The scalar product $\mathbf{A} \cdot \mathbf{B}$ of two vector operators is the simplest example of a scalar operator, since

$$\delta(\mathbf{A} \cdot \mathbf{B}) = \delta\mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \delta\mathbf{B} = -\delta\boldsymbol{\phi} \times \mathbf{A} \cdot \mathbf{B} - \mathbf{A} \cdot \delta\boldsymbol{\phi} \times \mathbf{B} = 0$$

We note in particular that the orbital angular momentum $\mathbf{L}$ commutes with any (potential energy) function $V(r)$ of the scalar r, with the kinetic energy $p^2/2m$ and with $\mathbf{L}^2$:

$$\boxed{[\mathbf{L}, \mathbf{L}^2] = 0} \tag{11.22}$$

Exercise 11.6. Verify that any component of $\mathbf{L}$, say L_z, commutes with $\mathbf{L}^2 = L_x^2 + L_y^2 + L_z^2$ by using the commutation relations (11.5).

Exercise 11.7. Prove that $\mathbf{r} \cdot \mathbf{L}$ and $\mathbf{p} \cdot \mathbf{L}$ are null operators.

Exercise 11.8. Does the equation

$$\mathbf{L} \times \mathbf{L} = i\hbar\mathbf{L} \tag{11.23}$$

make sense?

Exercise 11.9. If two rotations $\delta\boldsymbol{\phi}_1$ and $\delta\boldsymbol{\phi}_2$ are performed, in that sequence, to second order, the total displacement is

$$\begin{aligned}\mathbf{a}_{21} = {} & \delta\boldsymbol{\phi}_1 \times \mathbf{r} + \delta\boldsymbol{\phi}_2 \times \mathbf{r} + \delta\boldsymbol{\phi}_2 \times (\delta\boldsymbol{\phi}_1 \times \mathbf{r}) \\ & + \frac{1}{2}\,\delta\boldsymbol{\phi}_1 \times (\delta\boldsymbol{\phi}_1 \times \mathbf{r}) + \frac{1}{2}\,\delta\boldsymbol{\phi}_2 \times (\delta\boldsymbol{\phi}_2 \times \mathbf{r})\end{aligned}$$

The displacement $\mathbf{a}_{12}$ is obtained by interchanging the rotations 1 and 2. Show that the difference displacement $\mathbf{a}_{21} - \mathbf{a}_{12} = (\delta\boldsymbol{\phi}_2 \times \delta\boldsymbol{\phi}_1) \times \mathbf{r}$ is effected by the rotation $\delta\boldsymbol{\phi}_{12} = \delta\boldsymbol{\phi}_2 \times \delta\boldsymbol{\phi}_1$ and that this rotation correctly induces a second-order difference $\delta_{21}^2 f - \delta_{12}^2 f = -\frac{i}{\hbar}\,\delta\boldsymbol{\phi}_2 \times \delta\boldsymbol{\phi}_1 \cdot \mathbf{L}f$ owing to the validity of the commutation relations (11.23).

In summary, we conclude that it is not possible, in general, to specify and measure *more* than one component $\hat{\mathbf{n}} \cdot \mathbf{L}$ of orbital angular momentum. It is, however, possible to specify $\mathbf{L}^2$ simultaneously with any one component of $\mathbf{L}$. The Hamiltonian, $H = p^2/2m + V(r)$, for a particle moving in a central-force field commutes with $\mathbf{L}$, and it is therefore possible to require the energy eigenstates of a rotationally invariant system to be also eigenvectors of $\mathbf{L}^2$ and of one component of $\mathbf{L}$, which is usually chosen to be L_z. Thus, in preparation for solving the energy eigenvalue problem, it is useful first to derive the common eigenvectors of L_z and $\mathbf{L}^2$. Just as the eigenvalue problem for the harmonic oscillator could be solved by two methods, one analytic based on differential equations, the other algebraic and starting from the commutation relations, here also we can proceed by two quite different routes. We give precedence to the algebraic method, saving the analytic approach for subsequent sections.

2. *Algebraic Approach to the Angular Momentum Eigenvalue Problem.* We start with three Hermitian operators J_x, J_y, J_z, which are assumed to satisfy the same commutation relations (11.5) as the three Cartesian components of orbital angular momentum:

$$\boxed{[J_x, J_y] = i\hbar J_z \qquad [J_y, J_z] = i\hbar J_x, \qquad [J_z, J_x] = i\hbar J_y} \tag{11.24}$$

Nothing is different here from (11.5) except the names of the operators. We have replaced **L** by **J** in order to emphasize that the eigenvalue problem, which will be solved in this section by the algebraic method, has the capacity of representing a much larger class of physical situations than orbital angular momentum of a single particle.

Let us consider the eigenvalue problem of one of the components of **J**, say J_z. We construct the operators

$$J_+ = J_x + iJ_y, \qquad J_- = J_x - iJ_y \tag{11.25}$$

and

$$\mathbf{J}^2 = J_x^2 + J_y^2 + J_z^2 \tag{11.26}$$

Of these three operators only $\mathbf{J}^2$ is Hermitian. The operator J_- is the adjoint of J_+. From the commutation relations (11.24), we infer further commutation relations:

$$J_z J_+ - J_+ J_z = \hbar J_+ \tag{11.27}$$

$$J_- J_z - J_z J_- = \hbar J_- \tag{11.28}$$

$$J_+ J_- - J_- J_+ = 2\hbar\, J_z \tag{11.29}$$

$$\mathbf{J}^2\,\mathbf{J} - \mathbf{J}\,\mathbf{J}^2 = 0 \tag{11.30}$$

We note the useful identity

$$\mathbf{J}^2 - J_z^2 \pm \hbar J_z = J_\pm J_\mp \tag{11.31}$$

Exercise 11.10. Prove (11.31).

Since according to (11.30), J_z commutes with $\mathbf{J}^2$, it is possible to obtain simultaneous eigenvectors for these two operators. This option will help us to distinguish between the various independent eigenvectors of J_z. If we denote the eigenvalues of J_z by $m\hbar$ and those of $\mathbf{J}^2$ by $\lambda\hbar^2$, the eigenvalue problem can be written as

$$J_z|\lambda m\rangle = m\hbar|\lambda m\rangle \tag{11.32}$$

$$\mathbf{J}^2|\lambda m\rangle = \lambda\hbar^2|\lambda m\rangle \tag{11.33}$$

The eigenvalues m and λ, belonging to the same eigenvector, satisfy the inequality

$$\lambda \geq m^2 \tag{11.34}$$

To prove this inequality, we consider

$$\mathbf{J}^2 - J_z^2 = J_x^2 + J_y^2 = \frac{1}{2}(J_+J_- + J_-J_+) = \frac{1}{2}(J_+J_+^\dagger + J_+^\dagger J_+)$$

Since an operator of the form $AA^\dagger$ has only nonnegative expectation values, we conclude that

$$\langle \lambda m | \mathbf{J}^2 - J_z^2 | \lambda m \rangle \geq 0$$

from which the inequality (11.34) follows.

Next we develop again a ladder procedure similar to the method employed in Section 10.6 for the harmonic oscillator. If we act on Eq. (11.32) with J_+ and J_- and apply (11.27)* and (11.28), we obtain

$$J_z J_+ | \lambda m \rangle = (m + 1)\hbar J_+ | \lambda m \rangle \tag{11.35}$$

$$J_z J_- | \lambda m \rangle = (m - 1)\hbar J_- | \lambda m \rangle \tag{11.36}$$

Also,

$$\mathbf{J}^2 J_\pm | \lambda m \rangle = \lambda \hbar^2 J_\pm | \lambda m \rangle$$

Hence, if $|\lambda m\rangle$ is an eigenvector of J_z and $\mathbf{J}^2$ with eigenvalues $m\hbar$ and $\lambda\hbar^2$, then $J_\pm |\lambda m\rangle$ is also an eigenket of these same operators but with eigenvalues $(m \pm 1)\hbar$ and $\lambda\hbar^2$, respectively. We may therefore write

$$J_+ | \lambda m \rangle = C_+(\lambda m)\hbar | \lambda\ m{+}1 \rangle \tag{11.37}$$

$$J_- | \lambda m \rangle = C_-(\lambda m)\hbar | \lambda\ m{-}1 \rangle \tag{11.38}$$

where $C_\pm(\lambda m)$ are complex numbers yet to be determined.

For a given value of λ, the inequality $\lambda \geq m^2$ limits the magnitude of m. Hence, there must be a greatest value of m, $\text{Max}(m) = j$, for any given λ. Application of the raising operator J_+ to the eigenket $|\lambda j\rangle$ should not lead to any new eigenket; hence,

$$J_+ | \lambda j \rangle = 0$$

Multiplying on the left by J_-, we obtain

$$J_- J_+ | \lambda j \rangle = (\mathbf{J}^2 - J_z^2 - \hbar J_z) | \lambda j \rangle = (\lambda - j^2 - j)\hbar^2 | \lambda j \rangle = 0$$

from which the relation between j and λ follows:

$$\lambda = j(j + 1) \tag{11.39}$$

Similarly, there must be a lowest value of m, $\text{Min}(m) = j'$, such that

$$\lambda = j'(j' - 1) \tag{11.40}$$

Equations (11.39) and (11.40) are consistent only if

$$j' = -j \quad \text{or} \quad j' = j + 1$$

The second solution is meaningless because it violates the assumption that j is the greatest and j' the smallest value of m. Hence $j' = -j$.

Since the eigenvalues of J_z have both upper and lower bounds, it must be possible for a given value of λ or j to reach $|\lambda j'\rangle = |\lambda, -j\rangle$ from $|\lambda j\rangle$ in a sufficient number of steps descending the ladder by repeated application of the lowering operator J_-. In each downward step, m decreases by unity; it follows that $j - j' = 2j$ must be a nonnegative integer. Hence, j must be either a *nonnegative integer* or a *half-integer*, i.e., the only possible values for j are

$$\boxed{j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \ldots} \tag{11.41}$$

For a given value of j, the eigenvalues of J_z are

$$m\hbar = j\hbar, (j-1)\hbar, (j-2)\hbar, \ldots, -(j-1)\hbar, -j\hbar$$

These are $2j+1$ in number, and there are thus $2j+1$ orthogonal eigenvectors for every value of j. Since $2j+1$ can be any positive number, we see that for every dimension it is possible to construct a vector space that is closed under the operations of the algebra of the three operators J_x, J_y, J_z, which are constrained by the commutation relations. This is the key to the idea of an irreducible representation of the rotation group (see Chapter 17).

With the aid of the identity (11.31), we can now determine the coefficients $C_\pm$ in (11.37) and (11.38). Note that from (11.37),

$$\langle \lambda m | J_- = \langle \lambda m+1 | C_+^*(\lambda m)\hbar$$

Multiplying this and (11.37), we get

$$\langle \lambda m | J_- J_+ | \lambda\ m \rangle = |C_+(\lambda\ m)|^2 \hbar^2 \langle \lambda\ m+1 | \lambda\ m+1 \rangle$$

Let us assume that all eigenvectors are normalized to unity. Then, since

$$\begin{aligned} \langle \lambda m | J_- J_+ | \lambda m \rangle &= \langle \lambda m | \mathbf{J}^2 - J_z^2 - \hbar J_z | \lambda m \rangle \\ &= [j(j+1) - m^2 - m]\hbar^2 \langle \lambda m | \lambda m \rangle \end{aligned}$$

we conclude that

$$|C_+(\lambda m)|^2 = j(j+1) - m(m+1) = (j-m)(j+m+1)$$

The phases of C_+ are not determined and may be chosen arbitrarily. A usual choice is to make the phases equal to zero.

Exercise 11.11. Using the fact that J_- is the adjoint of J_+, show that

$$C_-(\lambda\ m) = C_+^*(\lambda\ m-1)$$

We then have

$$J_+ | \lambda m \rangle = \sqrt{(j-m)(j+m+1)}\ \hbar | \lambda\ m+1 \rangle \tag{11.42}$$

and

$$J_- | \lambda m \rangle = \sqrt{(j+m)(j-m+1)}\ \hbar | \lambda\ m-1 \rangle \tag{11.43}$$

Exercise 11.12. By the use of Eqs. (11.42) and (11.43), construct the matrices representing J_x, J_y, and J_z in a basis that consists of the common eigenvectors of J_z and $\mathbf{J}^2$.

Since it is impossible to specify two or more components of $\mathbf{J}$ simultaneously, it is of interest to ask what the physical implications of the noncommutivity of such operators as J_x and J_y are. If the commutation relations (11.24) are applied to the Heisenberg uncertainty relation (10.54), we have the inequality

$$\Delta J_x\ \Delta J_y \geq \frac{\hbar}{2} |\langle J_z \rangle| \tag{11.44}$$

Is there a state for which *all* components of $\mathbf{J}$ can be simultaneously determined, such that $\Delta J_x = \Delta J_y = \Delta J_z = 0$? From (11.44) and similar inequalities, we see that this can be the case only if the expectation values of all components of $\mathbf{J}$ vanish:

$$\langle \mathbf{J} \rangle = 0$$

But $(\Delta J_x)^2 = \langle J_x^2 \rangle - \langle J_x \rangle^2$; hence, if both ΔJ_x and $\langle J_x \rangle$ vanish, then we must also have

$$\langle J_x^2 \rangle = 0$$

and similar conditions must hold for the other components. In other words, the desired state has a sharp nonfluctuating angular momentum value of zero, or

$$\mathbf{J}|\psi\rangle = 0 \tag{11.45}$$

and therefore $\mathbf{J}^2|\psi\rangle = 0$. The only solution of (11.45) is the state $|00\rangle \equiv |j = 0, m = 0\rangle$. For all other states, quantum fluctuations make it impossible to specify J_x, J_y, and J_z simultaneously. As a consequence, for all states, except the state $|0\ 0\rangle$, (11.34) is a proper inequality and

$$j(j + 1) > m^2$$

The component J_z can never be as "long" as the vector **J**! In the vector model of angular momentum, in the "old quantum theory," the states $|jm\rangle$ were visualized by circular cones centered on the z axis (Figure 11.2).

We have thus completed the explicit construction of all the operators **J** which satisfy the commutation relations (11.24). The treatment of the eigenvalue problem given here has been a formal one. Only the commutation relations, the Hermitian nature of **J**, and certain implicit assumptions about the existence of eigenvectors were utilized, but nothing else. In particular, no explicit use was made of the con-

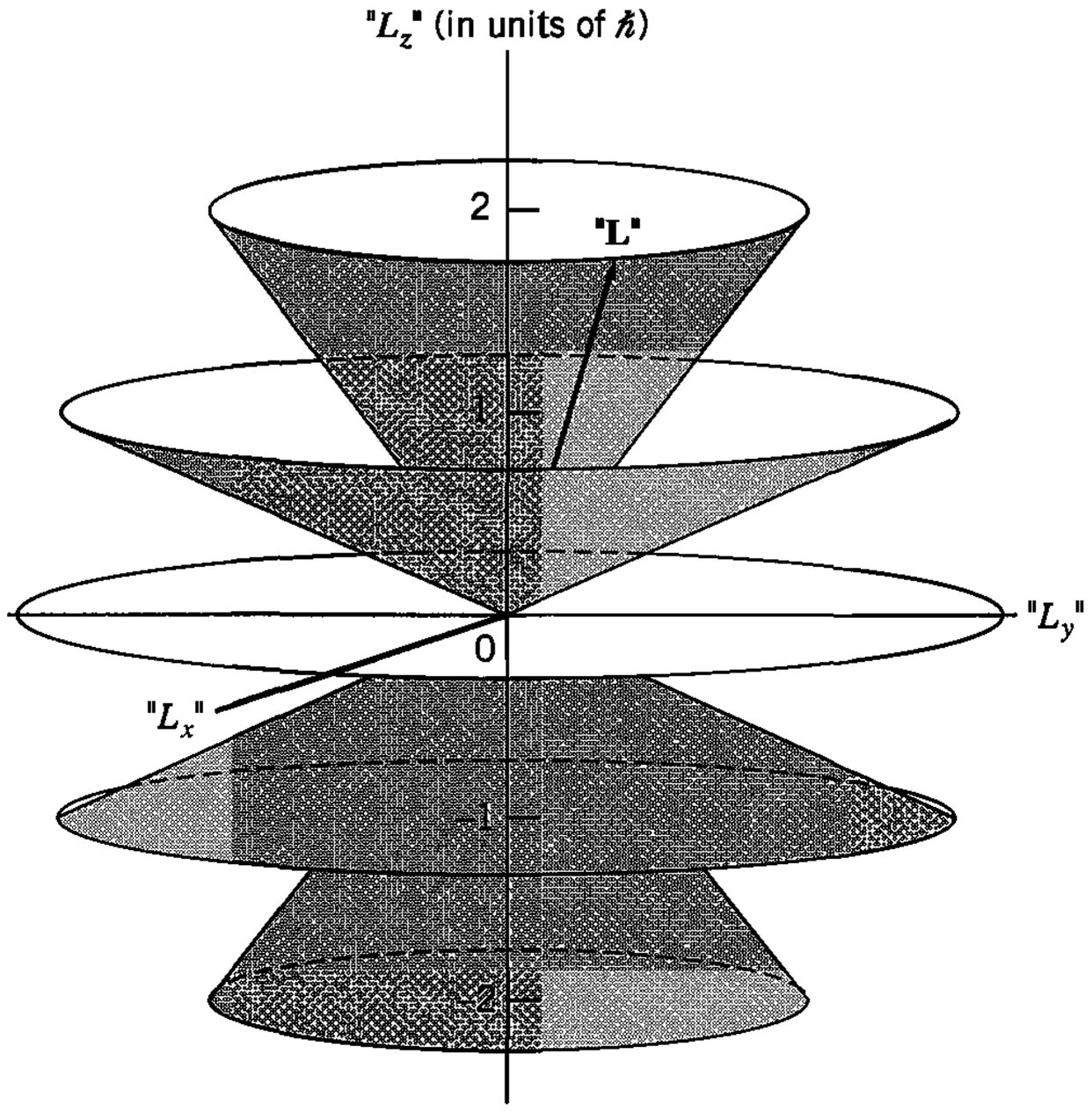


Figure 11.2. A cartoon illustrating the angular momentum eigenstates $|\ell m\rangle$ for $\ell = 2$. The semiclassical angular momentum vector "**L**" has length $\sqrt{\ell(\ell + 1)}\hbar = \sqrt{6}\hbar$, but its z component assumes the possible values "L_z" $= 0, \pm\hbar, \pm 2\hbar$. This is visualized by supposing that the "**L**" vector is stochastically distributed on one of the circular cones with uniform probability. The mean values of "L_x"' and "L_x^2" obtained from this model [zero and $(6 - m^2)\hbar^2/2$, respectively] agree with the expectation values of the corresponding quantum operators.

nection between **J** and spatial rotations, nor is **J** necessarily $\mathbf{r} \times \mathbf{p}$. Our solution of the eigenvalue problem thus extends to any three operators that satisfy commutation relations like (11.24), e.g., the isospin operator in the theory of elementary particles and in nuclear physics. We must now return to orbital angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ and analyze its eigenvalue problem in more detail.

3. ***The Eigenvalue Problem for L_z and $\mathbf{L}^2$.*** It is convenient to express the orbital angular momentum as a differential operator in terms of spherical polar coordinates defined by

$$x = r \sin\theta \cos\varphi, \qquad y = r \sin\theta \sin\varphi, \qquad z = r \cos\theta$$

The calculations become transparent if we note that the gradient operator can be written in terms of the unit vectors of spherical polar coordinates as

$$\boxed{\nabla = \hat{\mathbf{r}} \frac{\partial}{\partial r} + \hat{\boldsymbol{\varphi}} \frac{1}{r \sin\theta} \frac{\partial}{\partial \varphi} + \hat{\boldsymbol{\theta}} \frac{1}{r} \frac{\partial}{\partial \theta}} \tag{11.46}$$

where (Figure 11.3)

$$\begin{aligned} \hat{\mathbf{r}} &= \sin\theta \cos\varphi \hat{\mathbf{x}} + \sin\theta \sin\varphi \hat{\mathbf{y}} + \cos\theta \hat{\mathbf{z}} \\ \hat{\boldsymbol{\varphi}} &= -\sin\varphi \hat{\mathbf{x}} + \cos\varphi \hat{\mathbf{y}} \\ \hat{\boldsymbol{\theta}} &= \cos\theta \cos\varphi \hat{\mathbf{x}} + \cos\theta \sin\varphi \hat{\mathbf{y}} - \sin\theta \hat{\mathbf{z}} \end{aligned} \tag{11.47}$$

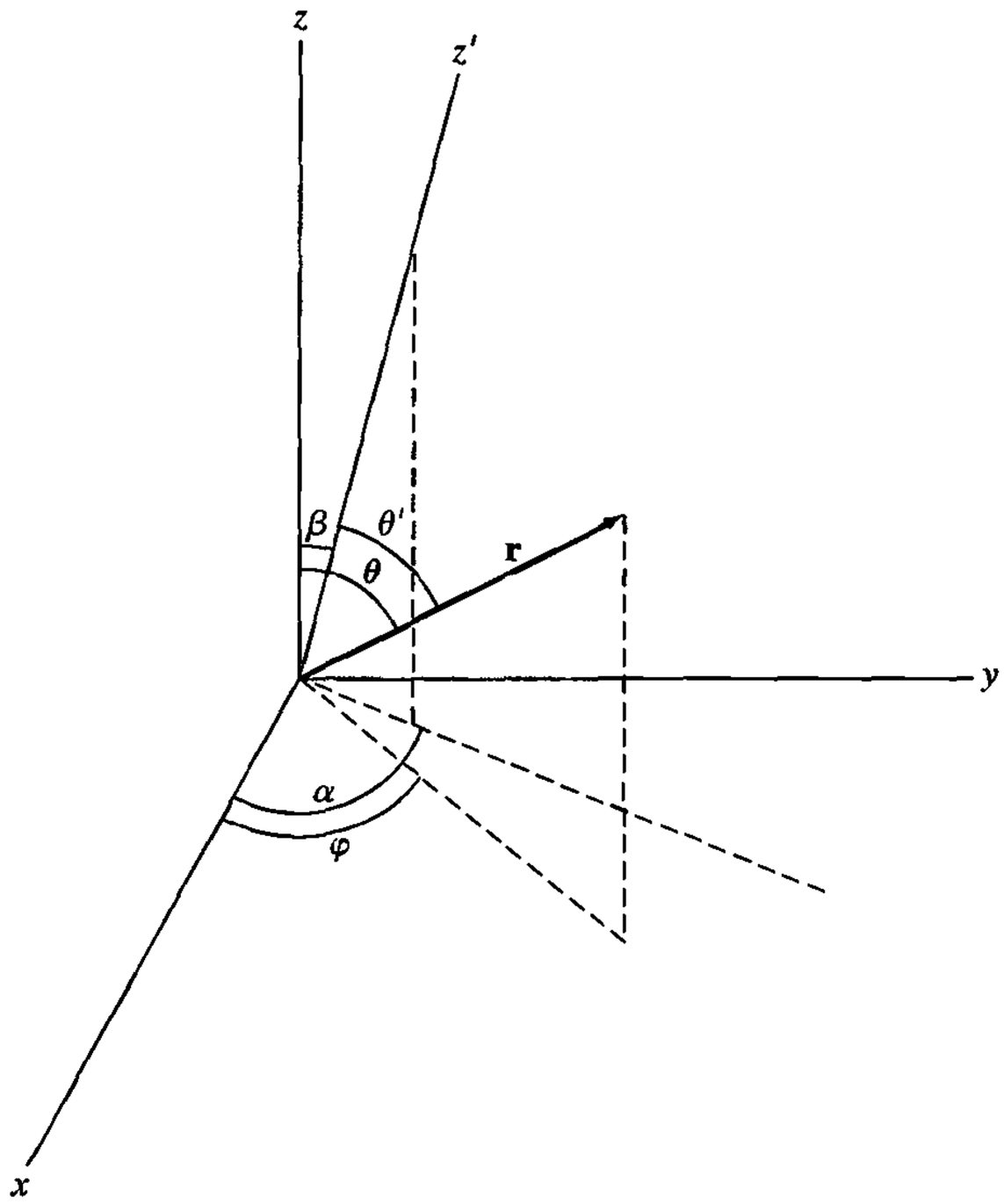


Figure 11.3. Angles used in the addition theorem of spherical harmonics. The angles α and β are the azimuth and the polar angle of the z' axis in the Cartesian xyz coordinate frame. They are also the first two Euler angles specifying the orientation of the Cartesian coordinate system $x'y'z'$ with respect to xyz (see Figure 17.1). The third Euler angle γ is left unspecified here, and the x' and y' axes are not shown. The projections of the z' axis and the vector **r** on the xy plane are dashed lines.

Exercise 11.13. Verify (11.46), which is done most easily by using the relation

$$df = d\mathbf{r} \cdot \nabla f$$

and considering displacements along the curves on which two of the three spherical polar coordinates r, φ, θ are held fixed.

From (11.46) it is evident that the three spherical polar components of the momentum operator $(\hbar/i)\,\nabla$, unlike its Cartesian components, do not commute. The angular momentum may now be expressed as

$$\boxed{\begin{aligned} L_x &= \frac{\hbar}{i}\left(-\sin\varphi\,\frac{\partial}{\partial\theta} - \cos\varphi\cot\theta\,\frac{\partial}{\partial\varphi}\right) \\ L_y &= \frac{\hbar}{i}\left(\cos\varphi\,\frac{\partial}{\partial\theta} - \sin\varphi\cot\theta\,\frac{\partial}{\partial\varphi}\right) \\ L_z &= \frac{\hbar}{i}\,\frac{\partial}{\partial\varphi} \end{aligned}} \tag{11.48}$$

From the representations (11.48), we obtain

$$\boxed{\mathbf{L}^2 = L_x^2 + L_y^2 + L_z^2 = -\hbar^2\left[\frac{1}{\sin^2\theta}\,\frac{\partial^2}{\partial\varphi^2} + \frac{1}{\sin\theta}\left(\sin\theta\,\frac{\partial}{\partial\theta}\right)\right]} \tag{11.49}$$

Exercise 11.14. Derive (11.49) from (11.48).

The spherical coordinate representation is particularly advantageous for treating the eigenvalue problem of L_z:

$$L_z\,\Phi(\varphi) = \frac{\hbar}{i}\,\frac{\partial\Phi}{\partial\varphi} = m\hbar\Phi(\varphi) \tag{11.50}$$

where use has been made of the conclusion of the last section that the eigenvalues of any component of angular momentum must have the form $m\hbar$, with m being restricted to integers and half-integers. The solutions of (11.50) are simply

$$\Phi(\varphi) = e^{im\varphi} \tag{11.51}$$

The simultaneous eigenfunctions of L_z and $\mathbf{L}^2$ must then be of the form

$$Y(\theta, \varphi) = \Phi(\varphi)\Theta(\theta) \tag{11.52}$$

What conditions must we impose on the solutions (11.52) to give us physically acceptable wave functions? It is implicit in the fundamental postulates of quantum mechanics that the wave function for a particle without spin must have a definite value at every point in space.[1] Hence, we demand that the wave function be a *single-valued* function of the particle's position. In particular, Φ must take on the same value whether the azimuth of a point is given by φ or $\varphi + 2\pi$.

[1]This and other arguments for the single-valuedness of the wave function are discussed in Merzbacher (1962). The conclusions depend strongly on the topology of the space of the coordinates of the system.

Applied to (11.51), the condition $\Phi(\varphi + 2\pi) = \Phi(\varphi)$ restricts us to those solutions

$$\Phi(\varphi) = e^{im\varphi} \tag{11.53}$$

for which $m = 0, \pm 1, \pm 2, \ldots$, i.e., an *integer*.

The half-integral values of m are unacceptable as eigenvalues of a component of orbital angular momentum, but we will see that they are admissible as eigenvalues of different kinds of angular momentum (spin) and also as eigenvalues of other physical quantities that satisfy the angular momentum commutation relations (e.g., isospin).

Equation (11.53) shows explicitly that the eigenvalues of L_z are $m\hbar$. Thus, a measurement of L_z can yield as its result only the value $0, \pm\hbar, +2\hbar, \ldots$. Since the z axis points in an arbitrarily chosen direction, it must be true that the angular momentum about any axis is *quantized* and can upon measurement reveal only one of these discrete values.

The term *magnetic quantum number* is frequently used for the integer m because of the part this number plays in describing the effect of a uniform magnetic field **B** on a charged particle moving in a central field.

Exercise 11.15. Use the Cartesian representation

$$L_z = \frac{\hbar}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)$$

to show that $(x \pm iy)^m$ is an eigenfunction of L_z.

With (11.49), the eigenvalue problem (11.33) for $\mathbf{L}^2$ now can be formulated explicitly as follows:

$$L^2\, Y(\theta, \varphi) = -\hbar^2\left[\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}\right]Y(\theta, \varphi) = \hbar^2\lambda Y(\theta, \varphi) \tag{11.54}$$

We require the functions $Y(\theta, \varphi)$ to be eigenfunctions of L_z as well. When we substitute from (11.52) and (11.51), we get the differential equation

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right) - \frac{m^2}{\sin^2\theta}\Theta + \lambda\Theta = 0 \tag{11.55}$$

By a change of variables

$$\xi = \cos\theta, \qquad F(\xi) = \Theta(\theta) \tag{11.56}$$

(11.55) is transformed into

$$\frac{d}{d\xi}\left[(1 - \xi^2)\frac{dF}{d\xi}\right] - \frac{m^2}{1 - \xi^2}F + \lambda F = 0 \tag{11.57}$$

For the particular case $m = 0$, (11.57) assumes an especially simple form, familiar in many problems of mathematical physics,

$$\frac{d}{d\xi}\left[(1 - \xi^2)\frac{dF}{d\xi}\right] + \lambda F = 0 \tag{11.58}$$

and known as *Legendre's differential equation.* Its examination follows a conventional pattern.

Equation (11.58) does not change its form when $-\xi$ is substituted for ξ. Hence, we need to look only for solutions of (11.58) which are even or odd functions of ξ. Since $\xi \to -\xi$ implies $\theta \to \pi - \theta$ and $z \to -z$, these functions are symmetric or antisymmetric with respect to the xy plane.

The solution of (11.58) that is regular at $\xi = 0$ can be expanded in a power series,

$$F(\xi) = \sum_{k=0}^{\infty} a_k \xi^k$$

Substitution into (11.58) yields the recursion relation

$$(k + 1)(k + 2)a_{k+2} + [\lambda - k(k + 1)]a_k = 0 \tag{11.59}$$

Equation (11.59) shows that in the even case ($a_1 = 0$) all even coefficients are proportional to a_0, and in the odd case ($a_0 = 0$) all odd coefficients are proportional to a_1. As an eigenvalue of a positive operator, λ must be a nonnegative number. If the series does not terminate at some finite value of k, the ratio $a_{k+2}/a_k \to k/(k + 2)$ as $k \to \infty$. The series thus behaves like $\Sigma\ (1/k)\xi^k$ for even or odd k, implying that it diverges logarithmically for $\xi = \pm 1$, that is, for $\theta = 0$ and π. For the same reason, we exclude the second linearly independent solution of (11.58).[2] Such singular functions, although solutions of the differential equation for almost all values of ξ, are not acceptable eigenfunctions of $\mathbf{L}^2$.

We conclude that the power series must terminate at some finite value of $k = \ell$, where ℓ is a nonnegative integer, and that all higher powers vanish. According to (11.59), this will happen if λ has the value

$$\lambda = \ell(\ell + 1) \tag{11.60}$$

We have thus rederived the law for the eigenvalues of $\mathbf{L}^2$, in agreement with the results of Section 11.2. The *orbital angular momentum quantum number* ℓ assumes the values 0, 1, 2, 3, . . . , and the measured values of $\mathbf{L}^2$ can only be 0, $2\hbar^2$, $6\hbar^2$, $12\hbar^2$, ... It is customary to designate the corresponding angular momentum states by the symbols S, P, D, F, . . . , which are familiar in atomic spectroscopy. If there are several particles in a central field, lower case letters s, p, d, . . . will be used to identify the angular momentum state of each particle, and capital letters S, P, D, . . . will be reserved for the total orbital angular momentum.[3]

The conventional form of the polynomial solutions of (11.58) is

$$P_\ell(\xi) = \frac{1}{2^\ell \cdot \ell!} \frac{d^\ell}{d\xi^\ell} (\xi^2 - 1)^\ell \tag{11.61}$$

These are called *Legendre polynomials*. The coefficient of ξ^k in the expansion of (11.61) is, for $\ell + k =$ even, easily seen to be

$$a_k = \frac{(-1)^{(\ell-k)/2}(\ell + k)(\ell + k - 1) \ldots (k + 2)(k + 1)}{2^\ell \cdot \ell!} \binom{\ell}{\frac{1}{2}(\ell + k)} \tag{11.62}$$

where the last factor is a binomial coefficient. For $\ell + k =$ odd, $a_k = 0$. We verify readily that (11.59) is satisfied by the coefficients a_k, and hence that $P_\ell(\xi)$ indeed

[2]For a rigorous treatment, see Bradbury (1984), p. 473.
[3]See Haken and Wolf (1993), p. 171.

solves (11.58). The peculiar constant factor in (11.61) has been adopted because it gives

$$P_\ell(\pm 1) = (\pm 1)^\ell \tag{11.63}$$

The first few Legendre polynomials are[4]

$$\begin{aligned} P_0(\xi) &= 1 & P_3(\xi) &= \tfrac{1}{2}(5\xi^3 - 3\xi) \\ P_1(\xi) &= \xi & P_4(\xi) &= \tfrac{1}{8}(35\xi^4 - 30\xi^2 + 3) \\ P_2(\xi) &= \tfrac{1}{2}(3\xi^2 - 1) & P_5(\xi) &= \tfrac{1}{8}(63\xi^5 - 70\xi^3 + 15\xi) \end{aligned} \tag{11.64}$$

Since $P_\ell(\cos\theta)$ is an eigenfunction of the Hermitian operator $\mathbf{L}^2$, it is clear from the general theorems of Chapter 10 that the Legendre polynomials must be orthogonal. Only the integration over the polar angle θ concerns us here—not the entire volume integral—and we expect that

$$\int_0^\pi P_\ell(\cos\theta)P_{\ell'}(\cos\theta)\sin\theta\, d\theta = 0 \quad \text{if } \ell' \neq \ell \tag{11.65}$$

No complex conjugation is needed because the Legendre polynomials are real functions. The orthogonality relation

$$\int_{-1}^{+1} P_\ell(\xi)P_{\ell'}(\xi)\, d\xi = 0 \quad \text{if } \ell' \neq \ell \tag{11.66}$$

can also be proved directly, using the definition (11.61) and successive integrations by parts. The normalization of these orthogonal polynomials can also be obtained easily by ℓ-fold integration by parts:

$$\begin{aligned} \int_{-1}^{+1} [P_\ell(\xi)]^2\, d\xi &= \left(\frac{1}{2^\ell \cdot \ell!}\right)^2 \int_{-1}^{+1} \left[\frac{d^\ell}{d\xi^\ell}(\xi^2 - 1)^\ell\right]\left[\frac{d^\ell}{d\xi^\ell}(\xi^2 - 1)^\ell\right] d\xi \\ &= (-1)^\ell \frac{1}{(2^\ell \cdot \ell!)^2} \int_{-1}^{+1} \left[\frac{d^{2\ell}}{d\xi^{2\ell}}(\xi^2 - 1)^\ell\right](\xi^2 - 1)^\ell\, d\xi \\ &= (-1)^\ell \frac{(2\ell)!}{(2^\ell \cdot \ell!)^2} \int_{-1}^{+1} (\xi^2 - 1)^\ell\, d\xi \doteq \frac{2}{2\ell + 1} \end{aligned} \tag{11.67}$$

Exercise 11.16. Prove the orthogonality relation (11.66) directly, using the definition (11.61).

As usual in the study of special functions, it is helpful to introduce a *generating function* for Legendre polynomials. Such a generating function is

$$(1 - 2\xi s + s^2)^{-1/2} = \sum_{n=0}^{\infty} P_n(\xi)s^n \qquad (|s| < 1) \tag{11.68}$$

To prove the identity of the coefficients $P_n(\xi)$ in (11.68) with the Legendre polynomials defined by (11.61), we derive a simple recurrence formula by differentiating (11.68) with respect to s:

$$(\xi - s)(1 - 2\xi s + s^2)^{-3/2} = \sum_{n=0}^{\infty} nP_n(\xi)s^{n-1} \tag{11.69}$$

[4]For pictorial representations of Legendre polynomials and other orbital angular momentum eigenfunctions, see Brandt and Dahmen (1985), Section 9.2.

or, by the use of (11.68),

$$(\xi - s)\sum_{n=0}^{\infty} P_n(\xi)s^n = (1 - 2\xi s + s^2)\sum_{n=0}^{\infty} nP_n(\xi)s^{n-1}$$

Equating the coefficients of each power of s, we obtain

$$(n + 1)P_{n+1}(\xi) = (2n + 1)\xi P_n(\xi) - nP_{n-1}(\xi) \tag{11.70}$$

By substituting $s^* = 0$ in (11.68) and (11.69), we see that

$$P_0(\xi) = 1, \qquad P_1(\xi) = \xi$$

in agreement with (11.64). The equivalence of the two definitions of the Legendre polynomials is completed by the demonstration that $P_n(\xi)$ as defined by (11.61) satisfies the recurrence formula (11.70).

Exercise 11.17. Prove the recurrence relation (11.70) for $P_n(\xi)$ defined in (11.61).

Having solved (11.58), it is not difficult to obtain the physically acceptable solutions of (11.57) with $m \neq 0$. If Legendre's equation (11.58) is differentiated m times and if the *associated Legendre functions*

$$P_\ell^m(\xi) = (1 - \xi^2)^{m/2}\frac{d^m P_\ell(\xi)}{d\xi^m} = \frac{1}{2^\ell \cdot \ell!}(1 - \xi^2)^{m/2}\frac{d^{\ell+m}}{d\xi^{\ell+m}}(\xi^2 - 1)^\ell \tag{11.71}$$

are defined for positive integers $m \leq \ell$, we deduce that

$$\frac{d}{d\xi}\left[(1 - \xi^2)\frac{dP_\ell^m(\xi)}{d\xi}\right] - \frac{m^2}{1 - \xi^2}P_\ell^m(\xi) + \ell(\ell + 1)P_\ell^m(\xi) = 0 \tag{11.72}$$

which is identical with (11.57) for $\lambda = \ell(\ell + 1)$. The associated Legendre functions with $m \leq \ell$ are the only nonsingular and physically acceptable solutions of (11.57). These functions are also called *associated Legendre functions of the first kind* to distinguish them from the second kind, $Q_\ell^m(\xi)$, which is the singular variety.

Exercise 11.18. Use the inequality (11.34) to verify that the magnetic quantum number cannot exceed the orbital angular momentum quantum number.

The associated Legendre functions are orthogonal in the sense that

$$\int_{-1}^{+1} P_\ell^m(\xi)P_{\ell'}^m(\xi)\,d\xi = 0 \quad \text{if } \ell \neq \ell' \tag{11.73}$$

Note that in this relation the two superscripts m are the same. Legendre functions with *different* values of m are generally *not* orthogonal. For purposes of normalization, we note that

$$\int_{-1}^{+1} [P_\ell^m(\xi)]^2\,d\xi = \frac{2}{2\ell + 1}\frac{(\ell + m)!}{(\ell - m)!} \tag{11.74}$$

We leave the proof to the interested reader. When ξ is changed to $-\xi$, $P_\ell^m(\xi)$ merely retains or changes its sign, depending on whether $\ell + m$ is an even or odd integer. It is natural to supplement the definition (11.71) by defining the associated Legendre functions for $m = 0$ as

$$P_\ell^0(\xi) = P_\ell(\xi) \tag{11.75}$$

Returning now to (11.52) and (11.54), we see that the solutions of (11.54), which are separable in spherical polar coordinates, are products of $e^{im\varphi}$ and $P_\ell^m(\cos\theta)$. Since (11.55) is unchanged if m is replaced by $-m$, and since P_ℓ^m is the only admissible solution of this equation, it follows that the same associated Legendre function must be used for a given absolute value of m.

Exercise 11.19. Justify the following alternative definition of associated Legendre functions:

$$P_\ell^m(\xi) = \frac{(-1)^m}{2^\ell \cdot \ell!}\frac{(\ell+m)!}{(\ell-m)!}(1-\xi^2)^{-m/2}\frac{d^{\ell-m}}{d\xi^{\ell-m}}(\xi^2-1)^\ell \tag{11.76}$$

The first few associated Legendre functions are

$$P_1^1(\xi) = \sqrt{1-\xi^2}, \qquad P_2^1(\xi) = 3\xi\sqrt{1-\xi^2}, \qquad P_2^2(\xi) = 3(1-\xi^2) \tag{11.77}$$

4. ***Spherical Harmonics.*** It is convenient to define the *spherical harmonics* $Y_\ell^m(\theta, \varphi)$ as the separable solutions (11.52) that are normalized with respect to an integration over the entire solid angle. For $m \geq 0$,

$$\boxed{Y_\ell^m(\theta,\varphi) = \sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}}(-1)^m e^{im\varphi}P_\ell^m(\cos\theta)} \tag{11.78}$$

Spherical harmonics with negative superscripts (subject to the restriction $-\ell \leq m \leq \ell$) will be defined by

$$Y_\ell^m(\theta,\varphi) = (-1)^m[Y_\ell^{-m}(\theta,\varphi)]^* \tag{11.79}$$

The spherical harmonics are normalized simultaneous eigenfunctions of L_z and $\mathbf{L}^2$ such that

$$L_z Y_\ell^m = \frac{\hbar}{i}\frac{\partial Y_\ell^m}{\partial\varphi} = m\hbar Y_\ell^m \tag{11.80}$$

$$\mathbf{L}^2 Y_\ell^m = \ell(\ell+1)\hbar^2 Y_\ell^m \tag{11.81}$$

The first few spherical harmonics are listed below:

$$\begin{aligned}
Y_0^0 &= \frac{1}{\sqrt{4\pi}} \\
Y_1^0 &= \sqrt{\frac{3}{4\pi}}\cos\theta = \sqrt{\frac{3}{4\pi}}\frac{z}{r} \\
Y_1^{\pm1} &= \mp\sqrt{\frac{3}{8\pi}}e^{\pm i\varphi}\sin\theta = \mp\sqrt{\frac{3}{8\pi}}\frac{x \pm iy}{r} \\
Y_2^0 &= \sqrt{\frac{5}{16\pi}}(3\cos^2\theta - 1) = \sqrt{\frac{5}{16\pi}}\frac{2z^2-x^2-y^2}{r^2} \\
Y_2^{\pm1} &= \mp\sqrt{\frac{15}{8\pi}}e^{\pm i\varphi}\cos\theta\sin\theta = \mp\sqrt{\frac{15}{8\pi}}\frac{(x \pm iy)z}{r^2} \\
Y_2^{\pm2} &= \sqrt{\frac{15}{32\pi}}e^{\pm 2i\varphi}\sin^2\theta = \sqrt{\frac{15}{32\pi}}\frac{(x \pm iy)^2}{r^2}
\end{aligned} \tag{11.82}$$

Under a coordinate reflection, or inversion, through the origin, which is realized by the transformation $\varphi \to \varphi + \pi$ and $\theta \to \pi - \theta$, the azimuthal wave function $e^{im\varphi}$ is multiplied by $(-1)^m$, and $P_\ell^m(\cos\theta)$ by $(-1)^{\ell+m}$. Hence, $Y_\ell^m(\theta, \varphi)$ is multiplied by $(-1)^\ell$, when $\mathbf{r}$ is changed to $-\mathbf{r}$.

The spherical harmonics are thus eigenfunctions of the parity operator U_p which changes $\mathbf{r}$ into $-\mathbf{r}$:

$$U_P\psi(r) = \psi(-r) \tag{11.83}$$

We have

$$U_P Y_\ell^m(\theta, \varphi) = (-1)^\ell Y_\ell^m(\theta, \varphi) \tag{11.84}$$

i.e., Y_ℓ^m has definite parity in consonance with the parity (evenness or oddness) of the angular momentum quantum number ℓ. This result is compatible with the reflection properties of orbital angular momentum. The operator $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ is invariant under reflection of all three Cartesian coordinates; it is an antisymmetric tensor of rank 2, or an axial vector, since both $\mathbf{r}$ and $\mathbf{p}$ change sign under reflection. Hence,

$$[U_P, \mathbf{L}] = 0 \tag{11.85}$$

and it follows that all eigenfunctions of L_z and $\mathbf{L}^2$ must have definite parity. Since Y_ℓ^m is obtained from $Y_\ell^{-\ell}$ by repeated application of the raising operator $L_+ = L_x + iL_y$ and since U_P commutes with L_+, all orbital angular momentum eigenfunctions with the same value of ℓ must have the same parity.

The spherical harmonics form an orthonormal set, since

$$\int_0^{2\pi}\int_0^{\pi} [Y_\ell^m(\theta, \varphi)]^* Y_{\ell'}^{m'}(\theta, \varphi) \sin\theta \, d\theta \, d\varphi = \delta_{\ell,\ell'}\delta_{m,m'} \tag{11.86}$$

Although no detailed proof will be given here, it is important to note that the spherical harmonics do form a complete set for the expansion of wave functions. Roughly, this can be seen from the following facts.

(a) The eigenfunctions $e^{im\varphi}$ of L_z are complete in the sense of Fourier series in the range $0 \leq \varphi \leq 2\pi$. Hence, a very large class of functions of φ can be expanded in terms of them.

(b) The Legendre polynomials $P_0(\xi)$, $P_1(\xi)$, $P_2(\xi)$, . . . , are the orthogonal polynomials that are obtained by applying the orthogonalization procedure described in Section 4.1 to the sequence of monomials $1, \xi, \xi^2, \ldots$, requiring that there be a polynomial of every degree and that they be orthogonal in the interval $-1 \leq \xi \leq +1$. Hence, ξ^k can be expressed in terms of Legendre polynomials, and any function that can be expanded in a uniformly converging power series of ξ can also be expanded in terms of Legendre polynomials. The same is true, though less obviously, for the associated Legendre functions of fixed order m, which also form a complete set as ℓ varies from 0 to ∞.

Hence, any wave function that depends on the angles θ and φ can be expanded in the mean in terms of spherical harmonics (see Section 4.2).

Exercise 11.20. Construct $P_2(\xi)$ by the orthogonalization procedure described above.

Some of the most frequently used expansions of angular functions in terms of spherical harmonics may be quoted without proof.

$$\cos\theta\, Y_\ell^m(\theta,\varphi) = \sqrt{\frac{(\ell+m+1)(\ell-m+1)}{(2\ell+1)(2\ell+3)}}\, Y_{\ell+1}^m + \sqrt{\frac{(\ell+m)(\ell-m)}{(2\ell+1)(2\ell-1)}}\, Y_{\ell-1}^m \tag{11.87}$$

$$\sin\theta\, e^{i\varphi}\, Y_\ell^m(\theta,\varphi) = -\sqrt{\frac{(\ell+m+1)(\ell+m+2)}{(2\ell+1)(2\ell+3)}}\, Y_{\ell+1}^{m+1} + \sqrt{\frac{(\ell-m)(\ell-m-1)}{(2\ell+1)(2\ell-1)}}\, Y_{\ell-1}^{m+1} \tag{11.88}$$

$$\sin\theta\, e^{-i\varphi}\, Y_\ell^m(\theta,\varphi) = \sqrt{\frac{(\ell-m+1)(\ell-m+2)}{(2\ell+1)(2\ell+3)}}\, Y_{\ell+1}^{m-1} - \sqrt{\frac{(\ell+m)(\ell+m-1)}{(2\ell+1)(2\ell-1)}}\, Y_{\ell-1}^{m-1} \tag{11.89}$$

The effect of the operators L_x and L_y on Y_ℓ^m is conveniently studied by employing the operators

$$L_+ = L_x + iL_y, \qquad L_- = L_x - iL_y \tag{11.90}$$

which, according to (11.48) may be written as

$$L_+ = \hbar e^{i\varphi}\left(\frac{\partial}{\partial\theta} + i\cot\theta\,\frac{\partial}{\partial\varphi}\right) \tag{11.91}$$

$$L_- = -\hbar e^{-i\varphi}\left(\frac{\partial}{\partial\theta} - i\cot\theta\,\frac{\partial}{\partial\varphi}\right). \tag{11.92}$$

The effect of $\partial/\partial\varphi$ on Y_ℓ^m is known from (11.80). To determine $\partial Y/\partial\theta$ we note that from the definitions (11.71) and (11.76)

$$\frac{dP_\ell^m}{d\xi} = \frac{1}{\sqrt{1-\xi^2}}\,P_\ell^{m+1} - \frac{m\xi}{1-\xi^2}\,P_\ell^m = \frac{-(\ell+m)(\ell-m+1)}{\sqrt{1-\xi^2}}\,P_\ell^{m-1} + \frac{m\xi}{1-\xi^2}\,P_\ell^m$$

With $\xi = \cos\theta$ and the definition (11.78), it is then easy to derive the relations

$$L_+ Y_\ell^m(\theta,\varphi) = \hbar\sqrt{(\ell-m)(\ell+m+1)}\,Y_\ell^{m+1}(\theta,\varphi) \tag{11.93}$$

$$L_- Y_\ell^m(\theta,\varphi) = \hbar\sqrt{(\ell+m)(\ell-m+1)}\,Y_\ell^{m-1}(\theta,\varphi) \tag{11.94}$$

These equations do not come as a surprise. The operators L_+ and L_- are raising and lowering operators for the magnetic quantum number, and Eqs. (11.93) and (11.94) are merely realizations of (11.42) and (11.43), which were derived directly from the commutation relations. The complete agreement between these two sets of equations shows that the choice of phase factor made for the spherical harmonics is consistent with the choice of phases for $C_+(\lambda m)$ in Section 11.2.

The addition theorem for spherical harmonics is a useful application. Consider two coordinate systems xyz and $x'y'z'$. The addition theorem is the formula express-

ing the eigenfunction $P_\ell(\cos\theta')$ of angular momentum about the z' axis in terms of the eigenfunctions $Y_\ell^m(\theta, \varphi)$ of L_z. Figure 11.3 indicates the various angles. The position vector $\mathbf{r}$ has angular coordinates θ, φ, and θ', φ' in the two coordinate systems. The direction of the z' axis in space is specified by its polar angle β and its azimuth α with respect to the unprimed system. Since P_ℓ is an eigenfunction of $\mathbf{L}^2$, only spherical harmonics with the same subscript ℓ can appear in the expansion. An interchange of θ, φ, and β, α is equivalent to the transformation $\theta' \to -\theta'$ and must leave the expansion unchanged, because $P_\ell(\cos\theta')$ is an even function of θ'. Hence, $P_\ell(\cos\theta')$ can also be expanded in terms of $Y_\ell^m(\beta, \alpha)$. In a rigid rotation of the figure about the z axis, α and φ change by equal amounts, and θ' remains constant. Hence, $P_\ell(\cos\theta')$ must be a function of $\varphi - \alpha$. All these requirements can be satisfied only if

$$P_\ell(\cos\theta') = \sum_{m=-\ell}^{\ell} c_m \, Y_\ell^{-m}(\beta, \alpha) Y_\ell^m(\theta, \varphi) \tag{11.95}$$

The coefficients c_m can be determined by using the condition

$$L_{z'} P_\ell(\cos\theta') = 0 \tag{11.96}$$

Since

$$\begin{aligned} L_{z'} &= \sin\beta\cos\alpha\, L_x + \sin\beta\sin\alpha\, L_y + \cos\beta\, L_z \\ &= \frac{1}{2}\sin\beta\, e^{-i\alpha} L_+ + \frac{1}{2}\sin\beta\, e^{i\alpha} L_- + \cos\beta\, L_z \end{aligned} \tag{11.97}$$

Eqs. (11.87), (11.88), (11.89), (11.93), and (11.94) may be used to evaluate $L_{z'} P_\ell(\cos\theta')$. If the linear independence of the spherical harmonics is invoked, we obtain, after some calculation, the simple result

$$c_{m\pm 1} = -c_m$$

Thus, $c_m = (-1)^m c_0$, and only c_0 need be determined. For this purpose we specialize to $\beta = 0$, or $\theta = \theta'$. Since, from the definitions of Y_ℓ^m and P_ℓ^m,

$$\boxed{Y_\ell^m(0, \varphi) = \sqrt{\frac{2\ell + 1}{4\pi}}\, \delta_{m0}} \tag{11.98}$$

and

$$\boxed{Y_\ell^0(\theta, \varphi) = \sqrt{\frac{2\ell + 1}{4\pi}}\, P_\ell(\cos\theta)} \tag{11.99}$$

it follows that

$$c_0 = \frac{4\pi}{2\ell + 1}$$

With (11.79), this proves the addition theorem in the form

$$\boxed{P_\ell(\cos\theta') = \frac{4\pi}{2\ell + 1} \sum_{m=-\ell}^{\ell} Y_\ell^{m*}(\beta, \alpha) Y_\ell^m(\theta, \varphi)} \tag{11.100}$$

The completeness of the spherical harmonics as basis functions for any function of the angles θ, φ, or of the direction of the vector $\mathbf{r}$, is expressed by the closure relation:

$$\sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} [Y_\ell^m(\beta, \alpha)]^* Y_\ell^m(\theta, \varphi) = \delta(\hat{\mathbf{r}}, \hat{\mathbf{r}}') \tag{11.101}$$

The solid-angle delta function on the right-hand side is equal to zero unless the two vectors $\hat{\mathbf{r}}(\theta, \varphi)$ and $\hat{\mathbf{r}}'(\beta, \alpha)$ coincide. It has the property

$$\int f(\hat{\mathbf{r}}')\delta(\hat{\mathbf{r}}, \hat{\mathbf{r}}')\, d\Omega' = f(\hat{\mathbf{r}}) \tag{11.102}$$

for any function $f(\hat{\mathbf{r}})$ of the spatial direction specified by θ, φ.[5] If (11.101) is combined with the addition theorem (11.100), the identity

$$\boxed{\sum_{\ell=0}^{\infty} (2\ell + 1)P_\ell(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') = 4\pi\delta(\hat{\mathbf{r}}, \hat{\mathbf{r}}')} \tag{11.103}$$

is obtained.

The delta function in three dimensions has a representation in spherical polar coordinates,

$$\delta(\mathbf{r} - \mathbf{r}') = \frac{\delta(r - r')}{r^2}\, \delta(\hat{\mathbf{r}}, \hat{\mathbf{r}}') \tag{11.104}$$

Hence, we infer the further identity:

$$\boxed{\delta(\mathbf{r} - \mathbf{r}') = \frac{\delta(r - r')}{r^2} \sum_{\ell=0}^{\infty} \frac{2\ell + 1}{4\pi} P_\ell(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')} \tag{11.105}$$

This formula will be useful in the theory of scattering from a spherical potential.

Exercise 11.21. Check (11.105) by integrating both sides over all of 3-space.

5. ***Angular Momentum and Kinetic Energy.*** Since the kinetic energy is represented in the coordinate (or momentum) representations by an operator proportional to ∇^2, it is expedient to relate $\mathbf{L}^2$ to the Laplacian. We make the calculation using the concise notation that takes advantage of the summation convention (summing from 1 to 3 over repeated indices) and of the Levi-Civita asymmetric (third-rank tensor) symbol ε_{ijk}. The Levi-Civita symbol is defined as follows:

$$\varepsilon_{ijk} = \begin{cases} 1 & ijk = 123, 231, 312 \text{ (even permutation of 123)} \\ -1 & ijk = 321, 213, 132 \text{ (odd permutation of 123)} \\ 0 & \text{when two or more indices are equal} \end{cases}$$

Using the summation convention, we formulate a simple identity,

$$\varepsilon_{ijk}\varepsilon_{nqk} = \delta_{in}\delta_{jq} - \delta_{iq}\delta_{jn} \tag{11.106}$$

[5]The delta function $\delta(\hat{\mathbf{r}}, \hat{\mathbf{r}}')$ is sometimes written as $\delta(\Omega - \Omega')$, but this is misleading, because it implies that Ω is a variable of integration. There is no such solid angle variable.

Exercise 11.22. Prove the equality (11.106).

It is evident that the components of orbital angular momentum can be written n the form

$$\boxed{L_k = -i\hbar\varepsilon_{ijk}\, x_i \frac{\partial}{\partial x_j}} \tag{11.107}$$

It follows that

$$\begin{aligned}
\mathbf{L}^2 = L_k L_k &= -\hbar^2 \varepsilon_{ijk}\varepsilon_{nqk} x_i \frac{\partial}{\partial x_j} x_n \frac{\partial}{\partial x_q} = -\hbar^2(\delta_{in}\delta_{jq} - \delta_{iq}\delta_{jn}) x_i \frac{\partial}{\partial x_j} x_n \frac{\partial}{\partial x_q} \\
&= -\hbar^2\left(x_i \frac{\partial}{\partial x_j} x_i \frac{\partial}{\partial x_j} - x_i \frac{\partial}{\partial x_j} x_j \frac{\partial}{\partial x_i} \right) \\
&= -\hbar^2\left[x_i \frac{\partial}{\partial x_i} + x_i x_i \frac{\partial^2}{\partial x_j \partial x_j} - \left(\frac{\partial x_j}{\partial x_j}\right) x_i \frac{\partial}{\partial x_i} - x_i x_j \frac{\partial^2}{\partial x_i \partial x_j} \right] \\
&= -\hbar^2\left(r^2\nabla^2 - 2\mathbf{r}\cdot\nabla - x_i x_j \frac{\partial^2}{\partial x_i \partial x_j} \right) = -\hbar^2\left(r^2\nabla^2 - 2r\frac{\partial}{\partial r} - r^2 \frac{\partial^2}{\partial r^2} \right)
\end{aligned}$$

and we arrive at the important identity

$$\boxed{\frac{\mathbf{L}^2}{\hbar^2} = -r^2\nabla^2 + \frac{\partial}{\partial r}\left(r^2 \frac{\partial}{\partial r} \right)} \tag{11.108}$$

Exercise 11.23. Line by line, work through the steps leading to Eq. (11.108).

An alternative, representation-independent, method for deriving (11.108) starts with the operator identity,

$$\begin{aligned}
\mathbf{L}^2 &= (\mathbf{r}\times\mathbf{p})\cdot(\mathbf{r}\times\mathbf{p}) = -(\mathbf{r}\times\mathbf{p})\cdot(\mathbf{p}\times\mathbf{r}) \\
&= -\mathbf{r}\cdot[\mathbf{p}\times(\mathbf{p}\times\mathbf{r})] = -\mathbf{r}\cdot[\mathbf{p}(\mathbf{p}\cdot\mathbf{r}) - p^2\mathbf{r}]
\end{aligned}$$

Since

$$[\mathbf{r}, p^2] = 2i\hbar\mathbf{p}$$

and

$$\mathbf{p}\cdot\mathbf{r} - \mathbf{r}\cdot\mathbf{p} = -3i\hbar I$$

we obtain

$$\mathbf{L}^2 = r^2p^2 + i\hbar\mathbf{r}\cdot\mathbf{p} - \mathbf{r}\cdot\mathbf{p}\,\mathbf{r}\cdot\mathbf{p} \tag{11.109}$$

The component of the gradient ∇f in the direction of $\mathbf{r}$ is $\partial f/\partial r$; hence, in the (spher ical) coordinate representation,

$$\mathbf{r}\cdot\mathbf{p} = \frac{\hbar}{i} r \frac{\partial}{\partial r}$$

and consequently

$$\mathbf{L}^2 = r^2p^2 + \hbar^2 r\frac{\partial}{\partial r} + \hbar^2 r\frac{\partial}{\partial r}\,r\frac{\partial}{\partial r} = r^2p^2 + 2\hbar^2 r\frac{\partial}{\partial r} + \hbar^2 r^2\frac{\partial^2}{\partial r^2} \qquad (11.110)$$
$$= r^2p^2 + \hbar^2\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)$$

in agreement with (11.108).

Since $\mathbf{L}$ and therefore also $\mathbf{L}^2$ commutes with any function of r, the kinetic energy operator is related to angular momentum by

$$\boxed{T = \frac{p^2}{2m} = \frac{\mathbf{L}^2}{2mr^2} - \frac{\hbar^2}{2mr^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)} \qquad (11.111)$$

Here we see explicitly that $\mathbf{L}$ commutes with T, since it is patently irrelevant whether differentiation with respect to r is performed before or after a rotation about the origin.

In order to establish the connection between the eigenfunctions of $\mathbf{L}^2$ and the solutions of Laplace's equation, we consider the eigenvalue problem for the last term in (11.108):

$$\frac{d}{dr}\,r^2\frac{df(r)}{dr} = Cf(r) \qquad (11.112)$$

This has the solution

$$f(r) = Ar^{\ell} + Br^{-\ell-1} \qquad (11.113)$$

and the eigenvalue

$$C = \ell(\ell + 1) \qquad (11.114)$$

Since $\mathbf{L}^2$ acts only on the variables θ and φ, and not on r, we see from (11.108) and (11.81), and by choosing the solution (11.113) which is not singular at the origin, that

$$\nabla^2 r^{\ell}Y_{\ell}^{m}(\theta, \varphi) = 0 \qquad (11.115)$$

Thus, the functions $r^{\ell}Y_{\ell}^{m}(\theta, \varphi)$ are regular solutions of Laplace's equation. From the definition (11.78) of spherical harmonics and inspection of the formula (11.71) for associated Legendre functions it follows that the functions $r^{\ell}Y_{\ell}^{m}(\theta, \varphi)$, when converted into functions of x, y, z, are homogeneous multinomials of degree ℓ. With nonnegative integer powers r, s, t, these functions can be expressed as $\Sigma a_{rst}\,x^r y^s z^t$, subject to the constraint $r + s + t = \ell$ and the requirement that they must be solutions of Laplace's equation, which accounts for their designation as *harmonic functions*.

Exercise 11.24. Show that the homogeneous multinomial

$$F = \sum_{r+s+t=\ell} a_{rst}\,x^r y^s z^t$$

has $(\ell + 1)(\ell + 2)/2$ coefficients and that the linear relations between them imposed by the requirement $\nabla^2 F = 0$ leave $2\ell + 1$ coefficients to be chosen arbitrarily, so

that the number of linearly independent harmonic multinomials of degree ℓ equals the number of orthogonal spherical harmonics of order ℓ.

Aided by the formula connecting the kinetic energy with angular momentum, we are now prepared to tackle the central-force problem in quantum mechanics.

Problems

1. For the state represented by the wave function

$$\psi = Ne^{-\alpha r^2}(x + y)z$$

(a) Determine the normalization constant N as a function of the parameter α.

(b) Calculate the expectation values of $\mathbf{L}$ and L^2.

(c) Calculate the variances of these quantities.

2. For a finite rotation by an angle α about the z axis, apply the rotation operator U_R to the function $f(\mathbf{r}) = ax + by$, and show that it transforms correctly.

3. Explicitly work out the $\mathbf{J}$ matrices for $j = 1/2$, 1, and 3/2.

4. Classically, we have for central forces

$$H = \frac{p_r^2}{2m} + \frac{\mathbf{L}^2}{2mr^2} + V(r)$$

where $p_r = (1/r)(\mathbf{r} \cdot \mathbf{p})$. Show that for translation into quantum mechanics we must write

$$p_r = \frac{1}{2}\left[\frac{1}{r}(\mathbf{r} \cdot \mathbf{p}) + (\mathbf{p} \cdot \mathbf{r})\frac{1}{r}\right]$$

and that this gives the correct Schrödinger equation with the Hermitian operator

$$p_r = \frac{\hbar}{i}\left(\frac{\partial}{\partial r} + \frac{1}{r}\right)$$

whereas $(\hbar/i)(\partial/\partial r)$ is not Hermitian.

5. Show that in D-dimensional Euclidean space the result of Problem 4 generalizes to

$$p_r = \frac{1}{2}\left[\frac{1}{r}(\mathbf{r} \cdot \mathbf{p}) + (\mathbf{p} \cdot \mathbf{r})\frac{1}{r}\right] = \frac{\hbar}{i}\left(\frac{\partial}{\partial r} + \frac{D-1}{2r}\right)$$

CHAPTER 12

Spherically Symmetric Potentials

If the potential energy is rotationally invariant, and thus dependent only on the distance r from a center of force, chosen as the coordinate origin, orbital angular momentum is conserved. This constant of the motion enables us to reduce the three-dimensional Schrödinger equation to an ordinary differential equation, the radial equation, analogous to the reduction of a central-force problem in classical mechanics to a dynamical problem for the radial coordinate r alone, provided that angular momentum conservation is used and the inertial centrifugal force introduced. As examples of central potentials, we solve the radial Schrödinger equation for the trivial case of a free particle ($V = 0$), the spherical square well, and the attractive Coulomb potential (the one-electron atom).

1. ***Reduction of the Central-Force Problem.*** Since the Hamiltonian

$$\boxed{H = \frac{p^2}{2m} + V(r)} \tag{12.1}$$

for a particle of mass m moving in a central-force field commutes with the orbital angular momentum operator,

$$[H, \mathbf{r} \times \mathbf{p}] = [H, \mathbf{L}] = 0 \tag{12.2}$$

angular momentum is a constant of the motion for a particle moving in a rotationally invariant potential. The operators H, L_z and $\mathbf{L}^2$ all commute with each other,

$$[H, L_z] = [H, \mathbf{L}^2] = [L_z, \mathbf{L}^2] = 0$$

in this case, and we can therefore require the energy eigenfunctions also to be eigenfunctions of L_z and $\mathbf{L}^2$. These eigenfunctions must then be of the separable form

$$\psi(\mathbf{r}) = R_E(r)Y_\ell^m(\theta, \varphi) \tag{12.3}$$

when spherical polar coordinates are used. The equation that is satisfied by the radial factor $R(r)$ is found if we express the Hamiltonian in terms of orbital angular momentum. Since $\mathbf{L}$ and therefore also $\mathbf{L}^2$ commutes with any function of r, we may use (11.111) to write the Schrödinger equation

$$H\psi \equiv [T + V(r)]\psi = E\psi$$

for central forces in the form

$$\left[-\frac{\hbar^2}{2mr^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\mathbf{L}^2\hbar^2}{2mr^2} + V(r)\right]\psi = E\psi \tag{12.4}$$

If the separated form (12.3) is substituted, this equation can be reduced to the ordinary differential equation for the radial eigenfunction $R_E(r)$:

$$\boxed{\left[-\frac{\hbar^2}{2mr^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right)+\frac{\ell(\ell+1)\hbar^2}{2mr^2}+V(r)\right]R_E(r)=ER_E(r)} \tag{12.5}$$

which is easier to solve than the original partial differential equation. Our procedure is entirely equivalent to the familiar *separation of variables* of the Laplacian operator in spherical polar coordinates, but we emphasize the physical meaning of the method.

Exercise 12.1. If you have never done it before, carry through the explicit calculation of ∇^2 in terms of spherical polar coordinates and derive (12.4) by comparison with (11.49). (See Appendix, Section 3, for a general formula for the Laplacian in curvilinear coordinates.)

It is sometimes convenient to introduce yet another radial wave function by the substitution

$$u(r) = r\,R(r) \tag{12.6}$$

From (12.6) we find that $u(r)$ obeys the *radial* equation

$$\boxed{-\frac{\hbar^2}{2m}\frac{d^2u}{dr^2}+\left[\frac{\hbar^2\ell(\ell+1)}{2mr^2}+V(r)\right]u=Eu} \tag{12.7}$$

This equation is identical in form with the one-dimensional Schrödinger equation except for the addition of the term $\hbar^2\ell(\ell+1)/2mr^2$ to the potential energy. This term is sometimes called the centrifugal potential, since it represents the potential whose negative gradient is the centrifugal force.

Although (12.7) is similar to the one-dimensional Schrödinger equation, the boundary conditions to be imposed on the solutions are quite different, since r is never negative. For instance, if ψ is to be finite everywhere, $u(r)$ must vanish at $r = 0$, according to the definition (12.6). A detailed discussion of these questions requires specific assumptions about the shape of the potential energy, and in this chapter the radial Schrödinger equation will be solved for several particular cases.

2. ***The Free Particle as a Central-Force Problem.*** In Section 4.4 the Schrödinger equation for a free particle ($V = 0$), with energy $E(\geq 0)$,

$$\frac{\hbar^2}{2m}\nabla^2\psi = E\psi \tag{12.8}$$

was treated quite naturally by the method of separation of variables using Cartesian coordinates, since these coordinates are particularly well suited for describing translations in Euclidean space. Nevertheless, it is useful also to look at the free particle problem as a special case of a potential that depends only on the radial coordinate r. The energy eigensolutions of Eq. (12.8) can then be assumed to be separable in spherical polar coordinates in the form (12.3). For $V = 0$, the function $R(r)$ (omitting the subscript E for brevity) must satisfy the radial equation

$$\left[-\frac{\hbar^2}{2mr^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right)+\frac{\ell(\ell+1)\hbar^2}{2mr^2}\right]R(r)=ER(r) \tag{12.9}$$

or, since, $\hbar k = \sqrt{2m\,E}$,

$$-\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \frac{\ell(\ell+1)}{r^2}R = k^2R \tag{12.10}$$

If we scale the radial coordinate by introducing the dimensionless variable

$$\rho = \sqrt{\frac{2m\,E}{\hbar^2}}\,r = kr, \tag{12.11}$$

the radial equation reduces to

$$\boxed{\frac{d^2R}{d\rho^2} + \frac{2}{\rho}\frac{dR}{d\rho} + \left[1 - \frac{\ell(\ell+1)}{\rho^2}\right]R = 0} \tag{12.12}$$

This differential equation is seen to be related to *Bessel's equation* if we make the transformation

$$R = \frac{J(\rho)}{\sqrt{\rho}} \tag{12.13}$$

and obtain

$$\frac{d^2J}{d\rho^2} + \frac{1}{\rho}\frac{dJ}{d\rho} + \left[1 - \frac{(\ell+1/2)^2}{\rho^2}\right]J = 0 \tag{12.14}$$

The *regular* solutions of this equation, which do not have a singularity at $\rho = 0$, are the Bessel functions $J_{\ell+1/2}(\rho)$. By (12.13) they are related to the *regular* solutions of Eq. (12.12), which are defined as

$$\boxed{j_\ell(z) = \sqrt{\frac{\pi}{2z}}\,J_{\ell+1/2}(z)} \tag{12.15}$$

and known as *spherical Bessel functions*. That the latter satisfy Eq. (12.12) is easily verified if their integral representation

$$j_\ell(z) = \frac{z^\ell}{2^{\ell+1}\cdot\ell!}\int_{-1}^{+1} e^{izs}(1-s^2)^\ell\,ds \tag{12.16}$$

is used. (The variable z, rather than ρ, is used in the last equations to emphasize that these formulas are valid for all complex values of the variable). The first term in the series expansion of (12.16) in powers of z is

$$j_\ell(z) = \frac{2^\ell\cdot\ell!}{(2\ell+1)!}\,z^\ell + O(z^{\ell+2}) \tag{12.17}$$

We note that the spherical Bessel functions are even or odd functions of their argument, depending on the parity of ℓ.

We thus see that the spherical Bessel function $j_\ell(kr)$ is the regular solution of the radial equation (12.12) and that the radial eigenfunction of the Schrödinger equation (12.10) for the free particle is

$$R(r) = Cj_\ell(kr) \tag{12.18}$$

A useful formula linking Bessel functions with Legendre polynomials is obtained by integrating (12.16) by parts ℓ times and using the definition (11.61). This leads to

$$j_\ell(z) = \frac{1}{2i^\ell} \int_{-1}^{+1} e^{izs} P_\ell(s)\, ds \tag{12.19}$$

The asymptotic form of the spherical Bessel functions can be derived from this expression by further integration by parts, and the leading term is

$$j_\ell(\rho) \cong \frac{\cos[\rho - (\ell + 1)\pi/2]}{\rho} \qquad (\rho \gg \ell) \tag{12.20}$$

All other solutions of (12.12) are singular at the origin and not admissible as energy eigenfunctions for the free particle.

Exercise 12.2. Verify that the asymptotic expression (12.20) for $j_\ell(\rho)$ satisfies the differential equation (12.12) to second order in ρ^{-1}.

A particularly simple *singular* solution of Eq. (12.12) is obtained from j_ℓ by noting that the differential equation is invariant under the substitution

$$\ell \to -\ell - 1 \tag{12.21}$$

If this transformation is applied to the asymptotic form (12.20), we obtain a linearly independent solution that can, for large positive ρ, be written in the form

$$n_\ell(\rho) \cong \frac{\sin[\rho - (\ell + 1)\pi/2]}{\rho} \qquad (\rho \gg \ell) \tag{12.22}$$

This particular singular solution of the radial equation, being asymptotically out of phase by $\pi/2$ compared to the regular solutions $j_\ell(\rho)$, is sometimes distinguished as "the" irregular solution, although any linear combination of j_ℓ and n_ℓ is also singular at the origin.

To exhibit the behavior of n_ℓ near the origin, it is merely necessary to subject the expression (12.17) for j_ℓ to the transformation (12.21). We see that the singularity n_ℓ at the origin is contained in the leading term which is proportional to $z^{-\ell-1}$. Its coefficient is most easily computed by applying Eq. (3.7) to the two solutions

$$\psi_1 = j_\ell(\rho)Y_\ell^m \quad \text{and} \quad \psi_2 = n_\ell(\rho)Y_\ell^m$$

of the Schrödinger equation. Since ψ_1 and ψ_2 correspond to the same energy, we infer from (3.7) by application of Gauss' divergence theorem that

$$\rho^2\left[j_\ell(\rho)\frac{dn_\ell(\rho)}{d\rho} - \frac{dj_\ell(\rho)}{d\rho}n_\ell(\rho)\right] = \text{constant} \tag{12.23}$$

This expression is the analogue of the Wronskian for the one-dimensional Schrödinger equation. Substitution of (12.20) and (12.22) into (12.23) shows that the constant has value unity for large ρ. Hence, its value must also be unity as $\rho \to 0$. Using the approximation (12.17) for j_ℓ near the origin, we derive from (12.23) for $z \approx 0$

$$n_\ell(z) = -\frac{(2\ell)!}{2^\ell \cdot \ell!} z^{-\ell-1} + O(z^{-\ell+1}) \tag{12.24}$$

The function $n_\ell(z)$ is known as the *spherical Neumann function*.

Exercise 12.3. Show that for any two solutions, R_1 and R_2, of the radial equation (12.5), the condition

$$r^2\left(R_1 \frac{dR_2}{dr} - \frac{dR_1}{dr} R_2\right) = \text{constant} \tag{12.25}$$

holds. Check this for the free particle as $r \to \infty$ and $r \to 0$.

Two other useful singular solutions of Eq. (12.12) are the *spherical Hankel functions of the first and second kind*, defined by

$$h_\ell^{(1)}(z) = j_\ell(z) + in_\ell(z) \tag{12.26}$$

and

$$h_\ell^{(2)}(z) = j_\ell(z) - in_\ell(z) \tag{12.27}$$

The generic name for the solutions of Bessel's equation (12.14) is *cylinder functions*, and the solutions of (12.12) are known, paradoxically, as *spherical cylinder functions*. The information we compile in this section about these special functions will be used in Section 12.3 and in Chapter 13 on scattering.

Like the Neumann function, the Hankel functions diverge as $z^{-\ell-1}$ near the origin:

$$h_\ell^{(1,2)}(z) = \mp\frac{i(2\ell)!}{2^\ell \cdot \ell!} z^{-\ell-1} + O^{(1,2)}(z^{-\ell+1}) \tag{12.28}$$

and their asymptotic behavior for large positive ρ is seen from (12.20) and (12.22) to be

$$\begin{aligned} h_\ell^{(1)}(\rho) &\cong \frac{1}{\rho} \exp\{i[\rho - (\ell + 1)\pi/2]\} \\ h_\ell^{(2)}(\rho) &\cong \frac{1}{\rho} \exp\{-i[\rho - (\ell + 1)\pi/2]\} \end{aligned} \tag{12.29}$$

The explicit forms of the spherical Bessel, Hankel, and Neumann functions for $\ell = 0$, 1 and 2, are given below:

$$\begin{aligned} j_0(z) &= \frac{\sin z}{z}, \qquad j_1(z) = \frac{\sin z}{z^2} - \frac{\cos z}{z} \\ j_2(z) &= \left(\frac{3}{z^3} - \frac{1}{z}\right)\sin z - \frac{3}{z^2}\cos z \end{aligned} \tag{12.30}$$

$$\begin{aligned} n_0(z) &= -\frac{\cos z}{z}, \qquad n_1(z) = -\frac{\cos z}{z^2} - \frac{\sin z}{z} \\ n_2(z) &= -\left(\frac{3}{z^3} - \frac{1}{z}\right)\cos z - \frac{3}{z^2}\sin z \end{aligned} \tag{12.31}$$

$$\begin{aligned} h_0^{(1)}(z) &= -i\frac{e^{iz}}{z}, \qquad h_1^{(1)}(z) = \left(-\frac{i}{z^2} - \frac{1}{z}\right)e^{iz} \\ h_2^{(1)}(z) &= \left(-\frac{3i}{z^3} - \frac{3}{z^2} + \frac{i}{z}\right)e^{iz} \end{aligned} \tag{12.32}$$

$$\begin{aligned} h_0^{(2)}(z) &= i\frac{e^{-iz}}{z}, \qquad h_1^{(2)}(z) = \left(\frac{i}{z^2} - \frac{1}{z}\right)e^{-iz} \\ h_2^{(2)}(z) &= \left(\frac{3i}{z^3} - \frac{3}{z^2} - \frac{i}{z}\right)e^{-iz} \end{aligned} \tag{12.33}$$

Exercise 12.4. Verify that for $E < 0$ none of the solutions of the free particle radial equation are physically acceptable owing to their asymptotic behavior.

The regular radial eigenfunctions of the Schrödinger equation for $V = 0$ constitute a complete set, as a consequence of a fundamental theorem concerning Sturm-Liouville differential equations,[1] of which (12.12) is an example. Hence, we have before us two alternative complete sets of eigenfunctions of the free particle Hamiltonian. They are the plane waves $e^{i\mathbf{k}\cdot\mathbf{r}}$ and the spherical waves $j_\ell(kr)Y_\ell^m(\theta, \varphi)$, where $\hbar k = \sqrt{2m\,E}$. Both sets exhibit an infinite degree of degeneracy, but for a given value of the energy, the number of plane waves is indenumerable, while the number of spherical waves is denumerable, corresponding to the countability of the integer quantum numbers ℓ and m. Nevertheless, these two sets of eigenfunctions are equivalent, and one kind must be capable of expansion in terms of the other, posing the problem of determining the coefficients in the expansion

$$e^{i\mathbf{k}\cdot\mathbf{r}} = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} c_{\ell,m} j_\ell(kr) Y_\ell^m(\theta, \varphi)$$

Actually, it is sufficient to specialize this relation to the case where $\mathbf{k}$ points along the z axis and consider the expansion

$$e^{ikz} = e^{ikr\cos\theta} = \sum_{\ell=0}^{\infty} a_\ell j_\ell(kr) P_\ell(\cos\,\theta)$$

From the orthogonality and normalization properties of Legendre polynomials we obtain (with $\xi = \cos\,\theta$)

$$a_\ell\, j_\ell(kr) = \frac{2\ell + 1}{2} \int_{-1}^{+1} e^{ikr\xi}\, P_\ell(\xi)\, d\xi \tag{12.34}$$

which we compare with (12.19) to establish the identity

$$\boxed{e^{ikz} = \sum_{\ell=0}^{\infty} (2\ell + 1) i^\ell j_\ell(kr) P_\ell(\cos\,\theta)} \tag{12.35}$$

This formula is especially useful in scattering theory. The more general expansion, with $\mathbf{k}$ pointing in an arbitrary direction, is obtained from (12.35) by use of the addition theorem for spherical harmonics:

$$\begin{aligned} e^{i\mathbf{k}\cdot\mathbf{r}} &= \sum_{\ell=0}^{\infty} (2\ell + 1) i^\ell j_\ell(kr) P_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{r}}) \\ &= 4\pi \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{+\ell} i^\ell j_\ell(kr) [Y_\ell^m(\beta, \alpha)]^* Y_\ell^m(\theta, \varphi) \end{aligned} \tag{12.36}$$

where α and β denote the spherical polar coordinates of the vector $\mathbf{k}$.

A useful asymptotic approximation to Eq. (12.36) is derived by substituting (12.20) on the right-hand side. For $kr \gg 1$ we get

$$\begin{aligned} e^{i\mathbf{k}\cdot\mathbf{r}} &\cong \sum_{\ell=0}^{\infty} (2\ell + 1) i^\ell \frac{\cos[kr - (\ell + 1)\pi/2]}{kr} P_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{r}}) \\ &= \frac{e^{ikr}}{2ikr} \sum_{\ell=0}^{\infty} (2\ell + 1) P_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{r}}) - \frac{e^{-ikr}}{2ikr} \sum_{\ell=0}^{\infty} (2\ell + 1)(-1)^\ell P_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{r}}) \end{aligned} \tag{12.37}$$

[1]Morse and Feshbach (1953), p. 738.

which by use of Eq. (11.103) reduces to

$$\boxed{e^{i\mathbf{k}\cdot\mathbf{r}} \cong \frac{2\pi e^{ikr}}{ikr}\,\delta(\hat{\mathbf{k}}, \hat{\mathbf{r}}) - \frac{2\pi e^{-ikr}}{ikr}\,\delta(\hat{\mathbf{k}}, -\hat{\mathbf{r}})} \tag{12.38}$$

a very convenient formula in scattering theory. It is seen that the leading term in the asymptotic expansion of the plane wave $e^{i\mathbf{k}\cdot\mathbf{r}}$ contributes only in the forward and backward directions, which is a physically reasonable result.

3. ***The Spherical Square Well Potential.*** The spherically symmetric square well in three dimensions is of interest because it is mathematically straightforward and approximates a number of real physical situations. Unlike the Coulomb potential, which gives rise to infinitely many discrete energy levels for bound states, the square well, owing to its finite range and finite depth, possesses only a finite number of such levels.

A square well is a central potential composed of two constant pieces: $V = -V_0$ for $r < a$ and $V = 0$ for $r > a$ (with $V_0 > 0$). The particle is free inside and outside the well, and subject to a force only at the discontinuity at $r = a$. In this section, the emphasis will be on the *bound states* of a particle in such a potential. The radial wave equation for a state of angular momentum ℓ is

$$-\frac{\hbar^2}{2mr^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \frac{\hbar^2\ell(\ell+1)}{2mr^2}R = (E + V_0)R \quad \text{for} \quad r < a \tag{12.39}$$

$$-\frac{\hbar^2}{2mr^2}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) + \frac{\hbar^2\ell(\ell+1)}{2mr^2}R = ER \quad \text{for} \quad r > a \tag{12.40}$$

For bound states $-V_0 \leq E \leq 0$.

The condition of regularity at the origin again restricts us to the spherical Bessel function for the solution inside the well. All the results of Section 12.2 apply provided that we take into account the fact that E must be replaced by the kinetic energy, $E - V = E + V_0$. Thus

$$R(r) = Aj_\ell\left(\sqrt{\frac{2m(E+V_0)}{\hbar^2}}\,r\right) \quad \text{for} \quad r < a \tag{12.41}$$

Outside the well we must exclude any solution of (12.40) that would increase exponentially at large distances. Since $E < 0$ for bound states, (12.40) has the same solutions as (12.10), but k is now an imaginary number. If we define

$$k = i\sqrt{\frac{-2mE}{\hbar^2}} = i\kappa$$

it is easily verified from the asymptotic forms in Section 12.2 that only the Hankel function of the first kind decreases exponentially. The eigenfunction outside the well must thus be of the form

$$R(r) = Bh_\ell^{(1)}\left(i\sqrt{\frac{-2mE}{\hbar^2}}\,r\right) \quad \text{for} \quad r > 0 \tag{12.42}$$

The interior and exterior solutions must be matched at $r = a$. In conformity with the analogous one-dimensional problem (see Section 3.5), the radial wave function and its derivative are required to be continuous at the discontinuity of the potential. Hence, the logarithmic derivative, $(1/R)\,dR/dr$ or $(1/u)\,du/dr$, must be continuous.

This condition, applied to (12.41) and (12.42) yields an equation for the allowed discrete energy eigenvalues:[2]

$$\sqrt{\frac{\alpha^2 - \kappa^2}{\kappa^2}} \frac{j_\ell'(\sqrt{\alpha^2 - \kappa^2}\, a)}{j_\ell(\sqrt{\alpha^2 - \kappa^2}\, a)} = \frac{h_\ell'^{(1)}(\kappa a)}{h_\ell^{(1)}(\kappa a)} \tag{12.43}$$

where $\alpha^2 = 2m\, V_0/\hbar^2$

The solutions for positive E are asymptotically oscillatory and correspond to scattering states in which the particle can go to infinity with a finite kinetic energy. They will be studied in Chapter 13.

Exercise 12.5. Compare the energy eigenvalues for S states in the three-dimensional square well with the energy eigenvalues of a one-dimensional square well of the same depth and width.

Exercise 12.6. If H is the sum of a Hermitian operator H_0 and a positive definite perturbation V, prove by a variational argument that the ground state energy of H_0 lies below the ground state energy of H. Apply this theorem to prove that in a central potential the ground state of a bound particle is an S state.

Exercise 12.7. Show that a spherical square well has no bound state unless

$$V_0\, a^2 > \frac{\hbar^2 \pi^2}{8m}$$

4. ***The Radial Equation and the Boundary Conditions.*** We now return to a general discussion of the radial equation for central forces. From Section 12.1 we know that the solutions of the Schrödinger equation can be constructed as

$$\psi(r, \theta, \varphi) = \frac{u(r)}{r} Y_\ell^m(\theta, \varphi) \tag{12.44}$$

Since r does not change under reflection, these wave functions have the same parity as Y_ℓ^m. Hence, for even ℓ we have states of even parity, and for odd ℓ we have states of odd parity.

The radial wave function $u(r)$ must satisfy the equation

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dr^2} + \left[\frac{\hbar^2 \ell(\ell + 1)}{2mr^2} + V(r)\right]u = Eu \tag{12.45}$$

The general principles of quantum mechanics require that the eigenfunctions (12.44) be normalizable. Since the spherical harmonics are normalized to unity, the eigenfunctions corresponding to discrete eigenvalues must satisfy the condition

$$\int\int \psi^*\psi\, r^2\, dr\, d\Omega = \int_0^\infty u^*u\, dr = 1 \tag{12.46}$$

If E lies in the continuous part of the spectrum, the eigenfunctions must be normalized in the sense of (4.33), or

$$\int_0^\infty u_E^* u_{E'}\, dr = \delta(E - E') \tag{12.47}$$

[2]Schiff (1968), p. 86, gives useful recurrence relations for spherical cylinder functions and their derivatives.

Most situations of practical interest are covered if we assume that $V(r)$ is finite everywhere except possibly at the origin and that near $r = 0$ it can be represented by[3]

$$V(r) \approx cr^{\alpha} \tag{12.48}$$

with α an integer and $\alpha \geq -1$. Furthermore, we assume that $V \to 0$ as $r \to \infty$.

We must not forget that, since division by r is involved, (11.111) is not a representation of the kinetic energy at the coordinate origin. For the same reason (12.45) is valid only for $r \neq 0$ and must be supplemented by a boundary condition at $r = 0$. Without going into detail, we note that the appropriate boundary condition is obtained by demanding that the Hamiltonian, or energy, operator must be self-adjoint in the sense of (4.35). This is the condition which consistency of the probability interpretation of quantum mechanics imposes on the eigenfunctions of H.

Applying this requirement to the operator

$$H = -\frac{\hbar^2}{2mr^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{\mathbf{L}^2}{2mr^2} + V(r)$$

we find, by integrating by parts, that any two physically admissible eigensolutions of (12.45) must satisfy the condition

$$\lim_{r\to 0}\left(u_1^*\frac{du_2}{dr} - \frac{du_1^*}{dr}u_2\right) = 0 \tag{12.49}$$

In applications, this condition usually may be replaced by the much simpler one requiring that $u(r)$ vanish at the origin:

$$u(0) = 0 \tag{12.50}$$

In most cases, this boundary condition singles out correctly the set of eigenfunctions that pass the test (12.49), but mildly singular wave functions are occasionally encountered (e.g., in the relativistic theory of the hydrogen atom, Section 24.9).

If in the immediate vicinity of the origin V can be neglected in comparison with the centrifugal term, which for $r \to 0$ increases as $1/r^2$, (12.45) reduces near $r = 0$ to

$$\frac{d^2u}{dr^2} - \frac{\ell(\ell+1)}{r^2}u = 0 \tag{12.51}$$

for states with $\ell \neq 0$. Potentials of the form (12.48) at small r, including the square well and the Coulomb potential, are examples of this. The general solution of (12.51) is

$$u = Ar^{\ell+1} + Br^{-\ell} \tag{12.52}$$

Since $\ell \geq 1$, the boundary condition (12.49) or (12.50) eliminates the second solution; hence, $B = 0$. Thus, for any but S states, $u(r)$ must be proportional to $r^{\ell+1}$ at the origin and ψ must behave as r^{ℓ}. Hence a power series solution of (12.45) must have the form

$$u(r) = r^{\ell+1}(a_0 + a_1 r + a_2 r^2 + \cdots) \tag{12.53}$$

[3]For potentials that are more singular at the origin, see Morse and Feshbach (1953), pp. 1665–1667.

If $\ell = 0$ (S states), the terms in (12.45) containing V and E cannot be neglected, and a separate investigation is required to obtain the behavior of the wave function near the origin. Even then the form (12.53) remains applicable for S states in most cases.

Exercise 12.8. For a potential $V = -C/r$ and angular momentum $\ell = 0$, show that the general solution of (12.45) is of the form

$$u = A\left(r - \frac{C}{2}r^2 + \cdots\right) + B(1 - C\log r + \cdots) \tag{12.54}$$

for small values of r and infer that for S states again we must require that $B = 0$.

Assuming that the potential energy vanishes at great distances, the radial equation (12.45) reduces to

$$\frac{d^2u}{dr^2} + \frac{2m\,E}{\hbar^2}u = 0 \tag{12.55}$$

as $r \to \infty$. Equation (12.55) possesses oscillatory solutions for positive E and exponential solutions for negative E, with the increasing exponential excluded by the condition that ψ must be normalizable in the sense of (12.46) or (12.47).

If $E < 0$, the eigenfunctions have the asymptotic behavior

$$u(r) \cong \exp\left(-\frac{\sqrt{-2m\,E}}{\hbar}r\right)$$

representing spatially confined, or *bound*, states. The boundary conditions will in general allow only certain discrete energy eigenvalues.

For bound states, the radial equation is conveniently transformed by the introduction of the dimensionless variable

$$\rho = \sqrt{\frac{-2m\,E}{\hbar^2}}\,r = \kappa r \tag{12.56}$$

Sometimes it is also convenient to remove from the unknown dependent variable the portions that describe its behavior at $r = 0$ and $r = \infty$. Thus, we introduce a new function $w(\rho)$ by setting

$$\boxed{u(\rho) = \rho^{\ell+1}e^{-\rho}w(\rho)} \tag{12.57}$$

Substituting this expression into (12.45), we obtain

$$\boxed{\frac{d^2w}{d\rho^2} + 2\left(\frac{\ell+1}{\rho} - 1\right)\frac{dw}{d\rho} + \left[\frac{V}{E} - \frac{2(\ell+1)}{\rho}\right]w = 0} \tag{12.58}$$

Of the solutions of this equation, we seek those that satisfy the boundary condition at infinity and at the origin.

5. ***The Coulomb Potential.*** Let us now suppose that V is the potential energy of the Coulomb attraction between a fixed charge Ze and a moving particle of charge $-e$,

$$V(r) = -\frac{Ze^2}{r} \tag{12.59}$$

For the hydrogen atom, $-e$ is the charge of the electron and $Z = 1$. According to the discussion of the last section and especially Eq. (12.52), the radial wave function $u(r)$ must behave as $r^{\ell+1}$ near the origin. This is also true for S states, as shown in Exercise 12.8.

The energy levels and eigenfunctions of *bound* states in the nonrelativistic theory will be discussed in this section and the next. The energy continuum ($E > 0$) of a particle in a Coulomb potential is the subject of Section 13.8 on scattering.

For convenience we introduce a dimensionless parameter

$$\rho_0 = \frac{Ze^2\kappa}{|E|} = \sqrt{\frac{2m}{|E|}}\frac{Ze^2}{\hbar} \tag{12.60}$$

such that

$$\frac{E}{V} = \frac{\rho}{\rho_0} \tag{12.61}$$

For this potential, the differential equation (12.58) can then be written as

$$\rho\frac{d^2w}{d\rho^2} + 2(\ell + 1 - \rho)\frac{dw}{d\rho} + [\rho_0 - 2(\ell + 1)]w = 0 \tag{12.62}$$

A simple two-term recursion relation is found if we expand $w(\rho)$ in a power series:

$$w(\rho) = a_0 + a_1\rho + a_2\rho^2 + \cdots \tag{12.63}$$

We substitute (12.63) into (12.62) and equate to zero the coefficient of ρ^k. The result is, for $k \geq 0$,

$$(k + 1)k\,a_{k+1} + 2(\ell + 1)(k + 1)a_{k+1} - 2k\,a_k + [\rho_0 - 2(\ell + 1)]a_k = 0$$

or

$$\frac{a_{k+1}}{a_k} = \frac{2(k + \ell + 1) - \rho_0}{(k + 1)(k + 2\ell + 2)} \tag{12.64}$$

This recursion relation shows that, for $k > 0$, the coefficients a_k are proportional to $a_0 \neq 0$. This power series must terminate at some finite maximum power. If it failed to do so, all terms with $k > (1/2)\rho_0 - (\ell + 1)$ would have the same sign. Furthermore, it is easy to verify that if $k > C\,\rho_0 + 2(C - 1)(\ell + 1)$, where C is a constant, $C > 1/2$, the ratio a_{k+1}/a_k is always greater than the corresponding ratio of coefficients in the power series expansion of $e^{(2-1/C)\rho}$. If C is chosen to be greater than 1, it follows from (12.57) that for large values of ρ the radial eigenfunction $u(\rho)$ *diverges* faster than $e^{(1-1/C)\rho}$. Such a strongly divergent wave function is not quadratically integrable and is not acceptable to represent the physical system. Hence, the series (12.63) must terminate, and $w(\rho)$ must be a polynomial. Let us suppose its degree to be N, that is, $a_{N+1} = 0$, but $a_N \neq 0$. Equation (12.64) leads to the condition

$$\boxed{\rho_0 = 2(N + \ell + 1)} \tag{12.65}$$

where $\ell = 0, 1, 2 \ldots$ and $N = 0, 1, 2\ldots$.

Exercise 12.9. Assume, contrary to the conventional procedure, that the radial eigenfunctions for a bound state of the hydrogen atom can be written as

$$u = \rho^k e^{-\rho}\left(b_0 + \frac{b_1}{\rho} + \frac{b_2}{\rho^2} + \cdots\right) \tag{12.66}$$

Obtain the recursion relations for the coefficients, and show that the boundary conditions give the same eigenvalues and eigenfunctions as usual. [A series like (12.66) in descending powers of r can be useful for calculating approximate radial eigenfunctions, if V behaves as the Coulomb potential at large, but not at small, distances.]

It is amusing to contemplate that as innocuous an equation as (12.65) is equivalent to the Balmer formula for the energy levels in hydrogenic atoms. To see this, we merely substitute (12.60) into (12.65) and define the *principal quantum number*

$$n \equiv N + \ell + 1 = \frac{\rho_0}{2} \tag{12.67}$$

The result is

$$\boxed{E_n = -\frac{Z^2 m e^4}{2\hbar^2 n^2}} \tag{12.68}$$

As is well known from the elementary Bohr theory,

$$\boxed{a = \frac{\hbar^2}{me^2}} \tag{12.69}$$

sets the length scale in the quantum description of the hydrogenic atom. The length a is termed the first *Bohr radius* of hydrogen if $m = m_e$ is the mass and $-e$ the charge of the electron. Its numerical value is

$$a = 0.529177 \times 10^{-8} \text{ cm}$$

Using this quantity, we can write the energy simply as

$$E_n = -\frac{Z^2 e^2}{2an^2} \tag{12.70}$$

Also, we see that

$$\kappa = \frac{Z}{na} \quad \text{and} \quad \rho = \frac{Z}{na} r \tag{12.71}$$

Since N is by its definition a nonnegative integer, it is obvious from (12.67) that n must be a positive integer,

$$n \geq 1 \tag{12.72}$$

The *ground state* of the hydrogen atom corresponds to $n = 1$, $\ell = 0$, with an energy of approximately -13.6 eV. There are infinitely many discrete energy levels; they have a point of accumulation at $E_0 = 0$ $(n \to \infty)$.

The fact that the energy depends only on the quantum number n implies that there are in general several linearly independent eigenfunctions of the Schrödinger equation for hydrogen corresponding to the same value of the energy. In the first place, $2\ell + 1$ different eigenfunctions of the same energy are obtained by varying the magnetic quantum number m in integral steps from $-\ell$ to ℓ. Second, there are n values of $\ell(\ell = 0, 1, 2, \ldots, n - 1)$ consistent with a given value of n. Hence, all energy levels with $n > 1$ are degenerate, and the number of linearly independent stationary states of hydrogen having the energy (12.68), the *degree of degeneracy*, is

$$\sum_{\ell=0}^{n-1} (2\ell + 1) = n^2 \tag{12.73}$$

For example, in standard spectroscopic notation (n followed by the symbol for ℓ), the first excited level of hydrogen ($n = 2$) is fourfold degenerate and consists of the $2S$ state and three $2P$ states. (The degeneracy is doubled if the spin is taken into account.)

The occurrence of degeneracy can often be ascribed to some transparent symmetry property of the physical system. For instance, the degeneracy with respect to magnetic quantum numbers is clearly present for any central potential. It has its origin in the absence of a preferred spatial direction in such systems, and it reflects their invariance with regard to rotations about the origin.

The degeneracy of energy eigenstates with different values of ℓ is not a general property of central forces. It occurs for the Coulomb potential because the Hamiltonian of this system is invariant under a less obvious symmetry, which generates an additional constant of the motion. Any departure from a strict $1/r$ dependence of the potential removes this degeneracy.

Exercise 12.10. Show that the addition of a small $1/r^2$ term to the Coulomb potential removes the degeneracy of states with different ℓ. The energy levels are still given by a Balmer-like formula (12.68), but n differs from an integer by an ℓ-dependent quantity, the *quantum defect* in the terminology of one-electron (alkali) spectra.

The new constant of the motion is the quantum-mechanical analogue of the *Runge-Lenz* vector (apparently first introduced by Laplace):[4]

$$\mathbf{K} = \frac{1}{2me^2}(\mathbf{L} \times \mathrm{p} - \mathrm{p} \times \mathbf{L}) + \frac{\mathbf{r}}{r} \tag{12.74}$$

This is a vector operator that commutes with the Hamiltonian for the hydrogen atom,

$$H = \frac{p^2}{2m} - \frac{e^2}{r} \tag{12.75}$$

and has the properties

$$\mathbf{L} \cdot \mathbf{K} = \mathbf{K} \cdot \mathbf{L} = 0 \tag{12.76}$$

Exercise 12.11. Show that $\mathbf{K}$ satisfies the condition (11.19) for a vector operator, that it commutes with the Hamiltonian (12.75), and that (12.76) holds.

For our purposes, the crucial property of the components of $\mathbf{K}$ is that they satisfy the commutation relations

$$\begin{aligned} [K_x, K_y] &= i\hbar\left(-\frac{2H}{me^4}\right)L_z, \quad [K_y, K_z] = i\hbar\left(-\frac{2H}{me^4}\right)L_x, \\ [K_z, K_x] &= i\hbar\left(-\frac{2H}{me^4}\right)L_y \end{aligned} \tag{12.77}$$

[4]See Goldstein (1980), p. 102.

From this relation it is seen that the vector operator

$$\mathbf{A} = \sqrt{-\frac{me^4}{2E}}\,\mathbf{K} \tag{12.78}$$

defined in the subspace of bound-state vectors with energy $E < 0$ satisfies the commutation relations

$$[A_x, A_y] = i\hbar L_z, \quad [A_y, A_z] = i\hbar L_x, \quad [A_z, A_x] = i\hbar L_y \tag{12.79}$$

Exercise 12.12. Check the commutation relations (12.77).

An important identity is obtained from the definition (12.74) through some operator manipulations:

$$\mathbf{K}^2 = \frac{2H}{me^4}(\mathbf{L}^2 + \hbar^2) + I \tag{12.80}$$

Restricting ourselves to the subspace of a definite (negative) bound-state energy E, we may write this, according to (12.78), as

$$\mathbf{A}^2 + \mathbf{L}^2 + \hbar^2 = -\frac{me^4}{2E} \tag{12.81}$$

Finally, it is convenient to introduce the operators

$$\mathbf{J}_1 = \frac{1}{2}(\mathbf{L} + \mathbf{A}) \quad \text{and} \quad \mathbf{J}_2 = \frac{1}{2}(\mathbf{L} - \mathbf{A}) \tag{12.82}$$

each of which satisfies the angular momentum commutation relations. They have the property

$$[\mathbf{J}_1, \mathbf{J}_2] = 0 \quad \text{and} \quad \mathbf{J}_1^2 = \mathbf{J}_2^2 \tag{12.83}$$

Exercise 12.13. Prove that $\mathbf{J}_1$ and $\mathbf{J}_2$ satisfy the commutation relations (11.24) and the conditions (12.83).

Equation (12.81) is then transformed into

$$2(\mathbf{J}_1^2 + \mathbf{J}_2^2) + \hbar^2 = -\frac{me^4}{2E} \tag{12.84}$$

Any state in the subspace spanned by the energy eigenstates corresponding to an eigenvalue E must be an eigenstate of $\mathbf{J}_1^2 = \mathbf{J}_2^2$. From Section 11.2 we know that this operator assumes the eigenvalues $j(j+1)\hbar^2$, where j can be any nonnegative integer or half-integer. From (12.84) we therefore deduce that

$$E = -\frac{me^4}{2\hbar^2(2j+1)^2} \tag{12.85}$$

which is identical with the standard form (12.68) for the energy levels of the nonrelativistic hydrogen atom, if we identify the positive integer $2j + 1$ with the principal quantum number n.

The two commuting "angular momentum" operators $\mathbf{J}_1$ and $\mathbf{J}_2$ can be linked to "rotations" in a four-dimensional Euclidean space. The invariance of the Ham-

iltonian for the Coulomb interaction under these "rotations" signals the new symmetry that accounts for the degeneracy of the energy levels of the hydrogen atom with different ℓ values. This symmetry also explains why it is possible to separate the corresponding Schrödinger equation in parabolic as well as spherical coordinates. We will take advantage of this property in discussing the positive-energy, unbound or scattering, eigenstates in Section 13.8. In the meantime, we return in the next section to the hydrogenic energy eigenfunctions in spherical polar coordinates.

6. ***The Bound-State Energy Eigenfunctions for the Coulomb Potential.*** This section summarizes the most important properties of the radial bound-state wave functions for the attractive Coulomb potential. These functions can be expressed in terms of confluent hypergeometric functions, as can be seen when the value given by (12.65) for ρ_0 is put into (12.64):

$$\frac{a_{k+1}}{a_k} = \frac{2(k - N)}{(k + 1)(k + 2\ell + 2)} \tag{12.86}$$

The confluent hypergeometric function has already been defined in (8.87) as

$${}_1F_1(a; c; z) = 1 + \frac{a}{c}\frac{z}{1!} + \frac{a(a + 1)}{c(c + 1)}\frac{z^2}{2!} + \cdot \tag{12.87}$$

Comparing its coefficients with (12.86), we see that

$$w(\rho) \propto {}_1F_1(-N; 2\ell + 2; 2\rho) \tag{12.88}$$

This can also be seen by comparing the differential equation

$$z\frac{d^2w}{dz^2} + (c - z)\frac{dw}{dz} - aw = 0 \tag{12.89}$$

which $w = {}_1F_1(a; c; z)$ satisfies, with the radial equation (12.62) if the latter by the use of (12.67) is cast in the form

$$\frac{\rho}{2}\frac{d^2w}{d\rho^2} + (\ell + 1 - \rho)\frac{dw}{d\rho} + Nw = 0 \tag{12.90}$$

The complete normalized wave function can be obtained only if we know the value of the normalization integral for the confluent hypergeometric functions. It can be shown that[5]

$$\int_0^\infty e^{-x}x^c[{}_1F_1(a; c; x)]^2\,dx = \frac{(c - 2a)(-a)![(c - 1)!]^2}{(c - a - 1)!} = \frac{(c - 2a)\Gamma(1 - a)[\Gamma(c)]^2}{\Gamma(c - a)} \tag{12.91}$$

Since the spherical harmonics are normalized to unity, the complete eigenfunction $\psi(r, \theta, \varphi)$ of (12.44) is normalized to unity if

$$\int_0^\infty |u(r)|^2\,dr = 1$$

[5]Bethe and Salpeter (1957), Section 3.

Since

$$\rho = \kappa r, \qquad a = -N, \qquad c = 2\ell + 2,$$
$$c - 2a = 2(\ell + 1 + N) = 2n, \qquad c - a - 1 = n + \ell$$

we can write the hydrogenic eigenfunctions, normalized to unity, as

$$\psi_{n,\ell,m}(r, \theta, \varphi) = \frac{e^{-\kappa r}(2\kappa r)^{\ell}}{(2\ell + 1)!}\left[(2\kappa)^3 \frac{(n + \ell)!}{2n(n - \ell - 1)!}\right]^{1/2} {}_1F_1(-n + \ell + 1; 2\ell + 2; 2\kappa r)Y_{\ell}^{m}(\theta, \varphi) \tag{12.92}$$

The radial eigenfunctions $R_{n\ell}$ for the three lowest energy eigenvalues are plotted in Figure 12.1 as a function of Zr/a.

Exercise 12.14. Calculate the radial hydrogen energy eigenfunctions for $n = 4$ and 5 explicitly, and sketch them graphically.[6]

The polynomial ${}_1F_1(-N; c; z)$ of degree N and positive integral argument c is proportional to the *associated Laguerre polynomials* of classical mathematical physics. The connection is established by the relation

$${}_1F_1(a; c; z) = \frac{\Gamma(1 - a)\Gamma(c)}{[\Gamma(c - a)]^2} L_{-a}^{c-1}(z) \tag{12.93}$$

An elementary definition of the associated Laguerre polynomial is

$$L_{q-p}^{p}(z) = (-1)^p \frac{d^p}{dz^p} L_q(z) \tag{12.94}$$

and the Laguerre polynomial of order q is

$$L_q^0(z) \equiv L_q(z) = e^z \frac{d^q}{dz^q}(e^{-z}z^q) \tag{12.95}$$

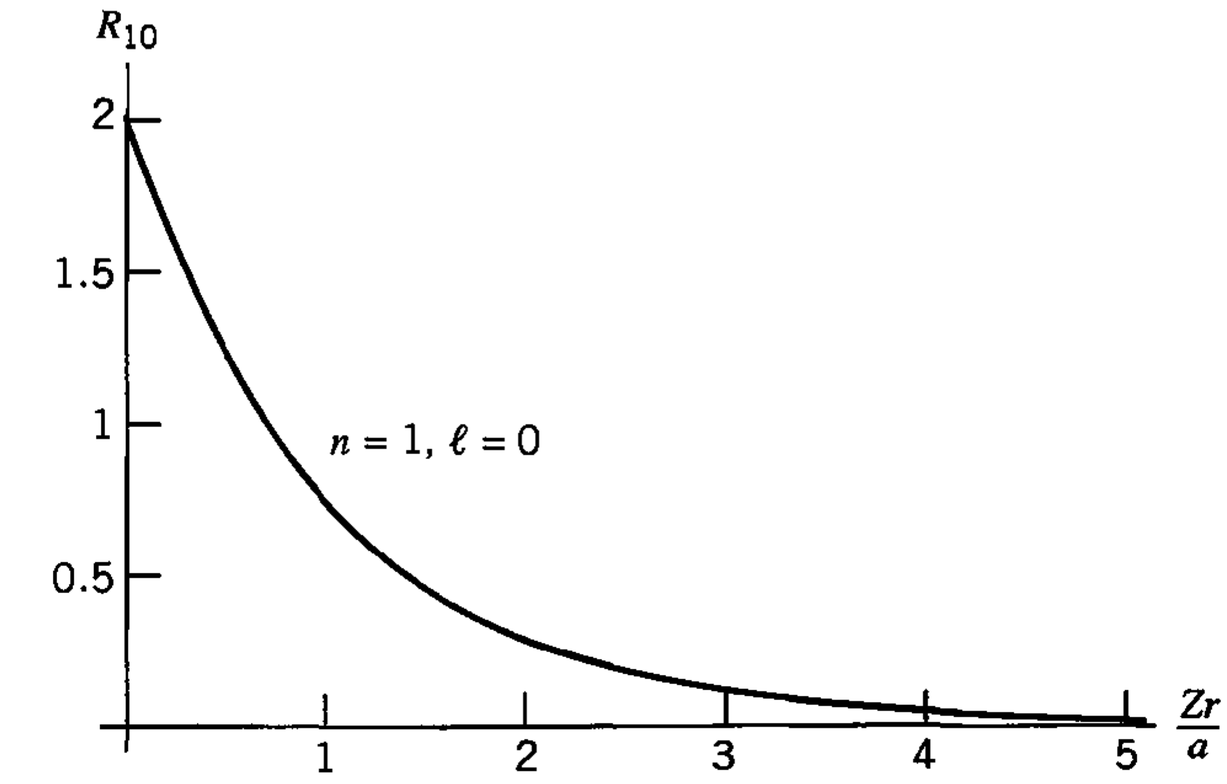
Exercise 12.15. Show that the generating function for the associated Laguerre polynomials (for $|s| \leq 1$) is

$$\frac{\exp\left(-\dfrac{sz}{1 - s}\right)}{(1 - s)^{p+1}} = \sum_{n=0}^{\infty} \frac{L_n^p(z)}{(n + p)!} s^n \tag{12.96}$$

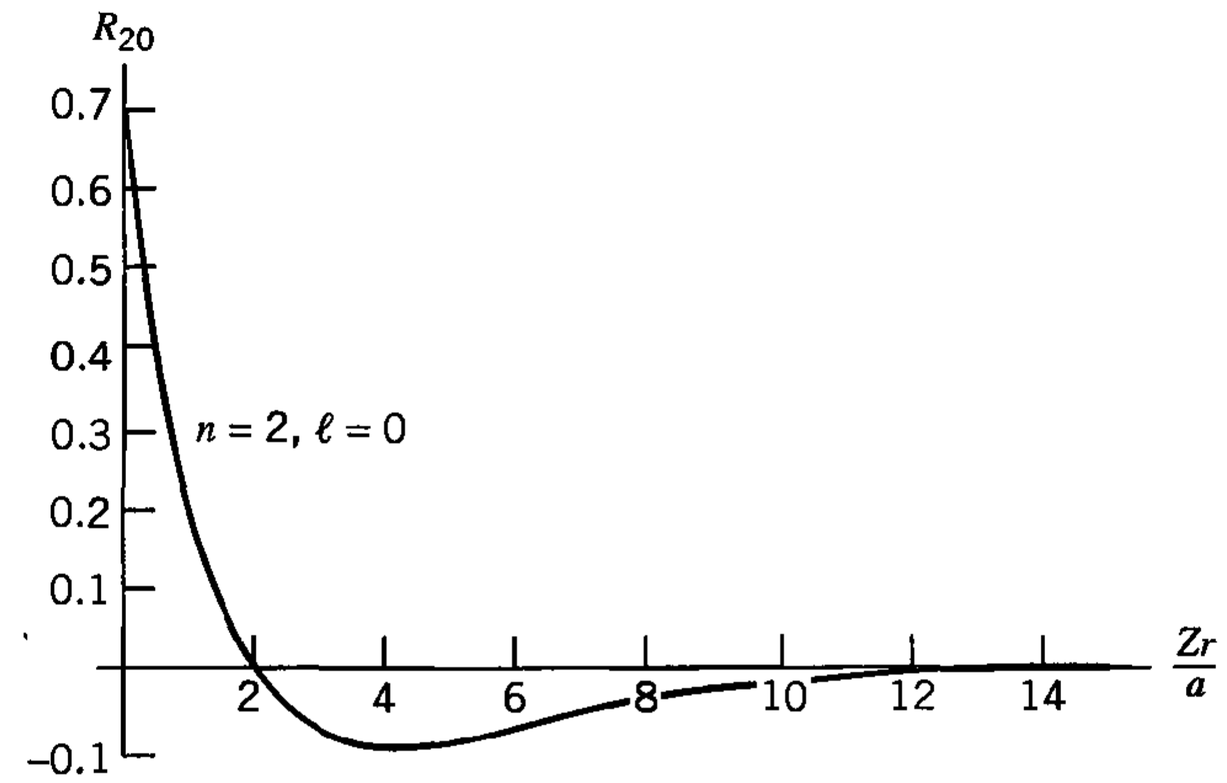
A few comments concerning the properties of the hydrogenic eigenfunctions are appropriate. The wave function possesses a number of nodal surfaces. These are surfaces on which $\psi = 0$. For these considerations, it is customary to refer instead of (12.92) to the *real-valued* eigenfunctions

$$\psi_{n,\ell,m} \propto r^{\ell}e^{-\kappa r}\, {}_1F_1(-n + \ell + 1; 2\ell + 2; 2\kappa r)P_{\ell}^{m}(\cos\theta)\begin{cases}\cos m\varphi \\ \sin m\varphi\end{cases} \tag{12.97}$$

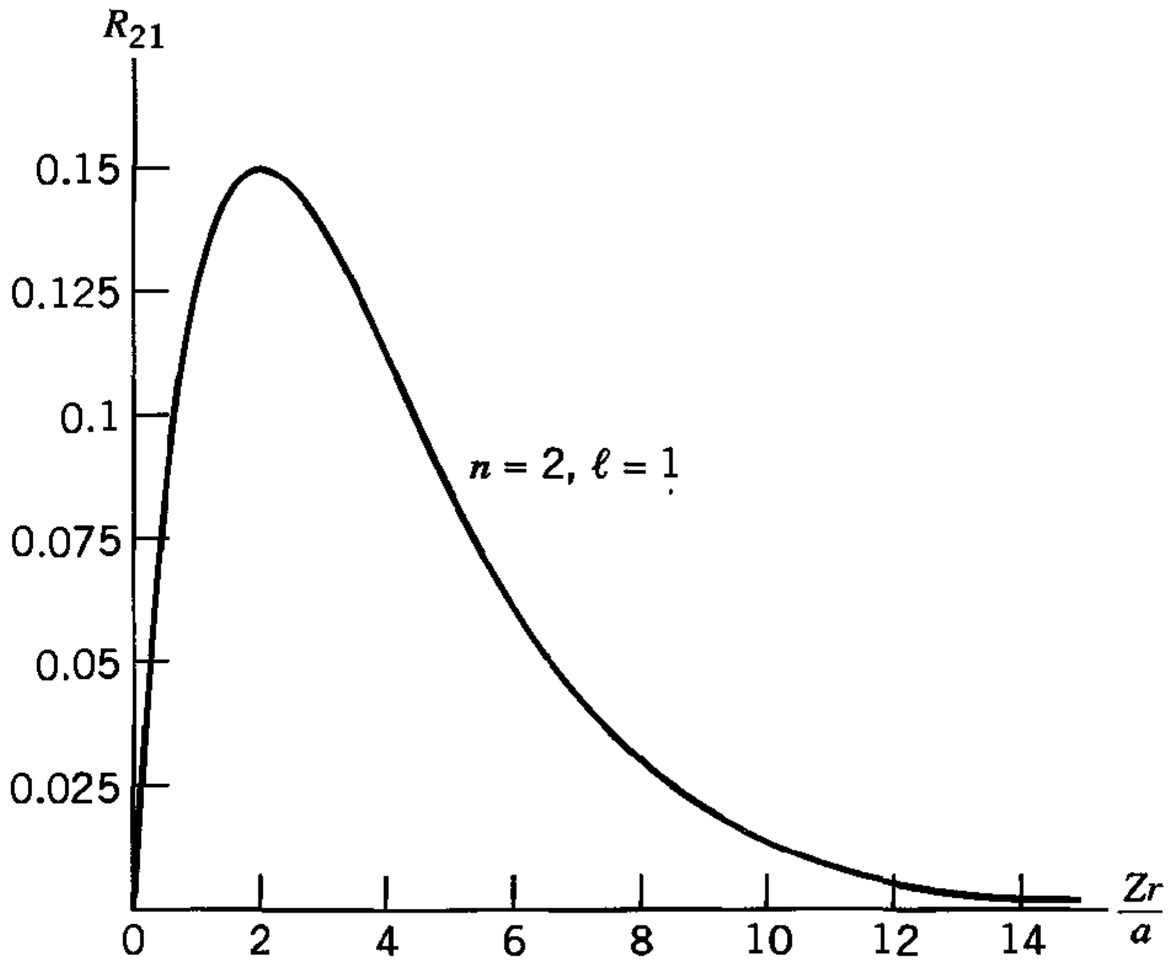
[6]Brandt and Dahmen (1985), Section 12.4. For good early graphic representations see White (1931). See also Pauling and Wilson, Jr. (1935), Section 5.21.



(*a*)



(*b*)



(*c*)

Figure 12.1. Radial bound-state energy eigenfunctions of the hydrogenic atom for $n = 1$, 2, and 3. The radial coordinate r is scaled in units of a/Z. The wave functions are normalized as $\int_0^\infty [R_{n\ell}(r)]^2 r^2 \, dr = 1$.

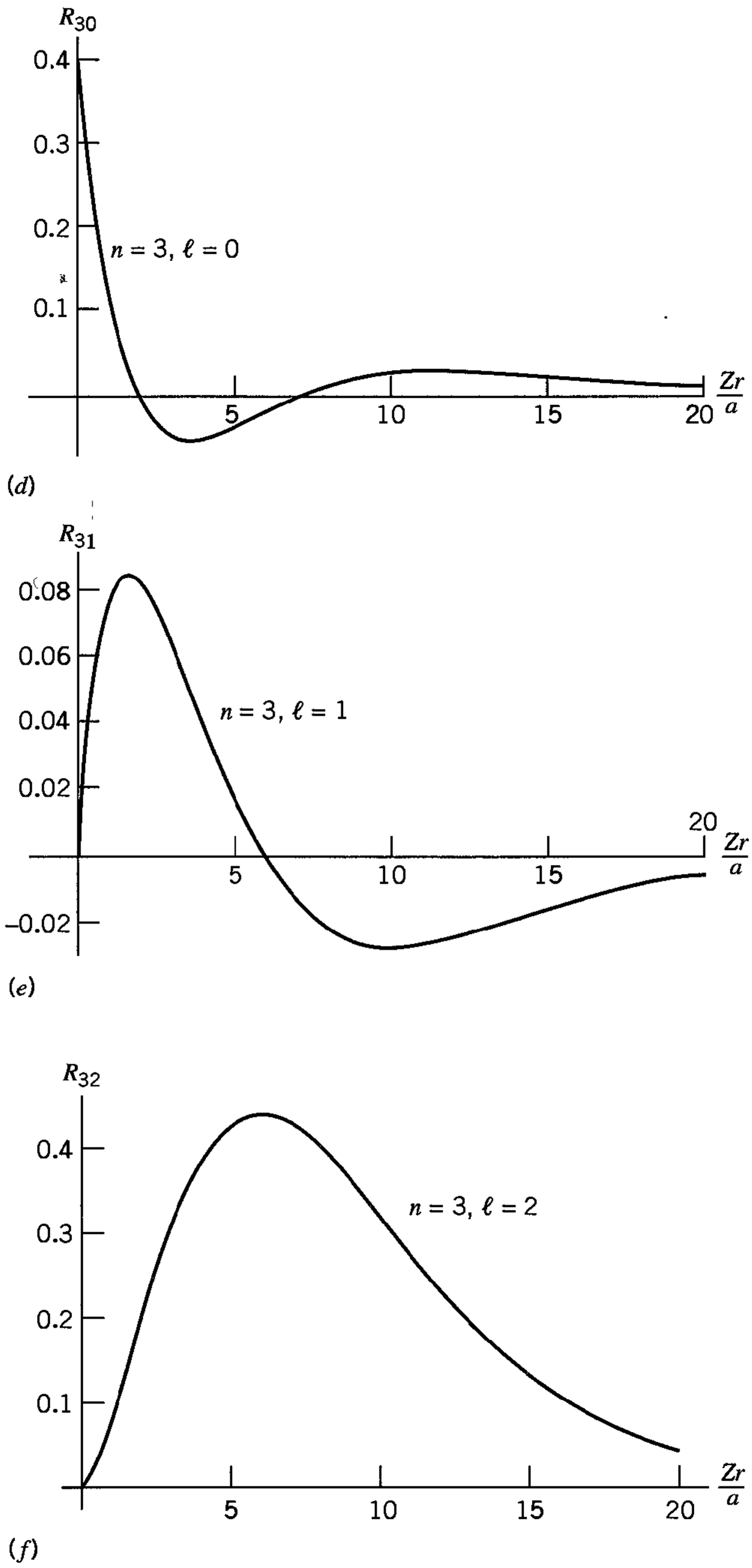


Figure 12.1. (continued)

There are $\ell - m$ values of θ for which $P_\ell^m(\cos\theta)$ vanishes; $\cos m\varphi$ and $\sin m\varphi$ vanish at m values of the azimuth, and the confluent hypergeometric function ${}_1F_1$ vanishes at $n - \ell - 1$ values of r. For $\ell \neq 0$, r^ℓ has a node at $r = 0$. Hence, except in S states, the total number of nodal surfaces is n, if $r = 0$ is counted as a surface.

Important consequences follow from the fact that the wave function vanishes at the origin except when $\ell = 0$ (S states). For instance, the capture by a nucleus of an atomic electron or any other orbiting negatively charged particle can occur with appreciable probability only from a level with $\ell = 0$, because these are the only states for which ψ has a nonzero finite value at the position of the nucleus. Similarly, in the phenomenon of internal conversion an atomic s electron interacts with the nucleus and is imparted enough energy to be ejected from the atom in an autoionizing transition.

The quantum mechanical significance of the Bohr radius $a = \hbar^2/me^2$, can be appreciated by observing that the wave function for the $1s$ ground state is

$$\boxed{\psi_{1,0,0}(r,\ \theta,\ \varphi) = \left(\frac{Z^3}{\pi a^3}\right)^{1/2} \exp\left(-\frac{Zr}{a}\right)} \tag{12.98}$$

The expectation value of r in this state is

$$\langle r\rangle = 4\left(\frac{Z}{a}\right)^3 \int_0^\infty \exp\left(-\frac{2Z}{a}\,r\right) r^3 dr = \frac{3}{2}\,\frac{a}{Z} \tag{12.99}$$

The maximum of the probability density for finding the particle in the ground state with a radial separation r from the nucleus, i.e., the maximum of the function

$$\left|\exp\left(-\frac{Z}{a}\,r\right)\right|^2 r^2$$

is located at a/Z.

Exercise 12.16. Evaluate the width of the probability distribution for the radial coordinate r in the ground state of the hydrogenic atom by calculating the uncertainty Δr.

The Bohr radius a is inversely proportional to the mass of the particle that moves around the nucleus. Hence, a muon, pion, or kaon in a mesic atom, or an antiproton that has been captured by a nucleus, is much closer to the nucleus than an electron is in an ordinary atom. The finite size of the nucleus will thus be expected to affect the discrete energy levels of exotic atoms appreciably, whereas nuclear size effects are very small for electronic states in the bound states of ordinary atoms.

Many other corrections, of course, must be taken into account when comparing the simple Balmer formula (12.68) with the amazingly accurate results of modern atomic spectroscopy. Most obviously, we must correct the error made in assuming that the nucleus is infinitely massive and therefore fixed. Since for central forces the actual two-body problem can, in the nonrelativistic approximation, be replaced by an effective one-body problem, if we substitute the reduced mass $m_1m_2/(m_1 + m_2)$ for m, this correction can be applied accurately and without difficulty. It gives rise to small but significant differences between the spectra of hydrogen and deuterium. For a positronium atom, composed of an electron and a positron, which have equal

masses, all energy levels are changed by a factor 1/2 compared to hydrogen. Further, and often more significant, corrections are due to the presence of the electron spin and the high speed of the electron, which necessitate a relativistic calculation (Section 24.9); hyperfine structure effects arise from the magnetic properties of the nucleus; and finally, there are small but measurable effects owing to the interactions between the electron and the electromagnetic field (Lamb shift).[7] The theory of some of these effects will be discussed later in this book; others lie outside its scope. But all are overshadowed in magnitude by the basic gross structure of the spectrum as obtained in this chapter by the application of nonrelativistic quantum mechanics to a charged particle subject to the Coulomb potential.

Problems

1. Compute (or obtain from mathematical tables) and plot the 10 lowest energy eigenvalues of a particle in an infinitely deep, spherically symmetric square well, and label the states by their appropriate quantum numbers.
2. If the ground state of a particle in a spherical square well is just barely bound, show that the well depth and radius parameters V_0 and a are related to the binding energy by the expansion

$$\frac{2mV_0a^2}{\hbar^2} = \frac{\pi^2}{4} + 2\kappa a + \left(1 - \frac{4}{\pi^2}\right)(\kappa a)^2 + \cdots$$

where

$$\hbar\kappa = \sqrt{-2mE}$$

The deuteron is bound with an energy of 2.226 MeV and has no discrete excited states. If the deuteron is represented by a nucleon, with reduced mass, moving in a square well with $a = 1.5$ fermi, estimate the depth of the potential.

3. Given an attractive central potential of the form

$$V(r) = -V_0e^{-r/a}$$

solve the Schrödinger equation for the S states by making the substitution

$$\xi = e^{-r/2a}$$

Obtain an equation for the eigenvalues. Estimate the value of V_0, if the state of the deuteron is to be described with an exponential potential (see Problem 2 for data).

4. Show that, if a square well just binds an energy level of angular momentum ℓ ($\neq 0$), its parameters satisfy the condition

$$j_{\ell-1}\left(\sqrt{\frac{2mV_0a^2}{\hbar^2}}\right) = 0$$

(Use recurrence formulas for Bessel functions from standard texts.)

5. Assuming the eigenfunctions for the hydrogen atom to be of the form $r^\beta e^{-\alpha r}\, Y_\ell^m$, with undetermined parameters α and β, solve the Schrödinger equation. Are all eigenfunctions and eigenvalues obtained this way?
6. Apply the WKB method to an attractive Coulomb potential, and derive an approximate formula for the S-state energy levels of hydrogen.

[7]Gross (1993). For a detailed account of the spectrum of hydrogenic and helium-like atoms, see Bethe and Salpeter (1957).

7. Compute the probability that the electron in a hydrogen atom will be found at a distance from the nucleus greater than its energy would permit on the classical theory. Make the calculation for the $n = 1$ and 2 levels.

8. Calculate the probability distribution for the momentum of the electron in the ground state of a hydrogen atom. Obtain the expectation value of p_x^2 from this or from the virial theorem. Also calculate $\langle x^2 \rangle$ from the ground state wave function, and verify the uncertainty relation for this state.

9. Solve the Schrödinger equation for the three-dimensional isotropic harmonic oscillator, $V = (1/2)m\omega^2 r^2$, by separation of variables in Cartesian and in spherical polar coordinates. In the latter case, assume the eigenfunctions to be of the form

$$\psi(r, \theta, \varphi) = r^\ell \exp\left(-\frac{m\omega}{2\hbar} r^2\right) f(r) Y_\ell^m(\theta, \varphi)$$

and show that $f(r)$ can be expressed as an associated Laguerre polynomial (or a confluent hypergeometric function) of the variable $m\omega r^2/\hbar$ with half-integral indices. Obtain the eigenvalues and establish the correspondence between the two sets of quantum numbers. For the lowest two energy eigenvalues, show the relation between the eigenfunctions obtained by the two methods.

10. For the isotropic harmonic oscillator of Problem 9, obtain a formula for the degree of degeneracy in terms of the energy. For large energies (large quantum numbers), compare the density of energy eigenstates in the oscillator and in a cubic box.

11. Starting with the radial equation

$$-\frac{\hbar^2}{2m}\frac{d^2u}{dr^2} + \frac{\hbar^2\ell(\ell+1)}{2mr^2}u - \frac{Ze^2}{r}u = Eu$$

for the hydrogenic atom, show that the transformation

$$r = \alpha\bar{r}^2, \qquad u = \sqrt{\bar{r}}\,\bar{u},$$

produces an equation for $\bar{u}(\bar{r})$ that, with appropriate choices of the constants, is equivalent to the radial equation for the isotropic oscillator. Exhibit the relation between the energy eigenvalues and the radial quantum numbers for the two systems.

12. The initial ($t = 0$) state of an isotropic harmonic oscillator is known to be an eigenstate of L_z with eigenvalue zero and a superposition of the ground and first excited states. Assuming that the expectation value of the coordinate z has at this time its largest possible value, determine the wave function for all times as completely as possible.

13. Solve the energy eigenvalue problem for the two-dimensional isotropic harmonic oscillator. Assume that the eigenfunctions are of the form

$$\psi = \rho^\ell \exp\left(-\frac{m\omega}{2\hbar}\rho^2\right) f(\rho) \exp(\pm i\ell\varphi)$$

where ρ and φ are plane polar coordinates and ℓ is a nonnegative integer. Show that $f(\rho)$ can be expressed as an associated Laguerre polynomial of the variable $m\omega\rho^2/\hbar$ and determine the eigenvalues. Solve the same problem in Cartesian coordinates, and establish the correspondence between the two methods. Discuss a few simple eigenfunctions.

14. Apply the variational method to the ground state ($\ell = 0$) of a particle moving in an attractive central potential $V(r) = Ar^n$ (integer $n \geq -1$), using

$$R(r) = e^{-\beta r}$$

as a trial wave function with variational parameter β. For $n = -1$ and $+2$, compare the results with the exact ground state energies.

15. Apply the variational method to the ground state ($\ell = 0$) of a particle moving in an attractive (Yukawa or screened Coulomb or Debye) potential

$$V(r) = -V_0 \frac{e^{-r/a}}{r/a} \qquad (V_0 > 0)$$

Use as a trial function

$$R(r) = e^{-\gamma r/a}$$

with an adjustable parameter γ. Obtain the "best" trial wave function of this form and deduce a relation between γ and the strength parameter $2mV_0\, a^2/\hbar^2$. Evaluate γ and calculate an upper bound to the energy for $2mV_0\, a^2/\hbar^2 = 2.7$.

Are there any excited bound states?

Show that in the limit of the Coulomb potential ($V_0 \to 0$, $a \to \infty$, $V_0 a$ finite) the correct energy and wave function for the hydrogenic atom are obtained.

16. Using first-order perturbation theory, estimate the correction to the ground state energy of a hydrogenic atom due to the finite size of the nucleus. Under the assumption that the nucleus is much smaller than the atomic radius, show that the energy change is approximately proportional to the nuclear mean square radius. Evaluate the correction for a uniformly charged spherical nucleus of radius R. Is the level shift due to the finite nuclear size observable? Consider both electronic and muonic atoms.

17. An electron is moving in the Coulomb field of a point charge Ze, modified by the presence of a uniformly charged spherical shell of charge $-Z'e$ and radius R, centered at the point charge. Perform a first-order perturbation calculation of the hydrogenic $1S$, $2S$, and $2P$ energy levels. For some representative values of $Z = Z'$, estimate the limit that must be placed on R so that none of the lowest three energy levels shift by more than 5 percent of the distance between the unperturbed first excited and ground state energy levels.

CHAPTER 13

Scattering

Much of what we know about the forces and interactions in atoms and subatomic particles has been learned from collision experiments, in which atoms in a target are bombarded with beams of particles. Particles that are scattered by the target atoms are subsequently detected by devices that may give us the intensity as a function of the scattering angle (and possibly of the energy of the scattered particles, if *inelastic* processes are also involved). We begin the discussion with a general introduction to the concept of a cross section, since it constitutes the meeting ground between experimentalists and theoreticians for all varieties of collisions. We then establish the connection between calculated quantum mechanical amplitudes and measured cross sections for collisions in which *elastic scattering* is the dominant process. The stock-in-trade of scattering theory is developed: incoming and outgoing Green's functions, quantum mechanical integral equations, the Born approximation, and partial wave and phase shift analysis. Later, in Chapter 20, we will take a second comprehensive look at collision theory in the context of the general principles of quantum dynamics, to be developed in Chapters 14 and 15.[1]

1. *The Cross Section.*

Some form of collision experiment is the most common tool for probing atomic and subatomic interactions. Collisions of nucleons with nuclei at various energies reveal information about the nuclear forces and the structure of the nucleus. Electrons of high energy, hence short wavelength, are particularly well suited to determine the charge distribution in nuclei, and indeed within nucleons. Electrons and heavier projectiles of low energy are scattered from atoms to obtain data that can serve as input information for calculations of kinetic processes in gases where low-energy collisions predominate. And collisions of hadrons and leptons with protons tell us about interactions of which we have no other direct information. These are just a few examples of the utility of collisions in studying the internal structure of atoms and nuclei and the interactions that govern elementary particles.

Usually, we know the nature of the particles used as projectiles, their momentum, and perhaps their polarization (defined in Chapter 16). The collision between a projectile and a target particle is sometimes referred to as a ***binary*** collision, in order to distinguish it from the kind of interaction that takes place when an incident beam interacts with a large number of target atoms, as happens in diffraction of electrons, neutrons, or even entire atoms from a lattice. The target particle is frequently at rest (or nearly so) before the collision, but its thermal motion cannot always be neglected. In some experiments, the target atoms are in gas jets, with controlled initial velocity, usually at right angles to the incident beam. In *merging* or *colliding* beam experiments, both projectile *and* target particles initially move

[1]Goldberger and Watson (1964), Newton (1982), and Taylor (1972) are treatises on collision theory. Mott and Massey (1965) and Bransden (1983) apply the theory to atomic collision problems.

along the same direction, either parallel or antiparallel, and with controlled initial velocity.

Experimentalists typically measure intensities and yields of certain collision or reaction products that result from the interaction of incident particles or waves with target "atoms," upon which the projectiles impinge, although the "atoms" can be any kind of object (particle or wave) that affects the incident beam. The comparison between measurements and theoretical predictions is made in terms of the *cross section* for the particular process under consideration. Broadly speaking, in physics, a cross section measures the size of the effective area that must be hit, if an incident charged or neutral projectile particle is to cause a certain specified effect in a given target particle, which may also be either neutral (stationary target) or charged (colliding beams). The relative velocity (collision energy) of the two interacting particles is usually specified at incidence. The experiment illustrated in Figure 13.1 involves a *collimated* homogeneous beam of monoenergetic particles moving in the same sharply defined direction toward the scatterer from a great distance. Good resolution in incident velocity (magnitude as well as direction) is prized in all collision experiments. The width of the beam is determined by slits, which, though quite narrow from an experimental point of view, are nevertheless very wide compared with the cross section. Experimentally, in the interest of securing "good statistics," it is desirable to employ beams with high incident intensity, or *luminosity*. Yet the beam density must be low enough that it can safely be assumed that the incident particles do not interact with one another. A precise knowledge of the number of projectiles, either per pulse or per unit time, is essential.

The interpretation of most experiments requires accurate knowledge of the physical properties of the target, especially the number of target particles that are exposed to the incident beam. If the target is solid, it is assumed (unless otherwise specified) that the beam is incident normal to the surface of the target.

Prototypically, a cross section measurement is similar to the ("Monte Carlo") determination of the size of the bull's eye on a dart board by recording the successful hits among randomly thrown missiles directed blindly toward the target, provided we know the average number of projectiles striking the target per unit area. However, in the atomic and subatomic domain the cross section, which is a physical property of the target particle and its interaction with the projectile, generally bears no direct relation to the geometric size of the target particle.

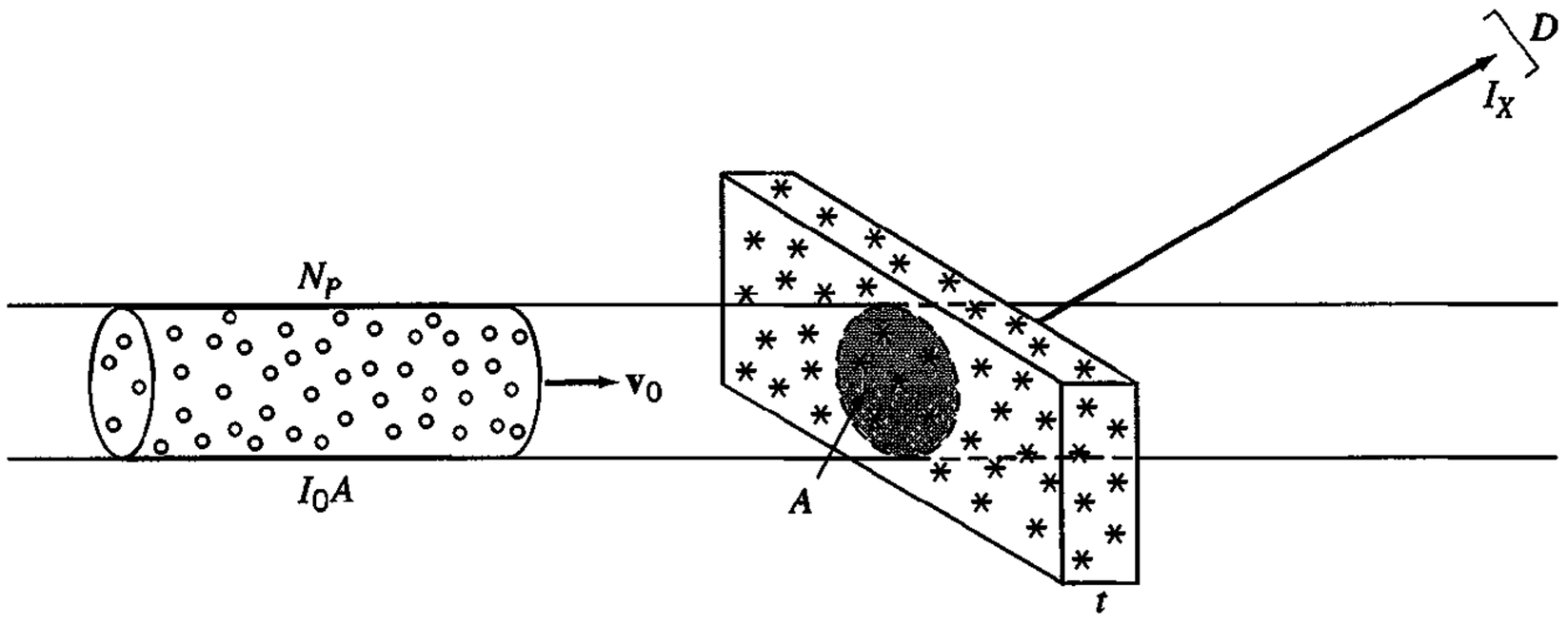


Figure 13.1. Sketch of a scattering experiment. A pulse of $N_P = I_0A$ projectile particles is incident from the left with velocity $\mathbf{v}_0$. The beam intersects an area A of the target (thickness t and particle density n) and encounters $N_T = ntA$ target particles. The detector D registers I_X outcomes X, and the corresponding cross section is $\sigma_X = I_X/N_TI_0$.

Let us suppose that we are interested in determining the likelihood of an outcome X in a binary collision. In measurements, the outcome X might refer to any of a number of physical processes, as for example:

(a) *Scattering*, with or without loss of energy, into a well-defined solid angle tilted at a specified direction with regard to the direction of incidence. This direction is defined by the position of the detector. If the collision is not elastic, the energetically degraded projectile and the excitation of the target may be observed simultaneously by a coincidence measurement of some outgoing radiation or particle emission. Sometimes the recoil of the target particle is observed.

(b) *Absorption* of the incident projectile by the target, measured either through the attenuation of the incident beam or through observation of reaction products, e.g., the creation of new particles, the emission of radiation or particles, or chemical changes in the composition of the target.

Let σ_X stand for the cross section that corresponds to the outcome X. All cross sections have the dimension (units) of an area, although differential cross sections may be defined and quoted as "area per unit measure of X." Let A denote the area of overlap of the beam of projectiles and the assemblage of target particles, measured at right angles to the relative velocity of the projectile-target system (Figure 13.1). (In the case of a stationary target, this is just the direction of the incident beam.) We assume that the experimental arrangement allows us to choose $A \gg \sigma_X$. The dimensionless ratio σ_X/A is the *probability* of the outcome X in a binary collision of a projectile particle with a target particle, provided that the presence of the other projectile and target particles does not affect the binary collision. Under these conditions if, within the common overlap area A, the assemblage of projectiles contains N_P particles and the target assemblage contains N_T particles, the total number I_X of detectable outcomes with property X is the product:

$$\boxed{I_X = \frac{\sigma_X}{A} N_P N_T} \tag{13.1}$$

Since experiments have to be conducted within a reasonable length of time, in laboratory practice steady and stable beams are characterized by the number of particles *per unit time*. If the product $N_P N_T$ is interpreted as the number of encounters between projectile and target particles per unit time, I_X measures the *rate* or *yield* of emerging particles with outcome property X per unit time. If the detector is "ideal," i.e., 100 percent efficient, the quantity I_X is known as the *counting rate* for the process that is being observed. Expression (13.1) is directly applicable to the colliding beam geometry in particle accelerators, and the quantity $L = N_P N_T/A$ is generally referred to as the *luminosity* (in units of $\text{sec}^{-1}\ \text{cm}^{-2}$). The magnitude of the luminosity is critical if reasonable counting rates are to be achieved. (In typical high-energy particle accelerators, luminosities are of the order $10^{32}\ \text{sec}^{-1}\ \text{cm}^{-2}$ for colliding beams.)

In the more common geometry of stationary targets exposed to projectile beams, Eq. (13.1) is usually written in a less symmetric form. The ratio $I_0 = N_P/A$, the incident *intensity*, is the number of particles incident on the target per unit area and unit time. The number N_T of target particles exposed to the beam can be expressed as

$$N_T = ntA \tag{13.2}$$

where n is the number of target particles per unit volume and t is the thickness of the target. If the binary collision between a projectile particle and any target atom is unaffected by the presence of the other target atoms, the product $\sigma_X nt$ is the fraction of the total area A over which beam and target particles interact that is effective in producing the outcome X. This must be less than unity:

$$N_T \frac{\sigma_X}{A} = nt\sigma_X < 1$$

setting an upper limit to the thickness t. The target is said to be "thin" if $t \ll 1/n\sigma_X$. For a thin target, Eq. (13.1) can be reexpressed as

$$nt\sigma_X = \frac{I_X}{AI_0} = \frac{\text{counting rate for detectable particles with property } X}{\text{number of particles incident on the target per unit time}} \tag{13.3}$$

For the cross section σ_X we thus obtain the formula

$$\boxed{\sigma_X = \frac{I_X}{ntAI_0}} \tag{13.4}$$

The outcome X in a collision experiment may be a *discrete* physical characteristic, such as the "yes/no" choice in a nuclear reaction, which either does or does not occur. Similarly, a *total* scattering cross section measures only the occurrence of the scattering process, and not any specific feature of the scattered particle. On the other hand, X may stand for a physical process that is described by one or more variables x with *continuous* values, such as scattering in a particular spatial direction or with an energy loss that lies in a certain continuous range. The cross section is a measure of the probability that the outcome lies in the interval between x and $x + dx$ and is usually proportional to dx. It is then customary to write this cross section in the form $d\sigma_X$ and to define the ratio $d\sigma_X/dx$ as the *differential cross section* for the outcome X with value x.

An example is the *differential scattering cross section*, $d\sigma/d\Omega$, corresponding to scattering into a solid angle $d\Omega$, defined as the angle subtended at the target by the detector in a given direction (Figure 13.2). Here, the outcome X is detection of the particles after scattering at a great distance from the scatterer into the solid angle $d\Omega$ in a direction specified by spherical coordinates θ, φ that are related to the direction of incidence, chosen as the polar axis. If the number of scattered particles per unit solid angle is denoted by $I(\theta, \varphi)$, the relation (13.4) becomes the expression for the differential cross section, or *angular distribution*,

$$\boxed{\frac{d\sigma}{d\Omega} = \frac{I(\theta, \varphi)}{ntAI_0}} \tag{13.5}$$

This is the quantity which the experimentalist delivers to the theoretician, who interprets the cross section in terms of probabilities calculated from wave functions.

The outcome X of a collision experiment is defined by the physical arrangement of the measurement. The initial collision parameters, including the relative velocity or kinetic energy of the incident particles, the intensity of the incident beam, and the velocity (if different from zero) of the target atoms, are usually determined and quoted with respect to a laboratory frame-of-reference. Yet, the theoretical analysis is often carried out in an entirely different frame of reference, such as the center-

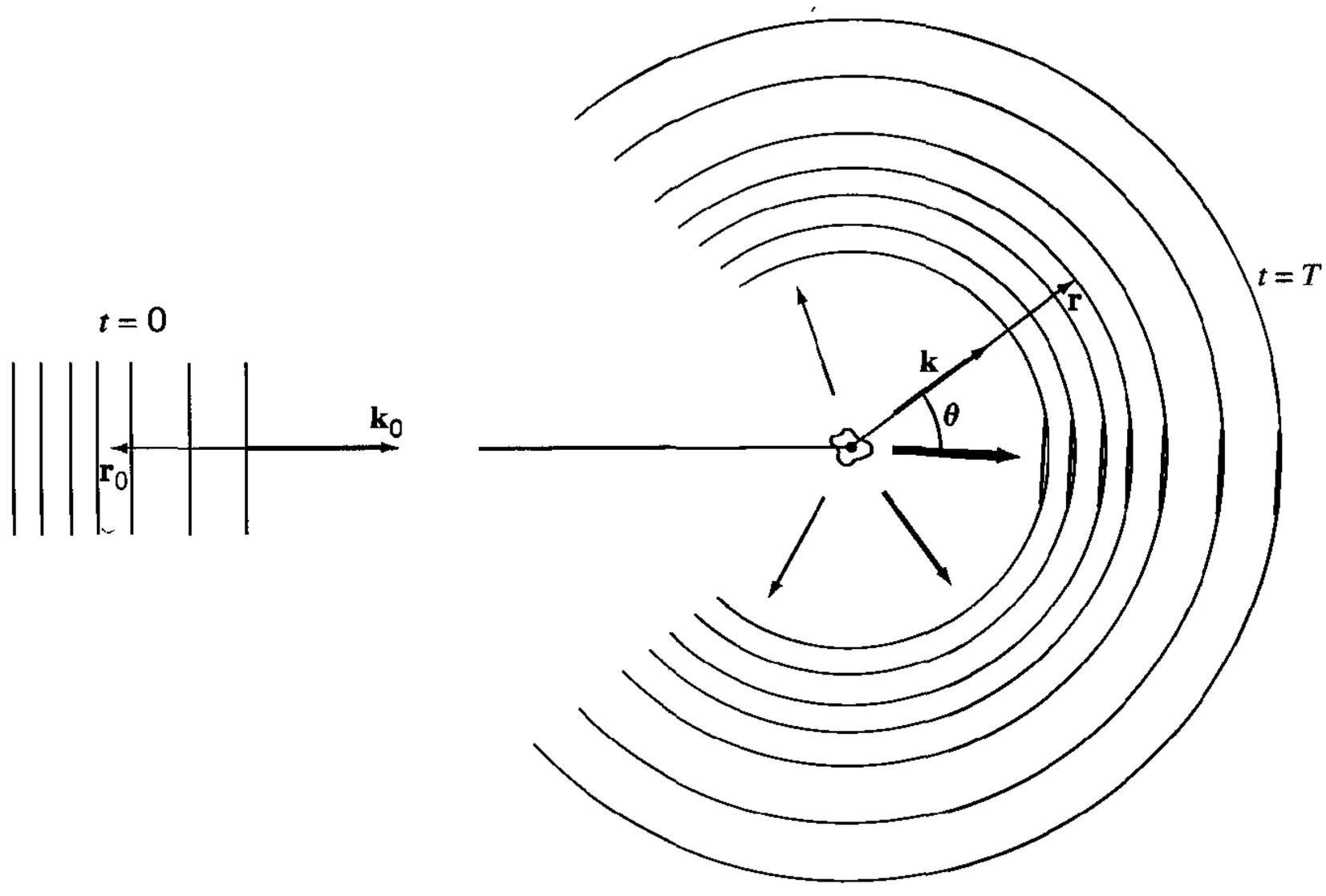


Figure 13.2. Scattering of a wave packet incident from the left with mean momentum $\hbar\mathbf{k}_0$. In elastic scattering, the wave packet is scattered with equal energy but varying amplitudes in different directions (indicated by the thickness of the arrows). In the forward direction, the scattered wave interferes with the advancing incident wave.

of-mass frame in which the total momentum of the system is zero, or with the use of relative coordinates appropriate to the interaction between two particles. The cross section depends on the collision parameters (incident energy, polarization, etc.) and the particular variables (scattering angle, energy loss, etc.) which quantify the outcome X. If these variables are denoted by x in the theoretically preferred frame-of-reference, a transformation to the corresponding variables, x', in the laboratory frame must be carried out. Formally, the differential cross-sections in the two frames are related by

$$\frac{d\sigma_X}{dx'} = \frac{d\sigma_X}{dx}\frac{dx}{dx'} \tag{13.6}$$

This relation cautions us that the two differential cross sections do not have the same value. A numerical example at the end of this section illustrates this for the angular distribution.

We now turn to the theoretical calculation of a scattering cross section. In an idealized scattering process, a *single* fixed scattering center or target particle is bombarded by particles incident along the z axis. It is assumed that the effect of the scattering center on the particles can be represented by a potential energy $V(\mathbf{r})$ which is appreciably different from zero only within a finite region. Although this assumption excludes as common a long-range force as the Coulomb field, represented by a potential proportional to $1/r$ and observable as Rutherford scattering, this limitation is not severe. In actual fact, the Coulomb field is screened at large distances by the presence of electrons in atoms and by other particles, and for large r the potential falls off faster than $1/r$.

In the classical limit, each incident particle can be assigned an *impact parameter*, i.e., a distance b from the z axis, parallel to which it approaches from infinity,

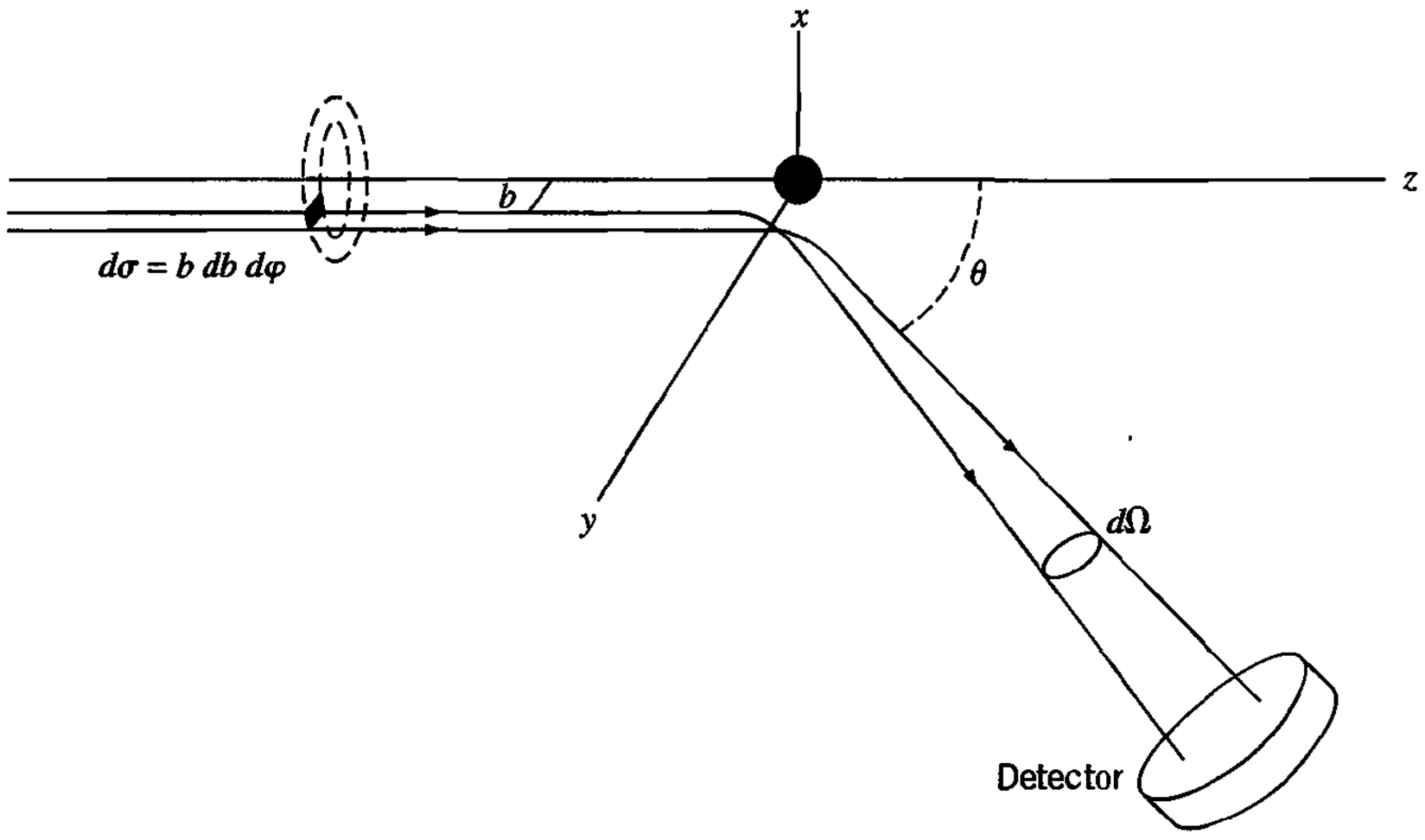


Figure 13.3. Scattering through an angle θ of a particle beam incident at impact parameter b and azimuth angle φ. Classically, the cylinder bounded by $b\ .\ .\ b + db$ and $\varphi\ .\ .\ \varphi + d\varphi$ defines the cross section area $d\sigma$.

and an azimuth angle φ, which together with z define its position in cylindrical coordinates. For particles moving on classical orbits, the final asymptotic direction θ, φ is determined by these initial coordinates and by the incident energy. The cross section $d\sigma$ is simply equal to the size of the area which, when placed at right angles to the incident beam, is traversed by all those orbits that asymptotically end up in the solid angle $d\Omega$ around the direction θ, φ. Figure 13.3 illustrates the point. Obviously, this discussion must be refined if the scattering angle is not a continuous smooth function of the impact parameter or of the azimuth angle of the incident particle, but we ignore such singular situations (e.g., *rainbow scattering*).

If the scattering potential is spherically symmetric, $V = V(r)$, the orbits lie in planes through the center of force, and the scattering becomes independent of the azimuth angle φ. Therefore, classically,

$$d\sigma = b\ db\ d\varphi = b\left|\frac{db}{d\cos\theta}\right||d\cos\theta|\ d\varphi = b\left|\frac{db}{d\cos\theta}\right|\ d\Omega \tag{13.7}$$

from which we see that

$$\frac{d\sigma}{d\Omega} = b\left|\frac{db}{d\cos\theta}\right| = \frac{b}{\sin\theta}\left|\frac{db}{d\theta}\right| \tag{13.8}$$

Hence, the differential cross section is calculable if, for the given energy, we know b as a function of the scattering angle θ. To determine this function from the equation of motion is the problem of classical scattering theory.

Exercise 13.1 In Rutherford scattering of a particle of mass m and incident energy E, with potential energy $V(r) = C/r$ (C constant), the functional relationship between the impact parameter b and the scattering angle θ is given by

$$b = \frac{C}{2E}\cot\frac{\theta}{2}$$

Derive the differential scattering cross section for Rutherford scattering:

$$\frac{d\sigma}{d\Omega} = \frac{Z_1^2 Z_2^2 e^4}{16E^2} \frac{1}{\sin^4(\theta/2)} \tag{13.9}$$

where $Z_1 e$ and $Z_2 e$ are the nuclear charges of the projectile and target particles, respectively.

Equation (13.8) is a good approximation only if the de Broglie wavelength of the incident particles is much smaller than the dimensions of the scattering region. It ceases to be useful when the wave description can no longer be adequately approximated by the geometrical optics limit. As the wavelength increases, quantum features make their appearance, and the quantum uncertainty begins to limit the simultaneous knowledge of, say, p_x, which for a well-collimated beam should be close to zero, and x, which is proportional to the impact parameter. Thus we must learn how to calculate cross sections from wave functions.

The most significant quantum mechanical aspect of scattering is that we are now concerned with unbound particles and the *continuous* part of the energy spectrum. The Schrödinger equation for unbound eigenstates must now be solved, corresponding to positive energy eigenvalues, and the connection between eigenfunctions and measured intensities must be established.

Quantum mechanics represents the particles in the beam by wave packets of various shapes and sizes. We would like to suppose that all the particles in the beam with incident intensity I_0 can be represented by very broad and very long wave packets propagating in the z direction and that, before they reach the neighborhood of the scatterer, these packets can be described simply and approximately by infinite plane waves $e^{i(kz-\omega t)}$, although strictly speaking the waves do not extend to infinity either in width or in length. After scattering, the particles are detected at a great distance from the scatterer, and we would like to describe them again, asymptotically, by simple, radially outgoing harmonic waves, but with direction-dependent amplitudes (Figure 13.2).

Only comparison between theory and experiment can tell us whether our assumption that all the particles in the beam can be represented by very broad and long wave packets truly reflects the properties of a real beam, consisting, as it does, of particles that were emitted from a source, that have perhaps undergone some acceleration, and that may have been selected as to energy and momentum by analyzers and slits. Fortunately, the mathematical analysis of the next section will show the cross section to be in general independent of the peculiarities of the incident wave packet, provided only that the packet is large compared with the scattering region toward which it is directed.

We conclude this section with a concrete illustrative example. Consider the elastic Rutherford scattering of 10-MeV alpha particles from an aluminum foil of $t = 1$ μm thickness. We will see in Section 13.8 that quantum mechanics gives essentially the same results for this force law as classical mechanics, allowing us to employ the differential cross section (13.9) for scattering of alpha particles. Since the atomic mass of aluminum is about 27 and its density is 2.7 g cm^{-3}, the number of Al atoms per cm^3 is $n = 6 \times 10^{22}$ cm^{-3}. The number of atoms per cm^2 of area is thus $nt = 6 \times 10^{18}$ cm^{-2}. The reduced mass of the α-Al system is $4 \times 27/31 = 3.48$ u (instead of approximately 4 u for an α particle colliding with an infinitely massive nucleus). Hence, the energy of the relative motion is 3.48/4 times the incident energy, or 8.7 MeV. Given that the charges of the colliding nuclei are $2e$ (for

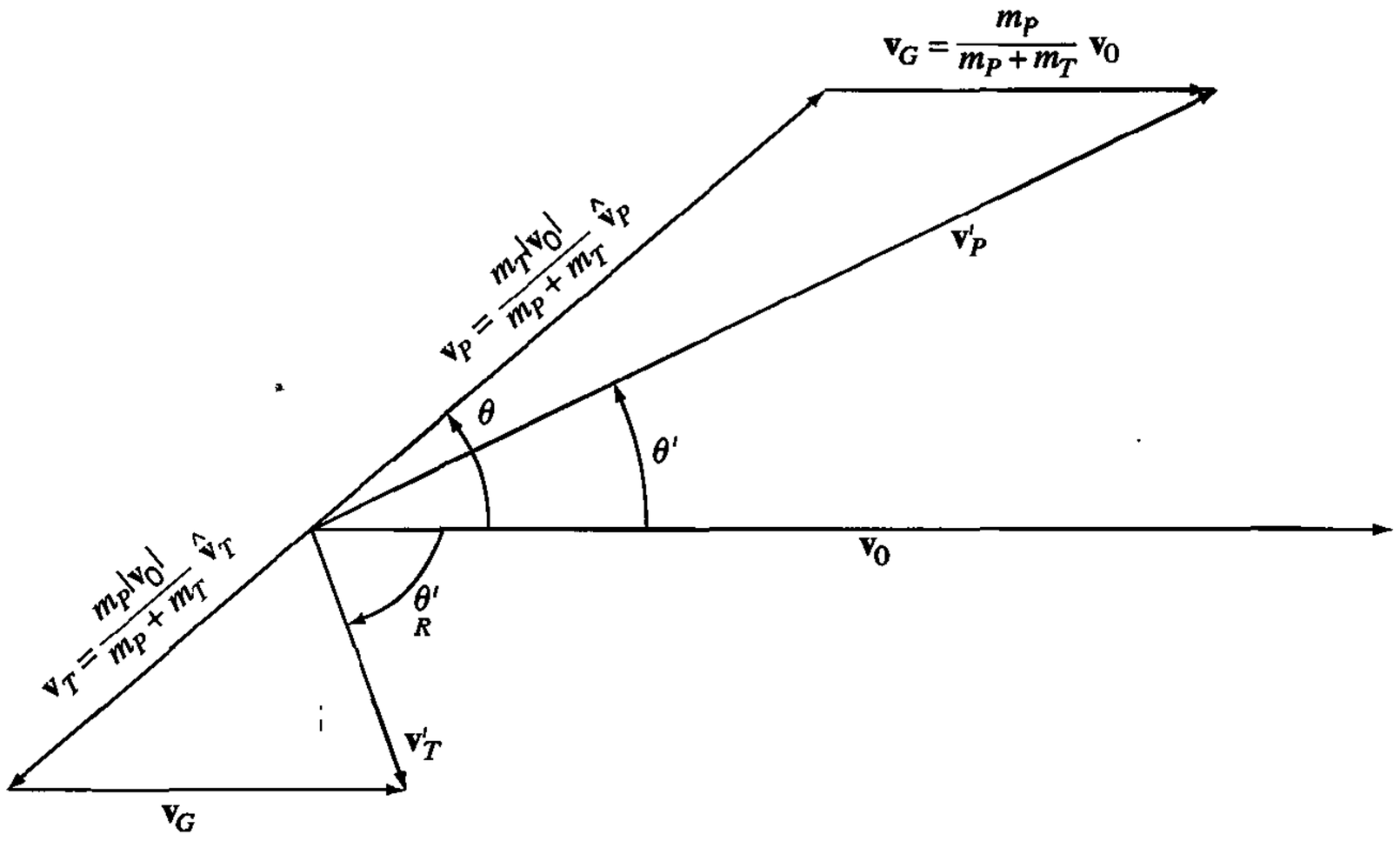


Figure 13.4. Kinematics of elastic scattering in the center-of-mass and laboratory coordinate frames. In the laboratory frame, the target particle is initially at rest, the projectile has velocity $\mathbf{v}_0$, and the center of mass has constant velocity $\mathbf{v}_G$. In the center-of-mass frame $\mathbf{v}_P$ and $\mathbf{v}_T$ are the velocities of the scattered projectile and the recoiling target particle, respectively. In the laboratory frame θ' and θ'_R are the scattering and recoil angles, respectively, and $\mathbf{v}'_P$ and $\mathbf{v}'_T$ are the final laboratory velocities. A mass ratio $m_p/m_T = 3/5$ has been assumed in this diagram.

He) and $13e$ (for Al), their classical distance of closest approach in a head-on collision is calculated from conservation of energy as 4.3×10^{-15} m or 4.3 fm.

For elastic collisions, the kinematic relation between the scattering angle θ in the center-of-mass or relative coordinate frame and the scattering angle θ' in the laboratory frame-of-reference is

$$\tan \theta' = \frac{\sin \theta}{\cos \theta + \dfrac{m_P}{m_T}} \tag{13.10}$$

where m_P is the mass of the incident projectile and m_T is the mass of the initially stationary target (Figure 13.4).

For example, corresponding to $\theta' = 30°$ in the laboratory frame, the scattering angle in the center-of-mass system (as well as in the relative coordinate frame) is $\theta = 34.2°$. The distance of closest approach corresponding to this scattering angle is 9.5 fm, which is close to the nuclear radius of aluminum, so that small departures from Rutherford scattering may occur. We then calculate the Rutherford differential cross section in the center-of-mass frame for this scattering angle as

$$\frac{d\sigma}{d\Omega} = 1.53 \times 10^{-28}\ \text{m}^2 = 1.53 \times 10^{-24}\ \text{cm}^2 = 1.53\ \text{barns}$$

This is the differential cross section per steradian. At this scattering angle one finds that

$$\frac{d\theta}{d\theta'} = 1.13$$

and, for the differential cross section in the laboratory frame,

$$\frac{d\sigma}{d\Omega'} = \frac{d\sigma}{d\Omega}\frac{\sin\theta}{\sin\theta'} = 1.95 \text{ barns} = 1.95 \times 10^{-24} \text{ cm}^2$$

If the solid angle subtended by the detector is $d\Omega' = 10^{-6}$ steradian, the calculated cross section for this outcome, is $d\sigma = 1.95 \times 10^{-30}$ cm^2, and the probability that an α particle is scattered into the detector with ideal detection efficiency is $nt\, d\sigma = 1.17 \times 10^{-11}$. In other words, it is expected that about one in 10^{11} of the α particles that are incident on the target will be scattered into the detector at this scattering angle. If the rate of incident α particles is $I_0 = 10^{12}$ sec^{-1}, the counting rate at the detector will be about 12 particles per second. (The luminosity in this experiment is $I_0 nt = 10^{12} \times 6 \times 10^{22} \times 10^4 = 6 \times 10^{38}$ sec^{-1} cm^{-2}.)

It is important to ask if the energy of a 10-MeV α particle is appreciably degraded as it passes through the 1 micron aluminum foil. Since the stopping power of aluminum for protons of the same velocity (2.5 MeV) is about 250 MeV/cm, an α particle (charge $2e$) loses about 0.1 MeV, or only 1 percent of the original energy, in its passage through the foil. Finally, one has to estimate the likelihood of multiple Rutherford scattering in the target. For the foil thickness assumed in our example, this effect is found to be quite small.

Exercise 13.2. Calculate the probability that in the example of Rutherford scattering just given an alpha particle is scattered through an angle greater than 10^{-2} radian in the laboratory frame. Use the result to estimate the importance of multiple scattering in this example.

2. ***The Scattering of a Wave Packet.*** Let us assume that the motion of the projectiles in the field of a single scatterer ($N_T = 1$) is described by a Hamiltonian

$$H = \frac{\mathbf{p}^2}{2m} + V = T + V = H_0 + V \tag{13.11}$$

where V is appreciably different from zero only within a sphere of radius a surrounding the coordinate origin. At $t = 0$ a particle is represented by a wave packet of the general form

$$\psi(\mathbf{r}, 0) = N \int \phi(\mathbf{k}) \exp[i\mathbf{k} \cdot (\mathbf{r} - \mathbf{r}_0)]\, d^3k \tag{13.12}$$

where $\phi(\mathbf{k})$ is a smooth function of narrow width, $\Delta\mathbf{k}$, centered around a mean momentum $\hbar\mathbf{k}_0$. The normalization constant N is arbitrary. The initial state $\psi(\mathbf{r}, 0)$, as expressed by (13.12), is thus an extended wave packet located in the vicinity of $\mathbf{r}_0$, but with fairly sharply defined momentum. We assume that $\mathbf{k}_0$ is parallel to $\mathbf{r}_0$, but in an opposite direction, so that the wave packet, if unhindered in its motion, would move freely toward the coordinate origin; and we assume $\mathbf{r}_0$ to be so large that ψ at $t = 0$ lies in its entirety well to the left of the scatterer (Figure 13.2).

The dynamical problem to be solved is as follows: What is the shape of the wave packet (13.12) at a much later time, when the packet has encountered the scatterer and been eventually dispersed by it? In principle, the answer can be given easily, if we succeed in expanding $\psi(\mathbf{r}, 0)$ in terms of eigenfunctions, $\psi_n(\mathbf{r})$, of H. Indeed, if we can establish an expansion

$$\psi(\mathbf{r}, 0) = \sum_{n1} c_n \psi_n(\mathbf{r})$$

then the wave packet at time t is

$$\psi(\mathbf{r}, t) = \sum_{n1} c_n \psi_n(\mathbf{r}) \exp\left(-\frac{i}{\hbar} E_n t\right)$$

How can this program be applied to a wave packet of the form (13.12)? True, it is an expansion in terms of a complete orthonormal set, the plane waves $e^{i\mathbf{k}\cdot\mathbf{r}}$, but these are eigenfunctions of the operator H_0, and not of H. However, it will be proved that it is possible to replace the plane wave functions in the expansion by *particular* eigenfunctions of H, which we will designate as $\psi_{\mathbf{k}}^{(+)}(\mathbf{r})$, if $\psi(\mathbf{r}, 0)$ is a wave packet of the kind described. Asymptotically, the eigenfunctions $\psi_{\mathbf{k}}^{(+)}(\mathbf{r})$ bear a considerable resemblance to plane waves, since they are of the form

$$\boxed{\psi_{\mathbf{k}}^{(+)}(\mathbf{r}) \cong N\left(e^{i\mathbf{k}\cdot\mathbf{r}} + f_{\mathbf{k}}(\hat{\mathbf{r}}) \frac{e^{ikr}}{r}\right) \qquad (r \text{ large})} \tag{13.13}$$

This differs from a plane wave at large $\mathbf{r}$ only by an *outgoing spherical wave*. As we will see in Section 13.3, such solutions of the Schrödinger equation

$$H\psi_{\mathbf{k}}^{(+)}(\mathbf{r}) = E\psi_{\mathbf{k}}^{(+)}(\mathbf{r}) \tag{13.14}$$

do indeed exist. In order not to interrupt the argument, we assume their existence in this section. It will also be shown in Section 13.3 that the special wave packet $\psi(\mathbf{r}, 0)$ has the same expansion coefficients, whether it is expanded in terms of plane waves, as in (13.12), or in terms of the set $\psi_{\mathbf{k}}^{(+)}(\mathbf{r})$:

$$\psi(\mathbf{r}, 0) = \int \phi(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r}_0} \psi_{\mathbf{k}}^{(+)}(\mathbf{r})\, d^3k \tag{13.15}$$

In other words, the outgoing wave in (13.13) asymptotically makes no contribution in the initial wave packet. The replacement of (13.12) by (13.15) is the critical step in this analysis.[2]

Once the initial wave function $\psi(\mathbf{r}, 0)$ has been expanded in terms of eigenfunctions of H, we can write at any time t,

$$\psi(\mathbf{r}, t) = \int \phi(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r}_0 - i\omega t} \psi_{\mathbf{k}}^{(+)}(\mathbf{r})\, d^3k \tag{13.16}$$

where

$$\hbar\omega = \frac{\hbar^2 k^2}{2m}$$

is the energy eigenvalue for the eigenfunction $\psi_{\mathbf{k}}^{(+)}(\mathbf{r})$.

Assuming that the scattering detectors are at a macroscopic distance of the order of r_0 from the scatterer, after a time

$$T = \frac{2mr_0}{\hbar k_0} \tag{13.17}$$

[2]The discussion of the time development of wave packets in scattering follows lecture notes by F. E. Low.

the broad pulse will be traveling through the position of the detectors (see Figure 13.2). When we examine the pulse at the position of the detector, $\psi_{\mathbf{k}}^{(+)}(\mathbf{r})$ can again be represented by its asymptotic expansion, but since the phases have changed with time the outgoing wave may no longer be neglected. However, as in (2.29), we can make an expansion about k_0:

$$k^2 = k_0^2 + 2(\mathbf{k} - \mathbf{k}_0) \cdot \mathbf{k}_0 + (\mathbf{k} - \mathbf{k}_0)^2 \tag{13.18}$$

In order to be able to neglect the last term, when k^2 is substituted at $t = T$ into the expression

$$\omega T = \frac{\hbar k^2}{2m} T$$

in Eq. (13.16), we require that T, though large, should still satisfy the inequality

$$\frac{\hbar}{2m} (\mathbf{k} - \mathbf{k}_0)^2 \, T \ll 1$$

or, using (13.17),

$$\frac{(\Delta k)^2 r_0}{k_0} \ll 1$$

This condition, familiar from Section 2.4, implies that the wave packet spreads negligibly when it is displaced by the macroscopic distance $\mathbf{r}_0$.

Exercise 13.3. Show, by numerical example, that in scattering of particles from atoms and nuclei the condition of no spreading can be easily attained by a minimum uncertainty wave packet ($\Delta k \Delta r \approx 1$), which is large compared with the scattering region but small compared with r_0.

Approximately and asymptotically, we can thus write (13.16) as

$$\psi(\mathbf{r}, t) = \int \phi(\mathbf{k}) \exp[-i\mathbf{k} \cdot (\mathbf{r}_0 + \mathbf{v}_0 t) + i\omega_0 t] \times N\left(e^{i\mathbf{k}\cdot\mathbf{r}} + f_{\mathbf{k}}(\hat{\mathbf{r}}) \frac{e^{ikr}}{r}\right) d^3k \tag{13.19}$$

where

$$\mathbf{v}_0 = \frac{\hbar \mathbf{k}_0}{m}, \qquad \hbar\omega_0 = \frac{1}{2} m v_0^2$$

Comparing (13.19) with (13.12), and assuming that $f_{\mathbf{k}}(\hat{\mathbf{r}})$, unlike $\phi(\mathbf{k})$, is a slowly varying function of $\mathbf{k}$, we can write

$$\psi(\mathbf{r}, t) = \psi(\mathbf{r} - \mathbf{v}_0 t, 0) e^{i\omega_0 t} + N \frac{f_{\mathbf{k}_0}(\hat{\mathbf{r}}) e^{i\omega_0 t}}{r} \int \phi(\mathbf{k}) \exp\{i[kr - \mathbf{k} \cdot (\mathbf{r}_0 + \mathbf{v}_0 t)]\} \, d^3k$$

with carets denoting unit vectors. Since for a well-collimated beam, $\phi(\mathbf{k})$ is appreciably different from zero only for $\mathbf{k} \approx \mathbf{k}_0$, we can write in the exponent effectively,

$$kr \approx \mathbf{k} \cdot \hat{\mathbf{k}}_0 r \tag{13.20}$$

and consequently, again by comparison with (13.12), asymptotically,

$$\psi(\mathbf{r}, t) = \psi(\mathbf{r} - \mathbf{v}_0 t, 0)e^{i\omega_0 t} + \frac{f_{\mathbf{k}_0}(\hat{\mathbf{r}})}{r}\psi(r\hat{\mathbf{k}}_0 - \mathbf{v}_0 t, 0)e^{i\omega_0 t} \tag{13.21}$$

Except for the phase factor $e^{i\omega_0 t}$, the first term on the right-hand side represents the *initial* wave packet displaced without change of shape, as if no scattering had occurred, i.e., as if V were absent from the Hamiltonian.

The factor $\psi(\mathbf{r}\mathbf{k}_0 - \mathbf{v}_0 t, 0)$ in the second term of (13.21) can be expressed as $\psi([r - v_0 t]\hat{\mathbf{k}}_0, 0)$. It differs from zero only when $r - v_0 t \approx -r_0$, within a range Δr that is the same as the spatial width of the initial wave packet, and represents a radially expanding image of the incident wave packet. A detector placed at distance r is reached by this wave packet at time $t = (r_0 + r)/v_0$, as expected for uniform particle motion. Thus, the second term in (13.21) is a *scattered spherical wave packet*, reduced in amplitude by the factor $1/r$ and modulated by the angular amplitude $f_{\mathbf{k}_0}(\hat{\mathbf{r}})$. Sensibly, the latter is called the *scattering amplitude*.

The assumption, made in deriving (13.21), of a slow variation of the scattering amplitude with $\mathbf{k}$, excludes from this treatment those scattering *resonances* which are characterized by an exceptionally rapid change of the scattering amplitude with the energy of the particle. When a resonance is so sharp that the scattering amplitude changes appreciably even over the narrowest tolerable momentum range Δk, the scattered wave packet may have a shape that is considerably distorted from the incident wave packet, because near such a resonance different momentum components scatter very differently. There is then strong dispersion. (See Section 13.6 for some discussion of resonances.)

The probability of observing the particle at the detector in the time interval between t and $t + dt$ is

$$v_0 \frac{|f_{\mathbf{k}_0}(\hat{\mathbf{r}})|^2}{r^2} |\psi([r - v_0 t]\hat{\mathbf{k}}_0, 0)|^2 r^2 \, d\Omega \, dt$$

Hence, the total probability for detecting it at any time is

$$v_0 |f_{\mathbf{k}_0}(\hat{\mathbf{r}})|^2 \, d\Omega \int_{-\infty}^{+\infty} |\psi([r - v_0 t]\hat{\mathbf{k}}_0, 0)|^2 \, dt = |f_{\mathbf{k}_0}(\hat{\mathbf{r}})|^2 \, d\Omega \int_{-\infty}^{+\infty} |\psi(\xi\hat{\mathbf{k}}_0, 0)|^2 \, d\xi \tag{13.22}$$

The limits of integration in (13.22) may be taken to be $\pm\infty$ with impunity, since $\psi(\mathbf{r}, t)$ describes a *wave packet* of finite length. The probability that the incident particle will pass through a unit area located perpendicular to the beam in front of the scatterer is

$$\int_{-\infty}^{+\infty} |\psi(\xi\hat{\mathbf{k}}_0, 0)|^2 \, d\xi$$

There is no harm in extending this integration from $-\infty$ to $+\infty$ either, since at $t = 0$ the wave packet is entirely in front of the scatterer. If the ensemble of projectiles contains N_P particles, all represented by the same general type of wave packet ψ_i ($i = 1, \ldots N_P$), the number of particles scattered into the solid angle $d\Omega$ is

$$I(\hat{\mathbf{r}}) \, d\Omega = |f_{\mathbf{k}_0}(\hat{\mathbf{r}})|^2 \, d\Omega \sum_{i=1}^{N_P} \int_{-\infty}^{+\infty} |\psi_i(\xi\hat{\mathbf{k}}_0, 0)|^2 \, d\xi$$

whereas

$$I_0 = \sum_{i=1}^{N_P} \int_{-\infty}^{+\infty} |\psi_i(\xi \hat{\mathbf{k}}_0, 0)|^2 \, d\xi$$

gives the number of particles incident per unit area. Hence, by (13.5) and setting $N_T = ntA = 1$ for a single scattering center, we arrive at the fundamental result:

$$\boxed{\frac{d\sigma}{d\Omega} = |f_{\mathbf{k}_0}(\hat{\mathbf{r}})|^2} \tag{13.23}$$

establishing the fundamental connection between the scattering amplitude and the differential cross section. This result is independent of the detailed behavior of the projectile wave packets and the normalization constant N.

3. ***Green's Functions in Scattering Theory.*** To complete the discussion, it is necessary to show that

(a) eigenfunctions of the asymptotic form (13.13) exist, and

(b) the expansion (13.15) is correct.

These two problems are not unconnected, and (b) will find its answer after we have constructed the solutions $\psi_{\mathbf{k}}^{(+)}(\mathbf{r})$. The method of the *Green's function* by which this may be accomplished is far more general than the immediate problem would suggest.

The Schrödinger equation to be solved is

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\psi = E\psi$$

or

$$(\nabla^2 + k^2)\psi = U\psi \tag{13.24}$$

where $k^2 = 2mE/\hbar^2$ and $U = 2mV/\hbar^2$. It is useful to replace the differential equation (13.24) by an integral equation. The transformation to an integral equation is performed most efficiently by regarding $U\psi$ on the right-hand side of (13.24) temporarily as a given inhomogeneity, even though it contains the unknown function ψ. Formally, then, a particular "solution" of (13.24) is conveniently constructed in terms of the Green's function $G(\mathbf{r}, \mathbf{r}')$, which is a solution of the equation

$$(\nabla^2 + k^2)G(\mathbf{r}, \mathbf{r}') = -4\pi\delta(\mathbf{r} - \mathbf{r}') \tag{13.25}$$

Indeed, the expression

$$-\frac{1}{4\pi}\int G(\mathbf{r}, \mathbf{r}')U(\mathbf{r}')\psi(\mathbf{r}') \, d^3r' \tag{13.26}$$

solves (13.24) by virtue of the properties of the delta function. To this particular solution we may add an arbitrary solution of the homogeneous equation

$$(\nabla^2 + k^2)\psi = 0 \tag{13.27}$$

which is the Schrödinger equation for a free particle (no scattering). Leaving a normalization factor N undetermined, we thus establish the *integral equation*

$$\boxed{\psi_{\mathbf{k}}(\mathbf{r}) = Ne^{i\mathbf{k}\cdot\mathbf{r}} - \frac{1}{4\pi}\int G(\mathbf{r}, \mathbf{r}')U(\mathbf{r}')\psi_{\mathbf{k}}(\mathbf{r}')\, d^3r'} \tag{13.28}$$

for a particular set of solutions of the Schrödinger equation (13.24). The vector $\mathbf{k}$ has a definite magnitude, fixed by the energy eigenvalue, but its direction is undetermined. Thus, the solution exhibits an infinite degree of degeneracy, which corresponds physically to the possibility of choosing an arbitrary direction of incidence.

Even if a particular vector $\mathbf{k}$ is selected, (13.28) is by no means completely defined yet: The Green's function could be any solution of (13.25), and there are infinitely many different ones. The choice of a particular $G(\mathbf{r}, \mathbf{r}')$ imposes definite boundary conditions on the eigenfunctions $\psi_{\mathbf{k}}(\mathbf{r})$. Two particularly useful Green's functions are

$$\boxed{G_{\pm}(\mathbf{r}, \mathbf{r}') = \frac{\exp(\pm ik|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|}} \tag{13.29}$$

Exercise 13.4. Verify that the Green's functions (13.29) are solutions of the inhomogeneous equation (13.25).

A host of Green's functions of the form

$$G(\mathbf{r}, \mathbf{r}') = G(\mathbf{r} - \mathbf{r}') \tag{13.30}$$

may be obtained by applying a Fourier transformation to the equation

$$(\nabla^2 + k^2)G(\mathbf{r}) = -4\pi\delta(\mathbf{r}) \tag{13.31}$$

which is a simplified version of (13.25). If we introduce the Fourier integral

$$G(\mathbf{r}) = \int g(\mathbf{k}')e^{i\mathbf{k}'\cdot\mathbf{r}}\, d^3k'$$

and the Fourier representation of the delta function, we obtain by substitution into (13.31)

$$g(\mathbf{k}') = \frac{1}{2\pi^2}\frac{1}{k'^2 - k^2}$$

hence, the Fourier representation

$$G(\mathbf{r}) = \int \frac{1}{2\pi^2}\frac{1}{k'^2 - k^2} e^{i\mathbf{k}\cdot\mathbf{r}}\, d^3k' \tag{13.32}$$

Integrating over the angles, we obtain, after a little algebra, the convenient form

$$G(r) = \frac{1}{i\pi r}\int_{-\infty}^{+\infty} \frac{e^{ik'r}}{k'^2 - k^2} k'\, dk' \tag{13.33}$$

Since the integrand has simple poles on the real axis in the complex k'-plane, at $k' = \pm k$, the integral in (13.33) does not really exist. This suggests that our attempt to represent the solutions of (13.25) as Fourier integrals has failed. Nevertheless,

this approach has the potential to succeed, because the integral in (13.33) can be replaced by another one that does exist, thus

$$G_{+\eta}(r) = \frac{1}{i\pi r}\int_{-\infty}^{+\infty} \frac{e^{ik'r}}{k'^2 - (k^2 + i\eta)}\, k'\, dk' \tag{13.34}$$

where η is a small positive number. So defined, $G_{+\eta}(r)$ exists but is no longer a solution to (13.31). The trick is to evaluate the expression (13.34) for $\eta \neq 0$ and to let $\eta \to 0$, i.e., $G_{+\eta}(r) \to G_{+}(r)$, after the integration has been performed.

Alternatively, we might say that a unique solution of (13.31) does not exist, for if it did G would be the inverse of the operator $-(1/4\pi)(\nabla^2 + k^2)$. But this operator has no inverse, because the homogeneous equation $(\nabla^2 + k^2)\psi = 0$ does have non-trivial solutions. However, the inverse of $-(1/4\pi)(\nabla^2 + k^2 + i\eta)$ exists and is the Green's function $G_{+\eta}(r)$ in (13.34). (See also Section 18.2.)

The integral

$$\int_{-\infty}^{+\infty} \frac{e^{ik'r}}{k'^2 - (k^2 + i\eta)}\, k'\, dk'$$

is most easily computed by using the complex plane as an auxiliary device (see Figure 13.5). The poles of the integrand are at

$$k' = \pm\sqrt{k^2 + i\eta} \approx \pm\left(k + \frac{i\eta}{2k}\right)$$

for small η. The path of integration leads along the real axis from $-\infty$ to $+\infty$.

Since r is necessarily positive, a closed contour may be used if we complete the path by a semicircle of very large radius through the upper half k'-plane. It encloses the pole in the right half plane. The result of the integration is not altered by introducing detours avoiding the two points $k' = +k$ and $k' = -k$, as indicated

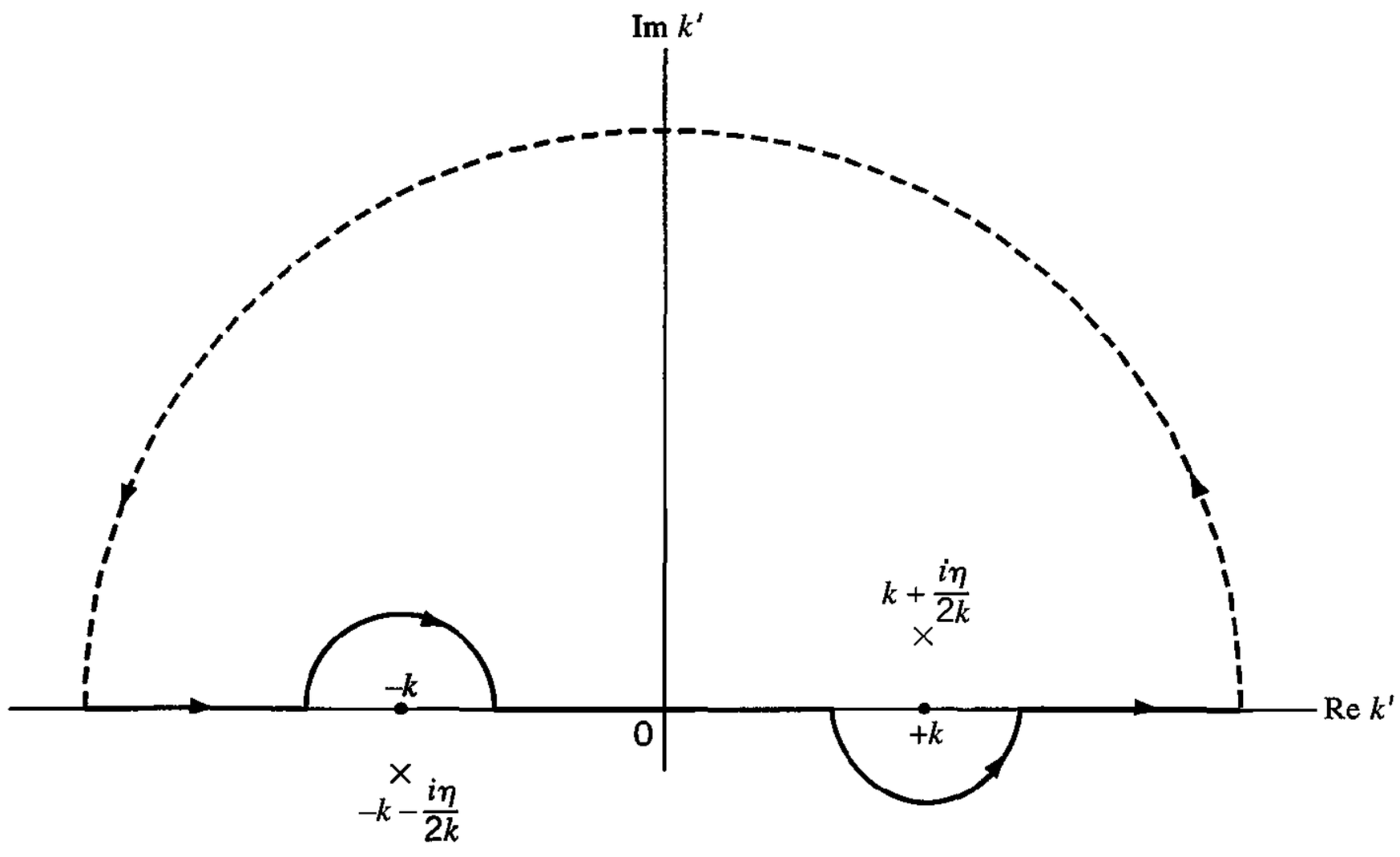


Figure 13.5. Path of integration in the complex k' plane.

in the figure. If this is done, the limit $\eta \to 0$ can be taken prior to the integration, and we may write

$$G_+(r) = \lim_{\eta \to 0} G_{+\eta}(r) = \frac{1}{i\pi r} \oint \frac{e^{ik'r}}{k'^2 - k^2} k' \, dk' \tag{13.35}$$

which by the use of the residue at $k' = k$ becomes

$$\boxed{G_+(r) = \frac{e^{ikr}}{r}} \tag{13.36}$$

If we replace η in (13.34) by $-\eta(\eta > 0)$, we obtain a second Green's function

$$G_-(r) = \lim_{\eta \to 0} G_{-\eta}(r) = \frac{e^{-ikr}}{r} \tag{13.37}$$

in agreement with (13.29). Still other Green's functions can be formed by treating the singularities that appear in (13.33) differently.

Exercise 13.5. Show that the Green's function

$$G_1(r) = \frac{1}{2} G_+(r) + \frac{1}{2} G_-(r) = \frac{\cos kr}{r} \tag{13.38}$$

is obtained if the integral in (13.33) is replaced by its (Cauchy) principal value.

The Green's functions (13.36), (13.37), and (13.38) may be identified as *outgoing*, *incoming*, and *standing* waves. To appreciate this terminology, we need only multiply them by the time factor $\exp(-iEt/\hbar)$. Furthermore, the description of scattering in terms of wave packets suggests the designation *retarded* for G_+ and *advanced* for G_-.

When the special forms (13.36) or (13.37) are substituted in (13.28), two distinct eigensolutions result, denoted by $\psi^{(+)}$ and $\psi^{(-)}$. They satisfy the integral equation,

$$\boxed{\psi_{\mathbf{k}}^{(\pm)}(\mathbf{r}) = Ne^{i\mathbf{k}\cdot\mathbf{r}} - \frac{1}{4\pi} \int \frac{\exp(\pm ik|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|} U(\mathbf{r}')\psi_{\mathbf{k}}^{(\pm)}(\mathbf{r}') \, d^3r'} \tag{13.39}$$

which is the Schrödinger equation rewritten in a form that is particularly convenient for use in scattering theory. We must show that in the asymptotic limit, as $r \to \infty$, the right-hand side of equation (13.39) assumes the simple form (13.13).

For large r the integrand can be closely approximated in view of the fact that $U \neq 0$ only for values of $r' < a$. In the exponent, we expand in powers of $\mathbf{r}'$:

$$k|\mathbf{r} - \mathbf{r}'| = k\sqrt{r^2 - 2\mathbf{r}\cdot\mathbf{r}' + r'^2} = kr - k\hat{\mathbf{r}}\cdot\mathbf{r}' + \frac{k(\hat{\mathbf{r}} \times \mathbf{r}')^2}{2r} + \cdots$$

If r is chosen so large that

$$ka^2/r \ll 1$$

then the quadratic term in the exponent can be neglected. If, furthermore, $\mathbf{r}'$ in the denominator of the integrand in (13.39) is neglected, we obtain for large r:

$$\psi_{\mathbf{k}}^{(\pm)}(\mathbf{r}) \cong Ne^{i\mathbf{k}\cdot\mathbf{r}} - \frac{e^{\pm ikr}}{4\pi r} \int e^{\mp i\mathbf{k}'\cdot\mathbf{r}'} U(\mathbf{r}')\psi_{\mathbf{k}}^{(\pm)}(\mathbf{r}') \, d^3r' \tag{13.40}$$

where we have set

$$\mathbf{k}' = k\hat{\mathbf{r}} \tag{13.41}$$

The asymptotic expression (13.40) can be written as

$$\psi_{\mathbf{k}}^{(\pm)}(\mathbf{r}) \cong N\left(e^{i\mathbf{k}\cdot\mathbf{r}} + \frac{e^{\pm ikr}}{r} f_{\mathbf{k}}^{(\pm)}(\mathbf{r})\right) \qquad (r \text{ large}) \tag{13.42}$$

where

$$\boxed{\begin{aligned} f_{\mathbf{k}}^{(\pm)}(\hat{\mathbf{r}}) &= -\frac{1}{4\pi N}\int e^{\mp i\mathbf{k}'\cdot\mathbf{r}'} U(\mathbf{r}')\psi_{\mathbf{k}}^{(\pm)}(\mathbf{r}')\, d^3r' \\ &= -\frac{m}{2\pi\hbar^2 N}\int e^{\mp i\mathbf{k}'\cdot\mathbf{r}'} V(\mathbf{r}')\psi_{\mathbf{k}}^{(\pm)}(\mathbf{r}')\, d^3r' \end{aligned}} \tag{13.43}$$

Equation (13.42) shows why $\psi^{(+)}$ and $\psi^{(-)}$, when supplemented by $\exp(-iEt/\hbar)$, are called the spherically *outgoing* and *incoming* solutions of the Schrödinger equation: They satisfy the appropriate boundary conditions at infinity. The outgoing solution is indeed asymptotically of the form (13.13), thus verifying assertion (a) at the beginning of this section. It is customary to omit the superscript symbol (+) qualifying the outgoing scattering amplitude and to write $f_{\mathbf{k}}(\hat{\mathbf{r}})$ for $f_{\mathbf{k}}^{(+)}(\hat{\mathbf{r}})$.

To prove assertion (b) we must employ the exact form (13.39) and demonstrate that for the initial wave packet

$$\int \phi(\mathbf{k}) \exp(-i\mathbf{k}\cdot\mathbf{r}_0)\, d^3k \int \frac{\exp(ik|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|} U(\mathbf{r}')\psi_{\mathbf{k}}^{(+)}(\mathbf{r}')\, d^3r' = 0$$

Since $U(\mathbf{r}') = 0$ for $r' > a$, it is sufficient to show that for $r' < a$

$$\int \phi(\mathbf{k}) \exp(-i\mathbf{k}\cdot\mathbf{r}_0 + ik|\mathbf{r}-\mathbf{r}'|)\psi_{\mathbf{k}}^{(+)}(\mathbf{r}')\, d^3k = 0 \tag{13.44}$$

It may usually be assumed that in this integral the variation of $\psi_{\mathbf{k}}^{(+)}(\mathbf{r}')$ with $\mathbf{k}$ can be neglected. It can be seen from the integral equation for $\psi_{\mathbf{k}}^{(+)}$ that this is true if the width of the wave packet $\Delta r \gg a$ and if we are not at an inordinately narrow resonance for which the scattering amplitude varies extremely rapidly with k. Most physical situations meet these conditions.

Again we note that the wave packet $\phi(\mathbf{k})$ is appreciably different from zero only for vectors $\mathbf{k}$ near the direction of $\mathbf{k}_0$, so that we may approximate as in (13.20),

$$k|\mathbf{r}-\mathbf{r}'| \approx \hat{\mathbf{k}}_0\cdot\mathbf{k}|\mathbf{r}-\mathbf{r}'|$$

Hence, the left-hand side of (13.44) is nearly equal to

$$\psi_{\mathbf{k}_0}^{(+)}(\mathbf{r}')\int \phi(\mathbf{k}) \exp[i\mathbf{k}\cdot(\hat{\mathbf{k}}_0|\mathbf{r}-\mathbf{r}'| - \mathbf{r}_0)]\, d^3k = N^{-1}\psi(\hat{\mathbf{k}}_0|\mathbf{r}-\mathbf{r}'|,0)\, \psi_{\mathbf{k}_0}^{(+)}(\mathbf{r}')$$

The right-hand side of this equation vanishes, because the vector $\hat{\mathbf{k}}_0|\mathbf{r}-\mathbf{r}'|$ points to a position behind the scatterer where the wave packet was assumed to vanish at $t = 0$. Hence, assertion (b) is proved, completing the discussion begun in Section 13.2.

Although the value of N does not affect the results, we may choose $N = (2\pi)^{-3/2}$, corresponding to $\mathbf{k}$-normalization. In Chapter 20 it will be shown

that, with this normalization, the energy eigenfunctions $\psi_{\mathbf{k}}^{(+)}(\mathbf{r})$ are orthonormal, since

$$\int \psi_{\mathbf{k}}^{(+)}(\mathbf{r})^* \psi_{\mathbf{k}'}^{(+)}(\mathbf{r})\, d^3r = \delta(\mathbf{k} - \mathbf{k}') \tag{13.45}$$

When supplemented by the bound-state energy eigenfunctions, they make up a complete set of orthonormal functions.

Exercise 13.6. If the potential V is real-valued, prove that $\psi_{\mathbf{k}}^{(+)}(\mathbf{r})$ and $\psi_{-\mathbf{k}}^{(-)}(\mathbf{r})$ are mutually time-reversed scattering-state solutions of the Schrödinger equation. Interpret this result.

Exercise 13.7. If the scattering potential has the translation invariance property $V(\mathbf{r} + \mathbf{R}) = V(\mathbf{r})$, where $\mathbf{R}$ is a constant vector, (a) prove that the scattering solutions $\psi_{\mathbf{k}}^{(\pm)}$ of the integral form of the Schrödinger equation are Bloch wave functions, since they satisfy the relation

$$\psi_{\mathbf{k}}^{(\pm)}(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{k}\cdot\mathbf{R}} \psi_{\mathbf{k}}^{(\pm)}(\mathbf{r})$$

and (b) show that the scattering amplitude vanishes unless $\mathrm{q} = \mathbf{k} - \mathbf{k}'$ is a reciprocal lattice vector (Exercise 4.14) which satisfies the condition

$$\mathrm{q} \cdot \mathbf{R} \equiv (\mathbf{k} - \mathbf{k}') \cdot \mathbf{R} = 2\pi n \tag{13.46}$$

where n is an integer. This relation is the Laue condition familiar in condensed-matter physics.[3]

4. ***The Born Approximation.*** Before we proceed, let us summarize the results of the preceding sections. If particles with an average momentum $\hbar\mathbf{k}$ are incident upon a scatterer represented by the potential $V(r)$, the differential cross section is given by (13.23),

$$\frac{d\sigma}{d\Omega} = |f_{\mathbf{k}}(\hat{\mathbf{r}})|^2$$

where the scattering amplitude $f_{\mathbf{k}}(\hat{\mathbf{r}})$ is defined as the coefficient of the outgoing wave in the asymptotic solution

$$\psi_{\mathbf{k}}^{(+)}(\mathbf{r}) \cong N\left(e^{i\mathbf{k}\cdot\mathbf{r}} + \frac{e^{ikr}}{r} f_{\mathbf{k}}(\mathbf{r})\right) \qquad (r \text{ large})$$

of the Schrödinger equation

$$(\nabla^2 + k^2)\psi = \frac{2m}{\hbar^2} V\psi$$

The scattering amplitude for elastic scattering in the direction $\hat{\mathbf{k}}'$ is given by the formula

$$f_{\mathbf{k}}(\hat{\mathbf{k}}') = -\frac{m}{2\pi\hbar^2 N} \int e^{-i\mathbf{k}'\cdot\mathbf{r}'} V(\mathbf{r}') \psi_{\mathbf{k}}^{(+)}(\mathbf{r}')\, d^3r \tag{13.47}$$

[3]See Christman (1988), Section 4.2.

but since $\psi^{(+)}$ appears in the integrand, this is not an explicit expression. However, it can be used to obtain an *estimate* of the scattering amplitude if we replace the exact eigenfunction $\psi^{(+)}$ in the integrand by the normalized plane wave term on the right-hand side of Eq. (13.40), neglecting the scattered wave. In this approximation we obtain

$$\boxed{f_{\mathbf{k}}(\hat{\mathbf{k}}) \approx -\frac{m}{2\pi\hbar^2}\int e^{-i\mathbf{k}'\cdot\mathbf{r}'}V(\mathbf{r}')e^{i\mathbf{k}\cdot\mathbf{r}'}\,d^3r'} \tag{13.48}$$

which is known as the scattering amplitude in the (first) *Born approximation*. Here the scattering amplitude appears proportional to the matrix element of the scattering potential between the plane waves, $e^{i\mathbf{k}\cdot\mathbf{r}}$ and $e^{i\mathbf{k}'\cdot\mathbf{r}}$, representing the free particle before and after the scattering. It should be evident that the Born approximation corresponds to a first iteration of (13.40), where the plane wave is substituted for $\psi^{(+)}$ under the integral. The iteration may be continued to obtain higher order Born approximations. (See Section 20.3.)

For a central-force potential, $V(r)$, the Born scattering amplitude (13.48) reduces to

$$f_{\text{Born}}(\theta) = -\frac{m.}{2\pi\hbar^2}\int V(r')e^{-i\mathbf{q}\cdot\mathbf{r}'}\,d^3r' \tag{13.49}$$

where

$$\mathbf{q} = \mathbf{k}' - \mathbf{k}$$

is known as the *momentum transfer* (in units of $\hbar$). As the Fourier transform of the potential, the Born amplitude (13.49) resembles the field amplitude in Fraunhofer diffraction. Figure 13.6 is a primitive Feynman diagram, depicting the first Born approximation scattering amplitude (13.48).

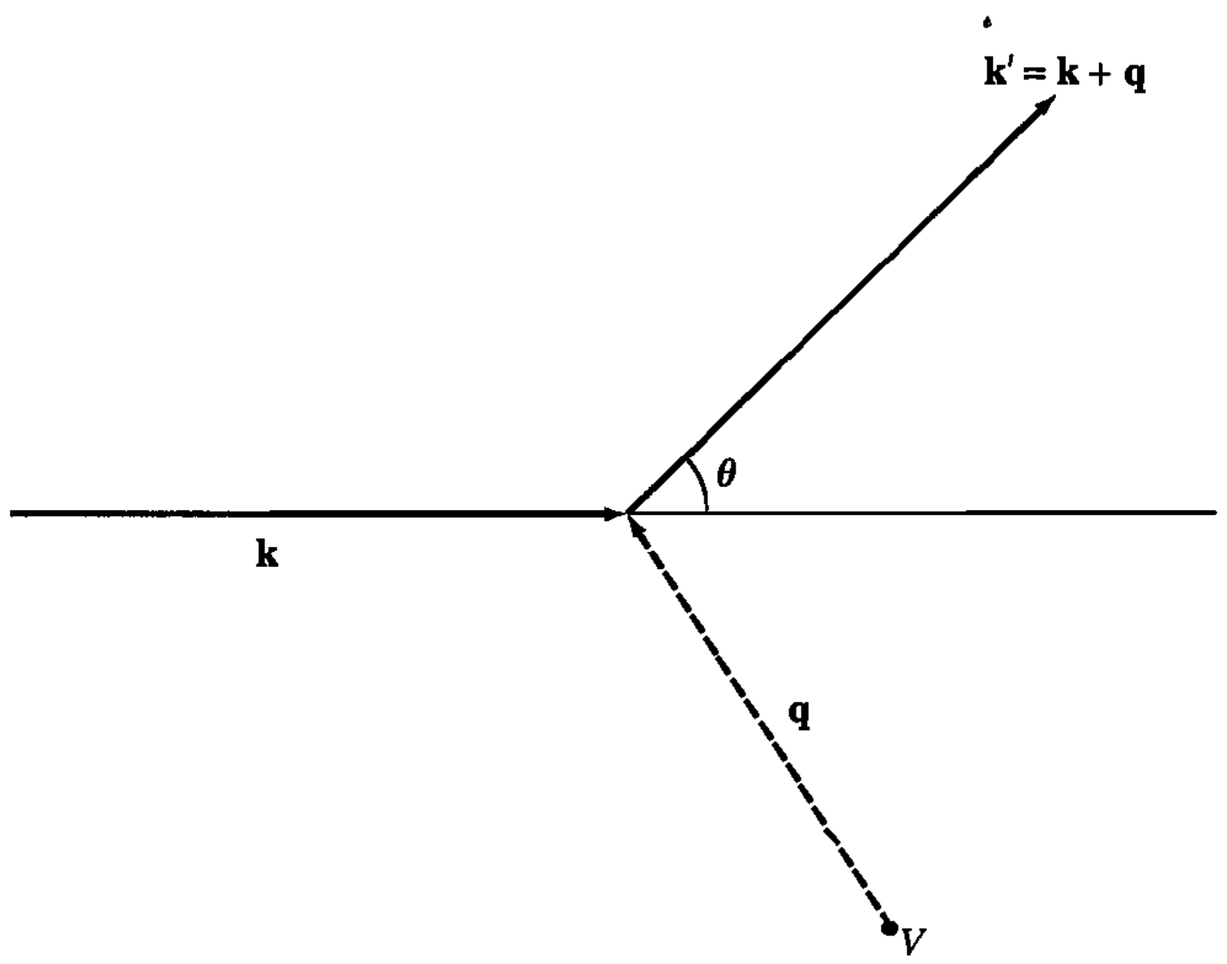


Figure 13.6. Diagram representing the first Born approximation matrix element, $\langle \mathbf{k}' | V | \mathbf{k} \rangle$, of the potential between eigenstates of incident momentum $\mathbf{k}$ and scattered momentum $\mathbf{k}' = \mathbf{k} + \mathbf{q}$ (in units of $\hbar$). The potential V causes the momentum transfer $\mathbf{q}$. The scattering angle is θ.

The integral over the solid angle in (13.49) is easily carried out and yields the result

$$f_{\text{Born}}(\theta) = -\frac{2m}{\hbar^2}\int_0^\infty V(r')\,\frac{\sin qr'}{qr'}\,r'^2\,dr' \tag{13.50}$$

Here we denote scattering angle between $\mathbf{k}$ and $\mathbf{k}'$ by θ, and note that $k' = k$ for elastic scattering, so that

$$q = 2k\sin\frac{\theta}{2} \tag{13.51}$$

as seen in Figure 13.6.

As an application, consider the screened Coulomb (or Yukawa) potential

$$V(r) = V_0\,\frac{e^{-\alpha r}}{\alpha r} \tag{13.52}$$

where the length $1/\alpha$ may be considered as the range of the potential. In the Born approximation, after a simple integration over angles, we find

$$\begin{aligned} f_{\text{Born}}(\theta) &= -\frac{2m}{\hbar^2}V_0\int_0^\infty \frac{e^{-\alpha r'}}{\alpha r'}\,\frac{\sin qr'}{qr'}\,r'^2\,dr' \\ &= -\frac{2m}{\hbar^2\alpha}V_0\,\frac{1}{q^2+\alpha^2} = -\frac{2mV_0}{\hbar^2\alpha}\,\frac{1}{4k^2\sin^2(\theta/2)+\alpha^2} \end{aligned} \tag{13.53}$$

The differential scattering cross section is obtained by taking the square of the amplitude (13.53).

The unscreened Coulomb potential between two charges q_1 and q_2 is a limiting case of the potential (13.52) for $\alpha \to 0$ and $V_0 \to 0$ with $V_0/\alpha = q_1q_2$. Hence, in the Born approximation,

$$\frac{d\sigma}{d\Omega} = \frac{m^2q_1^2q_2^2}{4p^4\sin^4(\theta/2)} = \frac{q_1^2q_2^2}{16E^2\sin^4(\theta/2)}$$

This result, even though obtained by an approximation method, is in exact agreement with both the classical Rutherford cross sections (13.9) and the exact quantum mechanical evaluation of the Coulomb scattering cross section—one of the many remarkable coincidences peculiar to the pure Coulomb potential. Note, however, that the exact Coulomb scattering amplitude differs from the Born amplitude by a phase factor (see Section 13.8).

Exercise 13.8. Calculate the total scattering cross section for the screened Coulomb potential (13.52) in the Born approximation and discuss the accuracy of this result.

Exercise 13.9. Apply the Born approximation to the scattering from a square well. Evaluate and plot the differential and total scattering cross sections.

Exercise 13.10. Obtain the differential scattering cross section in the Born approximation for the potential

$$V(r) = -V_0e^{-r/a} \qquad (V_0 > 0)$$

Exercise 13.11. If $V = C/r^n$, obtain the functional dependence of the Born scattering amplitude on the scattering angle. Discuss the reasonableness of the result qualitatively. What values of n give a meaningful answer?

A reliable estimate of the accuracy of the Born approximation is in general not easy to obtain, since the term that is neglected is itself an integral that depends on the potential, the wave function in the region where the potential does not vanish, and particularly on the momentum of the scattered particle. Qualitatively, it is easy to see that the Born approximation is a form of perturbation theory, in which the kinetic energy operator is the unperturbed Hamiltonian and the potential is the perturbation. The approximation is thus likely to be valid for weak potentials and high energies.

The Born approximation affords a rapid estimate of scattering cross sections and is valid for reasonably high energies in comparison with the interaction energy. Because of its simplicity, it has enjoyed great popularity in atomic and nuclear physics. Its usefulness does not, however, vitiate the need for an exact method of calculating scattering cross sections. To this task we must now attend.

5. ***Partial Waves and Phase Shifts.*** Let us assume that V is a central potential. It is to be expected that for a spherically symmetric potential the solution (13.39), representing an incident and a scattered wave, should exhibit cylindrical symmetry about the direction of incidence. Indeed, we can also see formally from (13.39) that $\psi_{\mathbf{k}}^{(+)}(\mathbf{r})$ depends on k and r and on the angle θ between $\mathbf{k}$ and $\mathbf{r}$ only, if V is a function of r alone. Hence, we may, without loss of generality, assume that $\mathbf{k}$ points in the positive z direction and that for a given value of k, $\psi_{\mathbf{k}}^{(+)}$ is a function of r and the scattering angle θ.

Exercise 13.12. Show that for a central potential $\psi_{\mathbf{k}}^{(+)}$ is an eigenfunction of the component of $\mathbf{L}$ in the direction of $\mathbf{k}$, with eigenvalue zero, and discuss the significance of this fact for the scattering of a wave packet.

We must thus look for solutions of the Schrödinger equation

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V(r)\psi = \frac{\hbar^2 k^2}{2m}\psi$$

which have the asymptotic form

$$\psi_{\mathbf{k}}^{(+)} \cong N\left(e^{ikr\cos\theta} + f_k(\theta)\frac{e^{ikr}}{r}\right) \tag{13.54}$$

It is desirable to establish the connection between these solutions and the separable solutions (12.3) of the central-force problem,

$$R_{\ell,k}(r)Y_\ell^m(\theta, \varphi) = \frac{u_{\ell,k}(r)}{r}Y_\ell^m(\theta, \varphi) \tag{13.55}$$

which are common eigenfunctions of H, $\mathbf{L}^2$, and L_z. The radial functions $R_{\ell,k}(r)$ and $u_{\ell,k}(r)$ satisfy the differential equations

$$\left[-\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right) + \frac{\ell(\ell+1)}{r^2} + \frac{2m}{\hbar^2}V(r) - k^2\right]R_{\ell,k}(r) = 0 \tag{13.56}$$

and

$$\left[-\frac{d^2}{dr^2} + \frac{\ell(\ell+1)}{r^2} + \frac{2m}{\hbar^2}V(r) - k^2\right]u_{\ell,k}(r) = 0 \tag{13.57}$$

respectively, as well as a boundary condition at the origin. In general, this boundary condition depends on the shape of V, but, as we saw in Section 12.4, in most practical cases it reduces to the requirement that the wave function $R_{\ell,k}(r)$ be finite at the origin, from which it follows that

$$u_{\ell,k}(0) = 0 \tag{13.58}$$

We will restrict ourselves to potentials that are in accord with this boundary condition. The value of $u'_{\ell,k}(0)$ implies a choice of the normalization of $u_{\ell,k}(r)$.

The radial equation for the external region $r > a$, where the scattering potential vanishes, is identical with Eq. (12.9) which was solved in the last chapter. The general solution of this equation is a linear combination of the regular and irregular solutions and has the form

$$R_{\ell,k}(r) = \frac{u_{\ell,k}(r)}{r} = A_\ell\, j_\ell(kr) + B_\ell n_l(kr) \qquad r > a \tag{13.59}$$

Using the asymptotic approximations (12.20) and (12.22), we get for large kr

$$R_{\ell,k}(r) = \frac{u_{\ell,k}(r)}{r} \cong A_\ell \frac{\sin(kr - \ell\pi/2)}{kr} - B_\ell \frac{\cos(kr - \ell\pi/2)}{kr} \qquad (kr \text{ large}) \tag{13.60}$$

In the complete absence of a scattering potential ($V = 0$ everywhere), the boundary condition at the origin would exclude the irregular solution, and we would have $B_\ell = 0$ for all values of ℓ. Hence, the magnitude of B_ℓ compared with A_ℓ is a measure of the intensity of scattering. The value B_ℓ/A_ℓ must be determined by solving the Schrödinger equation inside the scattering region ($r < a$), subject to the boundary condition, (13.58), and by joining the interior solution smoothly onto the exterior solution (13.59) at $r = a$. To do this, we must know V explicitly and solve (13.57) by numerical methods, if necessary.

A very useful expression for the cross section can be derived by introducing the ratios B_ℓ/A_ℓ as parameters. Since for a real-valued potential $u_{\ell,k}(r)$ may be assumed to be real, these parameters are real-valued numbers, and we may set

$$\frac{B_\ell}{A_\ell} = -\tan\delta_\ell \tag{13.61}$$

where δ_ℓ is a real angle that vanishes for all ℓ if $V = 0$ everywhere. The name *scattering phase shift* is thus appropriate for δ_ℓ, particularly if we note that (13.60) can now be written as

$$\frac{u_{\ell,k}(r)}{r} \cong \frac{\sin(kr - \ell\pi/2 + \delta_\ell)}{kr} \tag{13.62}$$

The phase shift δ_ℓ measures the amount by which the phase of the radial wave function for angular momentum ℓ differs from the no-scattering case ($\delta_\ell = 0$). Each phase shift is, of course, a function of the energy, or of k.

Exercise 13.13. Show that (13.62) implies the normalization

$$\int_0^\infty u_{\ell,k}(r)u_{\ell,k'}(r)\,dr = \frac{\pi}{2k^2}\,\delta(k - k') \tag{13.63}$$

for the radial eigenfunctions. (*Hint:* Use the radial Schrödinger equation and integration by parts.)

In order to express the differential scattering cross section, or the scattering amplitude, through the phase shifts we must expand $\psi_k^{(+)}(r, \theta)$ in terms of the separable solutions of the form (13.55), which (except for any bound states) are assumed to constitute a complete set of orthonormal eigenfunctions. Thus, we set

$$\psi_k^{(+)}(r, \theta) = \sum_{\ell=0}^{\infty} C_\ell(k)P_\ell(\cos\theta)\,\frac{u_{\ell,k}(r)}{r} \cong \sum_{\ell=0}^{\infty} C_\ell(k)P_\ell(\cos\theta)\,\frac{\sin(kr - \ell\pi/2 + \delta_\ell)}{r} \tag{13.64}$$

where use has been made of the fact that $\psi_k^{(+)}(r, \theta)$ depends on the angle θ between **k** and **r** and not on the directions of each of these vectors separately (see Exercise 13.12).

The expansion coefficients $C_\ell(k)$ can be determined by comparing the two asymptotic expressions of the wave function, (13.54) and (13.64). We make use of the asymptotic expansion of the plane wave, (12.37), with which (13.54) can be written as

$$\psi_k^{(+)} \cong N\left[-\frac{e^{-ikr}}{2ikr}\sum_{\ell=0}^{\infty}(2\ell + 1)(-1)^\ell P_\ell(\cos\theta) + \frac{e^{ikr}}{2ikr}\sum_{\ell=0}^{\infty}(2\ell + 1)P_\ell(\cos\theta) + f_k(\theta)\,\frac{e^{ikr}}{r}\right] \tag{13.65}$$

On the other hand, (13.64) takes the equivalent form

$$\psi_k^{(+)} \cong -\frac{e^{-ikr}}{2ikr}\sum_{\ell=0}^{\infty} C_\ell(k)i^\ell e^{-i\delta_\ell}P_\ell(\cos\theta) + \frac{e^{ikr}}{2ikr}\sum_{\ell=0}^{\infty} C_\ell(k)(-i)^\ell e^{i\delta_\ell}P_\ell(\cos\theta) \tag{13.66}$$

By comparing the incoming spherical waves in (13.65) and (13.66), we see that the expansion coefficients must be of the form

$$C_\ell = N(2\ell + 1)i^\ell e^{i\delta_\ell} \tag{13.67}$$

Substitution of these values into (13.64) gives the asymptotic expression for the wave function,

$$\psi_k^{(+)}(r, \theta) \cong N\sum_{\ell=0}^{\infty}(2\ell + 1)i^\ell e^{i\delta_\ell}\,\frac{\sin(kr - \ell\pi/2 + \delta_\ell)}{kr}\,P_\ell(\cos\theta) \tag{13.68}$$

This differs from a plane wave by the presence of the phase shifts and is called a *distorted plane wave*. Comparing now the coefficients of the outgoing spherical wave $\frac{e^{ikr}}{r}$ in (13.65) and (13.66), we obtain

$$f_k(\theta) = \sum_{\ell=0}^{\infty}(2\ell + 1)\,\frac{e^{2i\delta_\ell} - 1}{2ik}\,P_\ell(\cos\theta) \tag{13.69}$$

or

$$\boxed{f_k(\theta) = \frac{1}{k}\sum_{\ell=0}^{\infty}(2\ell + 1)e^{i\delta_\ell(k)}\sin\delta_\ell(k)\,P_\ell(\cos\theta)} \tag{13.70}$$

This important formula gives the scattering amplitude in terms of the phase shifts by making what is known as a *partial wave analysis* of the scattering amplitude. If we remember that each term (partial wave) in the sum (13.70) corresponds to a definite value of angular momentum ℓ, the formula may be seen in a more physical light. If the scattering potential is strongest at the origin and decreases in strength as r increases, then we may expect the low angular momentum components, which classically correspond to small impact parameters and therefore close collisions, to scatter more intensely than the high angular momentum components. More quantitatively, this semiclassical argument suggests that if the impact parameter

$$b = \frac{\ell\hbar}{p} = \frac{\ell}{k}$$

exceeds the range a of the potential, or when $\ell > ka$, no appreciable scattering occurs. Thus, if $ka \gg 1$, making the classical argument applicable, we expect the phase shifts δ_ℓ for $\ell > ka$ to be vanishingly small.

But this argument is of a more general nature. For suppose that $ka \ll 1$, as is the case for scattering at low energies. Then the incident waves are long, and at a given instant the phase of the wave changes very little across the scattering region. Hence, all spatial sense of direction is lost, and the scattering amplitude must become independent of the angle θ. By (13.70) this implies that all phase shifts vanish, except for $\delta_0(k)$ corresponding to $\ell = 0$. When this occurs we say that there is only S-wave scattering, and this is isotropic, since $P_0(\cos\theta) = 1$. (See Section 13.7 for a better estimate of δ_ℓ.)

We conclude that the partial wave sum (13.70) is a particularly useful representation of the scattering amplitude, if on physical grounds only a few angular momenta are expected to contribute significantly. In fact, if V is not known beforehand, we may attempt to determine the phase shifts as functions of k empirically by comparing scattering data at various energies with the formula

$$\frac{d\sigma}{d\Omega} = |f_k(\theta)|^2 = \frac{1}{k^2}\left|\sum_{\ell=0}^{\infty}(2\ell + 1)e^{i\delta_\ell(k)}\sin\delta_\ell(k)\,P_\ell(\cos\theta)\right|^2 \tag{13.71}$$

By integrating, we obtain the *total scattering cross section*:

$$\boxed{\sigma = \int\frac{d\sigma}{d\Omega}\,d\Omega = \frac{4\pi}{k^2}\sum_{\ell=0}^{\infty}(2\ell + 1)\sin^2\delta_\ell} \tag{13.72}$$

In the differential scattering cross section (13.71), or *angular distribution* as it is often called, contributions from different partial waves (angular momenta) interfere with each other, because the scattering amplitude is a sum of terms with different ℓ values. Such a sum is often said to be a "coherent" superposition of different angular momenta. No such interferences occur in the integrated total cross section (13.72), which is therefore said to constitute an "incoherent" sum of partial wave contributions, σ_ℓ.

Each angular momentum value contributes at most a partial cross section

$$(\sigma_\ell)_{\max} = \frac{4\pi}{k^2}(2\ell + 1) \tag{13.73}$$

to the total scattering cross section. This value is of the same order of magnitude as the maximum classical scattering cross section per unit $\hbar$ of angular momentum. Indeed, if we use the estimate $b = \ell/k$ for the impact parameter, applying (13.7) we obtain for the contribution to the total cross section from a range $\Delta\ell = 1$, or $\Delta b = 1/k$,

$$\sigma_\ell = 2\pi b \Delta b = 2\pi \frac{\ell}{k^2}$$

For large ℓ, this agrees with (13.73) except for a factor 4. The difference is due to the inevitable presence of diffraction effects for which the wave nature of matter is responsible.

Exercise 13.14. Show that the scattering amplitude (13.70) and the total cross section (13.72) are related by the identity

$$\boxed{\sigma = \frac{4\pi}{k} \operatorname{Im} f_{\mathbf{k}}(0)} \tag{13.74}$$

This formula, known as the *optical theorem*, holds for collisions in general (see Section 20.6).

Finally, we note that the quantity

$$S_\ell(k) = e^{2i\delta_l(k)} \tag{13.75}$$

which appears in the scattering amplitude (13.69), is the ratio of the coefficients multiplying the outgoing wave, $e^{i(kr-\ell\pi/2)}$, and the incoming wave, $e^{-i(kr-\ell\pi/2)}$, in the wave function (13.68). In Section 20.5 the quantities S_ℓ will be identified as eigenvalues of the S matrix.

6. ***Determination of the Phase Shifts and Scattering Resonances.*** The theoretical determination of phase shifts requires that we solve the Schrödinger equation for the given potential and obtain the asymptotic form of the solutions. If the interior wave function is calculated and joined smoothly onto the exterior solution (13.59) at $r = a$, the phase shifts can be expressed in terms of the logarithmic derivatives at the boundary $r = a$:

$$\beta_\ell = \left(\frac{a}{R_\ell}\frac{dR_\ell}{dr}\right)_{r=a}$$

Using (13.59) and (13.61), we easily find the desired relation between the phase shift and the logarithmic derivative at $r = a$:

$$\beta_\ell(k) = ka\frac{j_\ell'(ka)\cos\delta_\ell - n_\ell'(ka)\sin\delta_\ell}{j_\ell(ka)\cos\delta_\ell - n_\ell(ka)\sin\delta_\ell} \tag{13.76}$$

and conversely,

$$S_\ell = e^{2i\delta_\ell} = -\frac{j_\ell - in_\ell}{j_\ell + in_\ell} \times \frac{\beta_\ell - ka\dfrac{j'_\ell - in'_\ell}{j_\ell - in_\ell}}{\beta_\ell - ka\dfrac{j'_\ell + in'_\ell}{j_\ell + in_\ell}} \tag{13.77}$$

where all the spherical cylinder functions (j_ℓ and n_ℓ) and their (primed) derivatives are to be evaluated at the argument ka. The first factor

$$e^{2i\xi_\ell} = -\frac{j_\ell - in_\ell}{j_\ell + in_\ell} \tag{13.78}$$

in the expression (13.77) defines the (real) phase angles ξ_ℓ which have a simple and interesting interpretation. The quantities ξ_ℓ are seen to be the phase shifts if $\beta_\ell \to \infty$, which implies that $R_\ell(a) = 0$. The radial wave functions vanish at $r = a$ for all values of ℓ if the potential represents a hard spherical core of radius a, so that the wave function cannot penetrate into the interior at all. The angles ξ_ℓ are therefore referred to as the *hard sphere phase shifts*.

Exercise 13.15. Verify directly from (13.59) and (13.61) that (13.78) gives the phase shifts for a hard sphere, and plot their dependence on ka for the lowest few values of ℓ.

If we introduce the real parameters Δ_ℓ and s_ℓ,

$$ka\frac{j'_\ell + in'}{j_\ell + in_\ell} = \Delta_\ell + is_\ell \tag{13.79}$$

we obtain the simple relations

$$e^{2i(\delta_\ell - \xi_\ell)} = \frac{\beta_\ell - \Delta_\ell + is_\ell}{\beta_\ell - \Delta_\ell - is_\ell} \tag{13.80}$$

and

$$e^{i\delta_\ell} \sin \delta_\ell = e^{2i\xi_\ell}\left(\frac{s_\ell}{\beta_\ell - \Delta_\ell - is_\ell} + e^{-i\xi_\ell} \sin \xi_\ell\right) \tag{13.81}$$

which make explicit how the partial wave contributions to the scattering amplitude depend on β_ℓ.

From (13.79) and (12.23) we deduce that

$$s_\ell = ka\frac{j_\ell n'_\ell - j'_\ell n_\ell}{j_\ell^2 + n_\ell^2} = \frac{1}{ka(j_\ell^2 + n_\ell^2)} \tag{13.82}$$

which shows that s_ℓ is positive for all ℓ.

Let us apply these results now to the special case of the square well of range a and depth V_0. From (12.41),

$$R_\ell(r) = \frac{u_\ell(r)}{r} = j_\ell(k'r)$$

where $\hbar k' = \sqrt{2m(E + V_0)}$. Hence,

$$\beta_\ell = k'a \frac{j'_\ell(k'a)}{j_\ell(k'a)} \tag{13.83}$$

For $\ell = 0$, we find from (12.30) and (12.31) the simple expressions

$$\xi_0 = -ka, \qquad \Delta_0 = -1, \qquad s_0 = ka, \qquad \beta_0 = k'a \text{ cotan } k'a - 1$$

The logarithmic derivative β_0 is a monotonically decreasing function of energy—an important property that can be proved more generally for any potential and all β_ℓ. Applying these results to (13.81), we see that the S-wave scattering amplitude is

$$f_0 = \frac{1}{k} e^{i\delta_0} \sin \delta_0 = \frac{e^{-2ika}}{k} \left(\frac{k}{k' \text{ cotan } k'a - ik} - e^{ika} \sin ka \right) \tag{13.84}$$

In the limit $E \to 0$, $k \to 0$, this gives the nonvanishing isotropic S-wave cross section:

$$\sigma_0 \to 4\pi a^2 \left(\frac{\tan k'_0 a}{k'_0 a} - 1 \right)^2 \quad \text{for} \quad k'_0 = \sqrt{\frac{2mV_0}{\hbar^2}} \tag{13.85}$$

Exercise 13.16. Compute ξ_1, Δ_1, s_1, β_1, and f_1 for P-wave scattering from a square well and examine their energy dependence.

Figures 13.7 and 13.8, calculated for a particular square well, illustrate some important common features of scattering cross sections. The phase shifts δ_0, δ_1, and δ_2, which are determined only to within multiples of π, were normalized so as to go to zero as $E \to \infty$, when the particle is effectively free. At low energies, P waves (and waves of higher angular momentum) are scattered less than S waves, because the presence of the centrifugal potential makes it improbable for a particle to be found near the center of force. Generally, the partial cross sections tend to decrease with increasing angular momentum and increasing energy, but the figures also show that the smooth variation of the phase shifts and cross sections is interrupted by a dramatic change in one of the phase shifts and a corresponding pronounced fluctuation in the partial cross section. Thus, for the particular values of the parameters on which Figures 13.7 and 13.8 are based, the P-wave phase shift rises rapidly near $ka = 0.7$. Since it passes near the value $3\pi/2$ in this energy range, $\sin^2 \delta_1$ becomes close to unity, and the partial cross section approaches its maximum value, $\sigma_1 = 12\pi/k^2$.

It may happen that in a small energy range a rapid change of the logarithmic derivative β_ℓ can be represented by a *linear* approximation,

$$\beta_\ell(E) = c + bE \tag{13.86}$$

while the quantities ξ_ℓ, Δ_ℓ, and s_ℓ, which characterize the external wave function, vary slowly and smoothly with energy, and may be regarded as constant. If this approximation is substituted in (13.80), we get

$$e^{2i(\delta_\ell - \xi_\ell)} = \frac{E - E_0 - i\dfrac{\Gamma}{2}}{E - E_0 + i\dfrac{\Gamma}{2}} \tag{13.87}$$

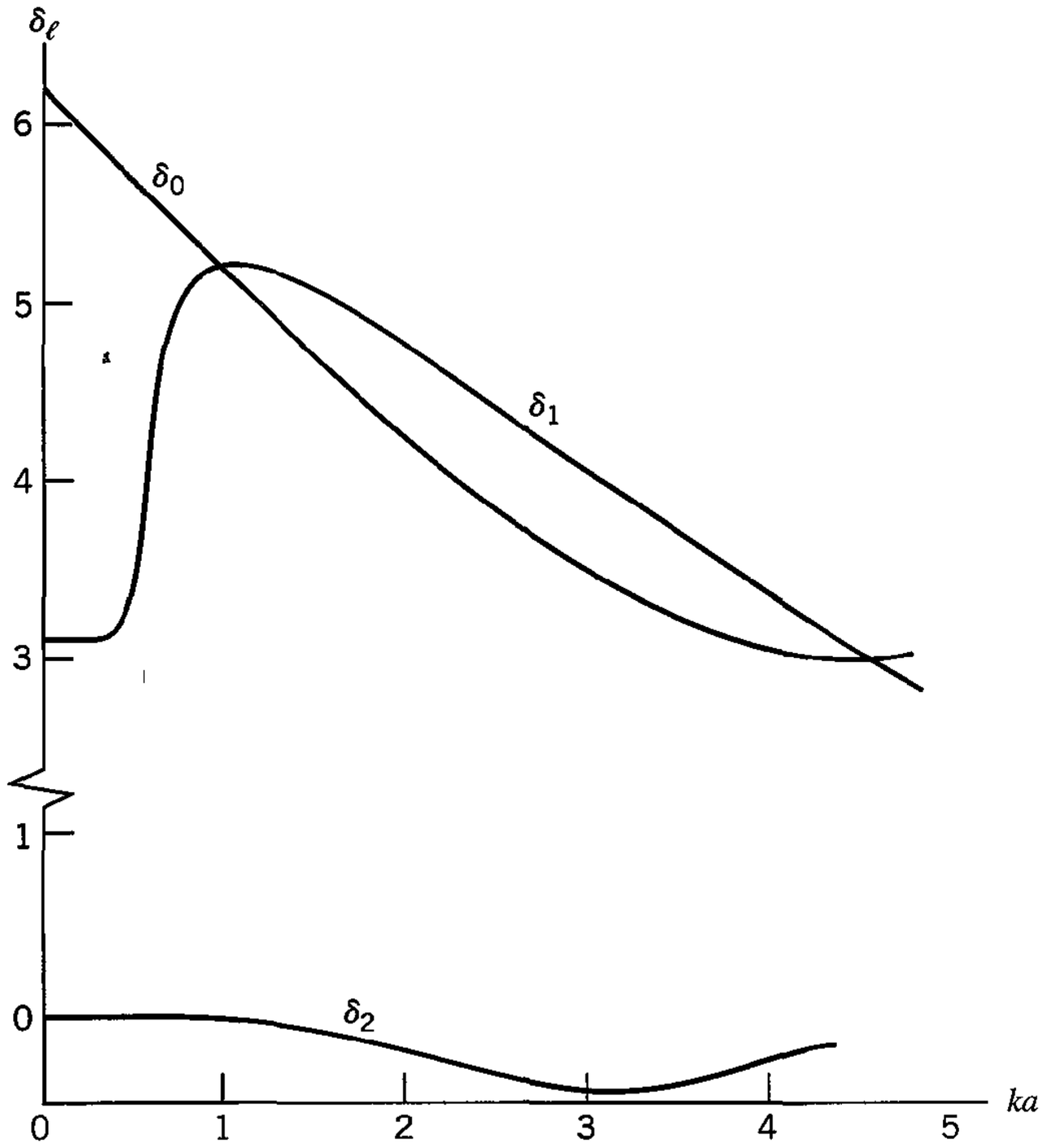


Figure 13.7 *S*, *P*, and *D* phase shifts (δ_0, δ_1, δ_2) for scattering from a square well of radius a with $k_0'a = \sqrt{2mV_0a^2/\hbar^2} = 6.2$.

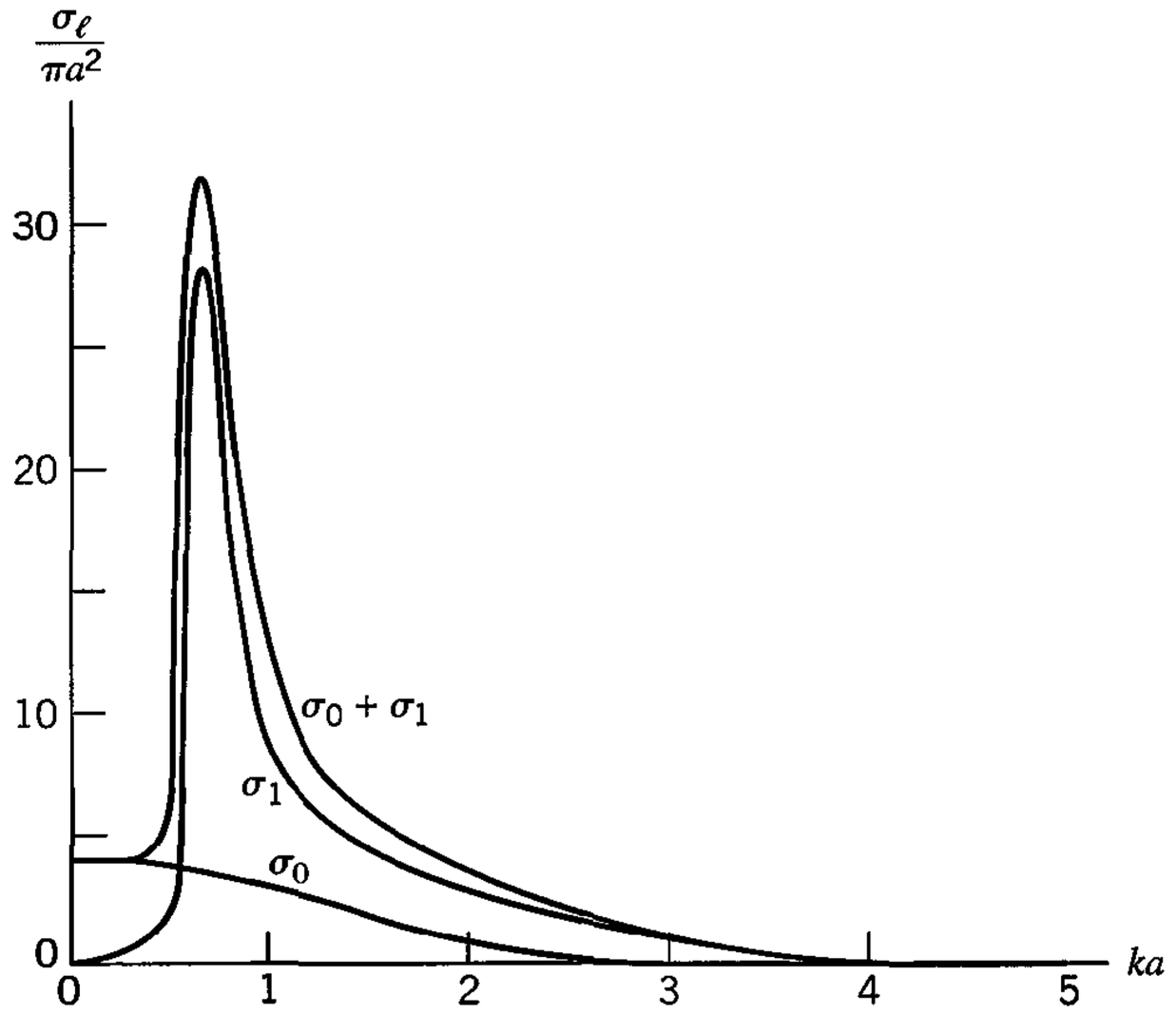


Figure 13.8. Momentum dependence of the partial cross sections (σ_0 and σ_1) for *S* and *P* waves corresponding to the phase shifts of Figure 13.7. The cross sections are given in units of πa^2.

where

$$E_0 = \frac{\Delta_\ell - c}{b} \quad \text{and} \quad \frac{\Gamma}{2} = -\frac{s_\ell}{b} \tag{13.88}$$

Hence, we arrive at the very simple approximate relation

$$\tan(\delta_\ell - \xi_\ell) = \frac{\Gamma}{2(E_0 - E)} \tag{13.89}$$

Since β_ℓ is a decreasing function of the energy, b must be negative. By its definition, s_ℓ is positive, and it thus follows that the quantity Γ defined in (13.88) is *positive*. The expressions (13.87) and (13.89) are useful if E_0 and Γ are reasonably constant and if the linear approximation (13.86) is accurate over an energy range large compared with Γ. Under these circumstances, it can be seen from (13.87) that the phase $2(\delta_\ell - \xi_\ell)$ changes by 2π as E varies from $E \ll E_0 - \frac{\Gamma}{2}$ to $E \gg E_0 + \frac{\Gamma}{2}$. Hence, if ξ_ℓ is also nearly constant in this interval, the phase shift δ_ℓ changes by π, and the partial cross section σ_ℓ, which is proportional to $\sin^2 \delta_\ell$, changes abruptly. Such sudden variations in the phase shifts are *resonances*, with E_0 being the *resonant energy* and Γ the *width* of the resonance.

If the phase shift δ_ℓ is near resonance, the contribution of the corresponding partial wave to the scattering amplitude can be written according to (13.81), (13.86) and (13.88) as

$$f_\ell = \frac{2\ell + 1}{k} e^{2i\xi_\ell} \left(\frac{\Gamma/2}{E_0 - E - i\,\Gamma/2} + e^{-i\xi_\ell} \sin \xi_\ell \right) P_\ell(\cos \theta) \tag{13.90}$$

giving a neat separation of the resonant part of the partial wave amplitude from the nonresonant part, which depends only on ξ_ℓ. If as is the case for low energies or high angular momenta, the hard sphere phase shifts are negligible, (13.89) reduces to

$$\tan \delta_\ell = \frac{\Gamma/2}{E_0 - E}$$

The resonant term in (13.90) predominates then and contributes to the total cross section an amount

$$\boxed{\sigma_\ell = \frac{4\pi}{k^2} (2\ell + 1) \sin^2 \delta_\ell = \frac{4\pi(2\ell + 1)}{k^2} \frac{\Gamma^2}{4(E - E_0)^2 + \Gamma^2}} \tag{13.91}$$

For a small width Γ, this represents a sharp maximum centered at E_0 with a symmetric shape similar to that of the transmission resonance peaks in Figure 6.8.[4] Profiles with this energy dependence are called ***Breit-Wigner*** cross sections. They are experimentally resolvable if the width Γ, while narrow, is still wide enough so that the particles in the beam may be represented by spatially broad wave packets with $\Delta E \ll \Gamma$.

Since we have discussed phase shifts and resonances in this section entirely in terms of the logarithmic derivatives at the boundary of an interior region within

[4]The *P*-wave resonance in Figure 13.8 is not accurately described by the simple formula (13.91) because the conditions under which this formula was derived ($\xi_\ell \ll 1$, Δ_1, and s_1 reasonably constant) are not well satisfied.

which the scattering forces are concentrated, our conclusions are independent of the particular mechanism that operates inside this region. The interior region is a "black box," characterized only by the values of the logarithmic derivatives at the surface. In particular, there is no need for the phase shifts to result from an interaction with a simple, fixed potential V. Rather, the scattering may arise from much more complicated interactions within the black box, and resonances may occur as a result of constructive interference at certain discrete frequencies, if some mechanism impedes the escape of the wave from the box. In Sections 6.4 and 7.4, where transmission resonances were considered, we saw how a precipitate change in the potential or a high potential barrier can produce narrow resonances, but in nuclear and particle physics resonances come about through a more complicated mechanism. For example, when a slow neutron enters a complex nucleus, it interacts in a complicated way with all the nucleons and its re-emission may thereby be considerably delayed. Owing to the wave properties of matter, such metastable states can occur only at certain discrete energies, and these are resonant energies. At low energies the neutron scattering cross section thus exhibits sharp maxima that may be regarded as the remnants of the discrete level structure which exists for $E < 0$.

The reciprocal width of a resonance $1/\Gamma$ is a direct measure of the stability of the resonant state or of the delay between the time of absorption of the particle by the black box and its re-emission. To see this, we now consider the extreme case of a very narrow resonance centered at energy E_0 and a sharply pulsed, well-collimated wave packet involving a broad energy range, ΔE, so that $\Delta E \gg \Gamma$. The analysis of Section 13.2, which assumed a slow energy variation of the scattering amplitude is no longer applicable. To simplify the discussion, without affecting any qualitative conclusions, we now assume that we have a very narrow isolated ℓ-wave resonance and that all nonresonant contributions to the scattering are negligible. The scattering amplitude is taken to be

$$f(\theta) = \frac{2\ell + 1}{k} \frac{\Gamma/2}{E_0 - E - i\,\Gamma/2} P_\ell(\cos\,\theta) \tag{13.92}$$

The wave packet is represented by a momentum wave function $\phi(\mathbf{k})$ that is appreciably different from zero only for $\mathbf{k}$ in the direction of incidence but that has an energy spread $\Delta E \gg \Gamma$. Using

$$\mathbf{k} \cdot \mathbf{r}_0 \approx -kr_0 \quad \text{and} \quad \mathbf{k} \cdot \mathbf{v}_0 \approx kv_0$$

we find that the scattered wave in Eq. (13.19) is, asymptotically, proportional to the expression

$$\psi_{\text{scatt}} \propto \int_0^\infty \frac{e^{ik(r+r_0-v_0t)}}{r} f(\theta)k^2\,dk \propto \int_{-\infty}^{+\infty} e^{i(E-E_0)(r+r_0-v_0t)/\hbar v_0} \frac{\Gamma/2}{E - E_0 + i\,\Gamma/2}\,dE \tag{13.93}$$

As in (7.68)–(7.69) and (A.22) in the Appendix, we evaluate the integral and obtain

$$\psi_{\text{scatt}} = \begin{cases} \Gamma\pi i \exp\left[\dfrac{\Gamma}{2\hbar}\left(\dfrac{r + r_0}{v_0} - t\right)\right] & v_0t > r + r_0 \\ 0 & v_0t < r + r_0 \end{cases} \tag{13.94}$$

Thus, once the scattered wave packet has reached a sphere of radius r, the probability of finding the particle at that location decreases exponentially with time. (See the analogous one-dimensional illustration in Figure 7.8). A narrow resonance corre-

sponds to a situation in which the incident particle spends a long time, of the order of the mean lifetime, $\tau = \hbar/\Gamma$, in the interaction region before being scattered.

It is possible to think of the quantity $E_0 - i\Gamma/2$, which appears in the denominator of the scattering amplitude as a complex energy of the resonant state. If the variables E and k are analytically continued into the complex plane, Eq. (13.87) shows that $E_0 - i\Gamma/2$ is a simple pole of $S_\ell(k) = e^{2i\delta_\ell(k)}$ and that the state with this complex E value has *no incoming wave.*

Exercise 13.17. Using Figure 13.8, estimate the mean lifetime of the metastable state responsible for the P-wave resonance in units of the period associated with the motion of the particle in the square well.

Exercise 13.18. Show that for a resonance the quantity $\hbar(d\delta_\ell/dE)$, evaluated at $E = E_0$ is a measure of the lifetime of the metastable state.

7. ***Phase Shifts and Green's Functions.*** Although the relation (13.76) between phase shifts and logarithmic derivatives of the radial wave functions is a very general and simple one, it does not shed any direct light on the dependence of the δ_ℓ on the scattering potential. This connection can be elucidated if a partial wave analysis is applied to the integral equation (13.39). To this end, the outgoing Green's functions must first be expanded in terms of Legendre polynomials.

Since $G_+(\mathbf{r}, \mathbf{r}')$ is a solution of the equations

$$(\nabla^2 + k^2)G_+ = 0 \quad \text{and} \quad (\nabla'^2 + k^2)G_+ = 0$$

if $\mathbf{r} \neq \mathbf{r}'$, it is seen from the separation of these equations in terms of spherical coordinates that the partial wave expansion for $r > r'$ must be of the form

$$G_+(\mathbf{r}, \mathbf{r}') = \frac{\exp(ik|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|} = \sum_{\ell=0}^{\infty} q_\ell(k) P_\ell(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') j_\ell(kr') h_\ell^{(1)}(kr) \quad (r > r') \tag{13.95}$$

where the particular choice of the spherical cylinder functions (Bessel and Hankel functions of the first kind) is dictated by the regular behavior of G_+ at $r' = 0$ and its asymptotic behavior, $G_+ \to e^{ikr}/r$ as $r \to \infty$. The remaining unknown coefficients, $q_\ell(k)$, in the expansion (13.95) can be determined by letting $r \approx r' \approx 0$ (but still $r > r'$). By using the first approximations (12.17) and (12.28) for Bessel and Hankel functions, (13.95) simplifies to

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \sum_{\ell=0}^{\infty} \frac{q_\ell(k)}{(2\ell + 1)ik} \frac{r'^\ell}{r^{\ell+1}} P_\ell(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')$$

This has to be compared with the expansion, familiar in electrostatics,

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \sum_{\ell=0}^{\infty} \frac{r'^\ell}{r^{\ell+1}} P_\ell(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}')$$

which comes from the generating function (11.68) for Legendre polynomials. Comparison of the last two equations yields the values of $q_\ell(k)$ and the desired identity:

$$\boxed{G_+(\mathbf{r}, \mathbf{r}') = \frac{\exp(ik|\mathbf{r} - \mathbf{r}'|)}{|\mathbf{r} - \mathbf{r}'|} = ik \sum_{\ell=0}^{\infty} (2\ell + 1) P_\ell(\hat{\mathbf{r}} \cdot \hat{\mathbf{r}}') j_\ell(kr') h_\ell^{(1)}(kr) \quad (r > r')} \tag{13.96}$$

We now substitute (13.96) and the partial wave expansion (13.64) of $\psi_k^{(+)}(r, \theta)$ into the integral equation (13.39), using (13.67), and carry out the integration over the direction of $\mathbf{r}'$. For $r > a$ we obtain the radial integral equation

$$e^{i\delta_\ell} \frac{u_{\ell,k}(r)}{r} = j_\ell(kr) - ike^{i\delta_\ell} \int_0^r j_\ell(kr')h_\ell^{(1)}(kr)u_{\ell,k}(r')U(r')r' \, dr' \tag{13.97}$$

Letting $r \to \infty$ and replacing $u_{\ell,k}(r)$ and the cylinder functions of argument kr by their asymptotic expansions, we finally arrive at the simple formula

$$\boxed{\sin \delta_\ell = -k \int_0^\infty j_\ell(kr')U(r')u_{\ell,k}(r')r' \, dr'} \tag{13.98}$$

This is an explicit expression for the phase shifts in terms of the potential and the radial eigenfunctions.

Exercise 13.19. Show that for *all* values of r the radial wave function $u_{\ell,k}(r)$ satisfies the integral equation

$$\begin{aligned} u_{\ell,k}(r) = r \cos \delta_\ell \, j_\ell(kr) &+ kr \int_0^r j_\ell(kr')n_\ell(kr)u_{\ell,k}(r')U(r')r' \, dr' \\ &+ kr \int_r^\infty j_\ell(kr)n_\ell(kr')u_{\ell,k}(r')U(r')r' \, dr' \end{aligned} \tag{13.99}$$

Exercise 13.20. Verify (13.98) by applying a partial wave analysis directly to the scattering amplitude (13.43) provided that $V(r)$ is a central potential.

Some useful estimates of phase shifts may be based on (13.98) and (13.99). For instance, if the potential is not strong enough to produce a resonance, these coupled equations may be solved by successive approximation in an iterative procedure. The zeroth approximation to the wave function is

$$u_{\ell,k}^{(0)}(r) = r \cos \delta_\ell j_\ell(kr)$$

When this wave function is substituted in (13.98), we get the approximate phase shift

$$\tan \delta_\ell \approx -k \int_0^\infty [j_\ell(kr')]^2 U(r')r'^2 \, dr' \tag{13.100}$$

Higher approximations may be obtained by iteration, but this is usually cumbersome. For values of $\ell > ka$, the spherical Bessel functions in the integral (13.100) may be approximated by the first term in their power series expansion [see (12.17)]. We thus obtain

$$\tan \delta_\ell \approx -\frac{2^{2\ell}(\ell!)^2}{[(2\ell + 1)!]^2} k^{2\ell+1} \int_0^\infty U(r')r'^{2\ell+2} \, dr' \tag{13.101}$$

as an estimate for the phase shifts. From this formula we deduce the rule of thumb that, for low energies and high angular momenta, $\delta_\ell \propto k^{2\ell+1}$.

Exercise 13.21. Show that a partial wave analysis of the scattering amplitude in the first Born approximation (13.47) gives the same estimate as (13.100) if $\delta_\ell \ll 1$.

8. ***Scattering in a Coulomb Field.*** Since a Coulomb field, $V = C/r$, has an infinite range, many of the results that we have derived in this chapter under the assumption of a potential with finite range are not immediately applicable without a separate investigation. For example, the concept of a total scattering cross section is meaningless for such a potential, because every incident particle, no matter how large its impact parameter, is scattered; hence, the total scattering cross section is infinite.

Exercise 13.22. Show that the total cross section obtained from the differential cross section (13.9) diverges. (See also Exercise 13.8.)

Owing to its great importance in atomic and nuclear physics, a vast amount of work has been done on the continuum ($E > 0$) eigenstates of the Coulomb potential. In this section, a particular method of solving the problem will be presented because of its intrinsic interest. This method depends on the observation that the Schrödinger equation with a Coulomb potential is separable in *parabolic coordinates*, which are defined by the equations

$$\xi = r + z \qquad \eta = r - z \qquad \varphi = \arctan \frac{y}{x} \tag{13.102}$$

The Schrödinger equation for the Coulomb potential,

$$V(r) = \frac{q_1 q_2}{r} \tag{13.103}$$

describing the electrostatic interaction between two charges q_1 and q_2, becomes in these coordinates [see Appendix, Eq. (A.58)]:

$$-\frac{\hbar^2}{2m}\left[\frac{4}{\xi + \eta}\frac{\partial}{\partial\xi}\left(\xi\frac{\partial}{\partial\xi}\right) + \frac{4}{\xi + \eta}\frac{\partial}{\partial\eta}\left(\eta\frac{\partial}{\partial\eta}\right) + \frac{1}{\xi\eta}\frac{\partial^2}{\partial\varphi^2}\right]\psi + \frac{2q_1q_2}{\xi + \eta}\psi = E\psi \tag{13.104}$$

For scattering, we are interested in a solution with axial symmetry about the direction of incidence, which we again choose to be the z axis. Hence, we must look for solutions that are independent of φ and that have the form

$$\psi(\xi, \eta) = f_1(\xi)f_2(\eta) \tag{13.105}$$

Separation of the parabolic variables ξ and η is achieved easily, and the Schrödinger equation is replaced by two *ordinary* differential equations

$$\frac{d}{d\xi}\left(\xi\frac{df_1}{d\xi}\right) + \frac{k^2}{4}\xi f_1 - c_1 f_1 = 0, \qquad \frac{d}{d\eta}\left(\eta\frac{df_2}{d\eta}\right) + \frac{k^2}{4}\eta f_2 - c_2 f_2 = 0 \tag{13.106}$$

where $\hbar^2k^2 = 2mE$, and the constants of separation, c_1 and c_2, are related by the condition

$$c_1 + c_2 = \frac{q_1 q_2 m}{\hbar^2} \tag{13.107}$$

Scattering is described by a wave function that asymptotically has an incident plane wave part,

$$\exp(ikz) = \exp[ik\,(\xi - \eta)/2] \tag{13.108}$$

and a radially outgoing part, proportional to

$$\exp(ikr) = \exp[ik\,(\xi + \eta)/2] \tag{13.109}$$

The behavior of these two portions of the wave function suggests that we look for a particular solution with

$$f_1(\xi) = e^{ik\xi/2} \tag{13.110}$$

Substituting this into (13.106), we see that the *Ansatz* (13.110) is indeed a solution if we choose

$$c_1 = i\frac{k}{2}$$

The remaining equation for f_2,

$$\frac{d}{d\eta}\left(\eta\frac{df_2}{d\eta}\right) + \frac{k^2}{4}\eta f_2 - \left(\frac{q_1q_2m}{\hbar^2} - i\frac{k}{2}\right)f_2 = 0 \tag{13.111}$$

is conveniently transformed if we define a new function $g(\eta)$:

$$f_2(\eta) = \exp(-ik\eta/2)\, g(\eta) \tag{13.112}$$

For a stationary *bound* state, we know from Chapter 12 that for an attractive Coulomb potential,

$$E_n = -\frac{q_1^2q_2^2m}{2\hbar^2n^2} = -\langle T\rangle$$

where in the last equality the virial theorem (Exercise 3.26) has been used. If the mean kinetic energy is expressed as $mv^2/2$, defining an effective velocity v, we obtain the relation

$$n = -\frac{q_1q_2}{\hbar v} \tag{13.113}$$

linking the principal quantum number n with the effective orbital speed v. If we apply the same defining relation (13.113) to the positive energy *unbound* scattering states in a Coulomb potential, a generally noninteger, positive or negative, quantum number n is defined. This allows us to cast the differential equation for $g(\eta)$ in the form

$$\eta\frac{d^2g}{d\eta^2} + (1 - ik\eta)\frac{dg}{d\eta} + nkg = 0 \tag{13.114}$$

which should be compared with Eq. (12.89) for the confluent hypergeometric function. We see that the solution of (13.114), which is regular at the origin, is

$$g(\eta) = {}_1F_1(in;\, 1;\, ik\eta) \tag{13.115}$$

For purely imaginary argument z, the asymptotic expansion of this function has the leading term[5]

$${}_1F_1(a;\, c;\, z) \cong \frac{\Gamma(c)}{\Gamma(c-a)}|z|^{-a}\exp\left(ia\frac{\pi}{2}\right) + \frac{\Gamma(c)}{\Gamma(a)}|z|^{a-c}e^z\exp\left[i(a-c)\frac{\pi}{2}\right] \tag{13.116}$$

[5]Morse and Feshbach (1953), Eq. (5.3.51).

Applying this approximation to $g(\eta)$, we obtain for large values of η,

$$g(\eta) \cong \frac{\exp(-n\pi/2)}{|\Gamma(1+in)|}\left\{\exp[-i(n\ln k\eta - \sigma_n)] + \frac{n}{k\eta}\exp[i(k\eta + n\ln k\eta - \sigma_n)]\right\} \tag{13.117}$$

where $\Gamma(1+in) = |\Gamma(1+in)|e^{i\sigma_n}$.

Collecting all these results together, we conclude that the particular solution

$$\psi(\xi, \eta, \varphi) = \exp[ik(\xi - \eta)/2]_1F_1(in; 1; ik\eta) \tag{13.118}$$

of the Schrödinger equation has the asymptotic form

$$\psi \cong \frac{\exp(-n\pi/2)}{|\Gamma(1+in)|}\left\{\exp\{i[kz - n\ln k(r-z) + \sigma_n]\} + \frac{n}{k(r-z)}\exp\{i[kr + n\ln k(r-z) - \sigma_n]\}\right\} \tag{13.119}$$

for large values of $r - z = r(1 - \cos\theta)$, where θ is the scattering angle. Equation (13.119) is valid at large distance from the scattering center, except in the forward direction. It shows clearly why it is impossible to obtain for a Coulomb potential an asymptotic eigenfunction that has the simple form,

$$\psi \cong N\left[e^{ikz} + f_k(\theta)\frac{e^{ikr}}{r}\right]$$

deduced for a potential of finite range. The Coulomb force is effective even at very great distances and prevents the incident wave from ever approaching a plane wave. Similarly, the scattered wave fails to approach the simple free particle form, e^{ikr}/r, as $r \to \infty$. However, it is important to note that the modification, for which the Coulomb potential is responsible at large distances from the scattering center, affects only the phase of the incident and scattered waves. Since the asymptotic contributions from the Coulomb potential to the phases are logarithmic functions of r, and hence vary but slowly, the analysis of the fate of an incident wave packet in a scattering process can still be carried through, and the result is precisely the same as before. The differential cross section is [see (13.23)]

$$\frac{d\sigma}{d\Omega} = |f_k(\theta)|^2$$

where the scattering amplitude $f_k(\theta)$ is again the factor multiplying the radial part of the asymptotically outgoing wave, provided the incident (modified) plane wave is normalized so that the probability density is unity. Hence, for all finite values of the scattering angle θ,

$$\boxed{f_k(\theta) = \frac{n}{k(1-\cos\theta)}\exp[in\ln(1-\cos\theta)]} \tag{13.120}$$

and the differential scattering cross section is

$$\frac{d\sigma}{d\Omega} = \frac{n^2}{4k^2\sin^4(\theta/2)} = \frac{q_1^2q_2^2}{16E^2}\frac{1}{\sin^4(\theta/2)} \tag{13.121}$$

in exact agreement with the classical Rutherford scattering cross section (13.9) and the Born approximation. An angle-independent addition to the phase has been omitted in (13.120).

Since ${}_1F_1(a; c; 0) = 1$, the normalization of the eigenfunction (13.118) is such that $\psi = 1$ at the scattering center (origin). On the other hand, the incident wave in the large brace of (13.119) is normalized to probability density unity far from the scattering center, and

$$|\psi(\infty)|^2 = \frac{\exp(-n\pi)}{|\Gamma(1 + in)|^2}.$$

Hence, the penetration probability of finding the particle at the origin relative to the probability of finding it in the incident beam is

$$\frac{|\psi(0)|^2}{|\psi(\infty)|^2} = |\Gamma(1 + in)|^2 e^{n\pi} = \frac{2\pi n}{1 - e^{-2\pi n}} \tag{13.122}$$

where in obtaining the last expression use was made of (8.90) and the definition of the Γ function: $\Gamma(1 + z) = z\Gamma(z)$.

If the Coulomb potential is repulsive ($q_1 q_2 > 0$ and $n < 0$) and strong compared with the kinetic energy of the particle, we have as a measure of penetration to the origin,

$$\frac{|\psi(0)|^2}{|\psi(\infty)|^2} = -2\pi n e^{2\pi n} \quad (n \ll 0) \tag{13.123}$$

The significance of the penetrability,

$$e^{2\pi n} = \exp\left(-\frac{2\pi q_1 q_2}{\hbar v}\right) \equiv e^{-2G} \tag{13.124}$$

was already discussed in connection with Eq. (7.52) for a one-dimensional WKB model of Coulomb scattering. The exponent G is called the *Gamow factor* in nuclear physics.

Exercise 13.23. Discuss other limiting cases (fast particles, attractive potentials) for (13.122).

Exercise 13.24. Calculate the wave function (13.118) in the forward direction. What physical conclusions can you draw from its form?

Since the Schrödinger equation is separable in parabolic coordinates only if the potential behaves strictly as $1/r$ at *all* distances, the method of this section is not appropriate if the potential is Coulombic only at large distances but, as in the case of nuclear interactions with charged hadrons, has a different radial dependence near the origin. It is then preferable to use spherical polar coordinates and attempt a phase shift analysis. By expanding the eigenfunction (13.118) in terms of Legendre polynomials, phase shifts can be calculated for the Coulomb potential, and the theory of Section 13.5 can be extended to include the presence of a long-range $1/r$ potential.

Problems

1. Using the first three partial waves, compute and display on a polar graph the differential cross section for an impenetrable hard sphere when the de Broglie wavelength of the incident particle equals the circumference of the sphere. Evaluate the total cross section and estimate the accuracy of the result. Also discuss what happens if the wavelength becomes very large compared with the size of the sphere.
2. If the scattering potential has the translation invariance property $V(\mathbf{r} + \mathbf{R}) = V(\mathbf{r})$, where $\mathbf{R}$ is a constant vector, show that in the first Born approximation, as in the exact formulation (Exercise 13.7), scattering occurs only when the momentum transfer $\mathbf{q}$ (in units of $\hbar$) equals a reciprocal lattice vector $\mathbf{G}$.

CHAPTER 14

The Principles of Quantum Dynamics

If the Hamiltonian operator is known, an initial wave function $\psi(\mathbf{r}, 0)$ develops in time into $\psi(\mathbf{r}, t)$ according to the time-dependent Schrödinger equation (Chapters 2 and 3). This algorithm for calculating the future behavior of a wave packet from its past history was used in Chapter 13 for scattering calculations. We now extend the general principles of quantum mechanics (Chapters 9 and 10) to the laws governing the time evolution of quantum systems, utilizing several equivalent dynamical pictures (Schrödinger's, Heisenberg's, Dirac's). The canonical quantization of systems with classical analogues is discussed and applied to the forced harmonic oscillator, which is the prototype of a system in interaction with its environment.

1. ***The Evolution of Probability Amplitudes and the Time Development Operator.*** We now add the time parameter to the description of quantum states and generalize the fundamental postulate of quantum mechanics (Section 9.1) by asserting that:

> *The maximum information about the outcome of physical measurements on a system at time t is contained in the probability amplitudes* $\langle K_i|\Psi(t)\rangle$, *which correspond to a complete set of observables K for the system.*

The only new feature here is that we now recognize formally that the state Ψ is a function of time.

For the simple system of a particle with coordinates x, y, z as observables, the amplitude is

$$\psi(\mathbf{r}, t) = \langle x, y, z|\Psi(t)\rangle \tag{14.1}$$

which is the time-dependent wave function in the coordinate representation. The same state is represented in the momentum representation as

$$\phi(\mathbf{p}, t) = \langle p_x, p_y, p_z|\Psi(t)\rangle \tag{14.2}$$

The basic question of quantum dynamics is this: Given an initial state $|\Psi(t_0)\rangle$ of the system, how is the state $|\Psi(t)\rangle$ at time t determined from this, if indeed it is so determined? Or, in terms of the amplitudes that specify the state, how do the amplitudes $\langle L_j|\Psi(t)\rangle$ evolve in time from the initial amplitudes $\langle K_i|\Psi(t_0)\rangle$?

The assertion that $|\Psi(t_0)\rangle$ determines $|\Psi(t)\rangle$ is the quantum mechanical form of the *principle of causality*, and we shall assume it. The dynamical law that connects the initial and final amplitudes is contained in the further assumption that the composition rule (9.8) can be generalized to the time-dependent form

$$\boxed{\langle L_j|\Psi(t)\rangle = \sum_i \langle L_j|T(t, t_0)|K_i\rangle\langle K_i|\Psi(t_0)\rangle} \tag{14.3}$$

where the coefficients

$$\langle L_j|T(t, t_0)|K_i\rangle \tag{14.4}$$

are independent of the state $|\Psi(t_0)\rangle$. They have a direct and simple interpretation: The expression (14.4) signifies the probability amplitude for finding the system at time t in the eigenstate of $|L_j\rangle$ of the observables symbolized by L, if at time t_0 it was known to be in the eigenstate $|K_i\rangle$ of the observables K. This quantity is called a *transition amplitude*. With forethought it has been written in the form of a matrix element of an operator $T(t, t_0)$, because from (14.3), which is valid for any state $|\Psi(t_0)\rangle$, we can derive the transformation equation

$$\langle L_j|T(t, t_0)|K_i\rangle = \sum_p \sum_\ell \langle L_j|N_p\rangle\langle N_p|T(t, t_0)|M_\ell\rangle\langle M_\ell|K_i\rangle \tag{14.5}$$

which shows that (14.4) defines a representation-independent linear operator $T(t, t_0)$. Equation (14.3) is consistent with the composition rule (9.8) if we require that

$$\langle L_j|T(t, t)|K_i\rangle = \langle L_j|K_i\rangle \quad \text{for all times } t$$

or

$$T(t, t) = 1 \tag{14.6}$$

Exercise 14.1. Prove the relation (14.5) from (14.3) without *assuming* that (14.4) is the matrix element of an operator.

It now follows from the composition rule (14.3) that the *time development* or *evolution operator* $T(t, t_0)$ relates the initial state $|\Psi(t_0)\rangle$ to the final state $|\Psi(t)\rangle$ according to

$$\boxed{|\Psi(t)\rangle = T(t, t_0)|\Psi(t_0)\rangle} \tag{14.7}$$

Since $T(t, t_0)$ does not depend on $|\Psi(t_0)\rangle$, the *principle of superposition* applies to the time development of states. This means that if $|\Psi_a(t_0)\rangle$ and $|\Psi_b(t_0)\rangle$ separately evolve into $|\Psi_a(t)\rangle$ and $|\Psi_b(t)\rangle$, then a superposition $c_a|\Psi_a(t_0)\rangle + c_b|\Psi_b(t_0)\rangle$ develops into $c_a|\Psi_a(t)\rangle + c_b|\Psi_b(t)\rangle$, i.e., each component of the state moves independently of all the others, expressing the fundamental *linearity* of quantum dynamics.

From (14.7) it follows that

$$|\Psi(t_2)\rangle = T(t_2, t_1)|\Psi(t_1)\rangle = T(t_2, t_1)T(t_1, t_0)|\Psi(t_0)\rangle = T(t_2, t_0)|\Psi(t_0)\rangle$$

Hence, the time development operator has the property

$$\boxed{T(t_2, t_0) = T(t_2, t_1)T(t_1, t_0)} \tag{14.8}$$

From (14.6) and (14.8) we infer that

$$T(t, t_0)T(t_0, t) = T(t_0, t)T(t, t_0) = 1$$

or

$$[T(t, t_0)]^{-1} = T(t_0, t) \tag{14.9}$$

For small ε we may write

$$T(t + \varepsilon, t) = 1 - \frac{i}{\hbar}\varepsilon H(t) \tag{14.10}$$

defining an operator $H(t)$. (The reason for introducing the factor $i/\hbar$, apparently capriciously, will become evident forthwith.) Since, by (14.8),

$$T(t + \varepsilon, t_0) = T(t + \varepsilon, t)T(t, t_0)$$

we have with (14.10) the differential equation for T,

$$\frac{dT(t, t_0)}{dt} = \lim_{\varepsilon \to 0} \frac{T(t + \varepsilon, t_0) - T(t, t_0)}{\varepsilon} = -\frac{i}{\hbar} H(t)T(t, t_0)$$

or

$$\boxed{i\hbar \frac{dT(t, t_0)}{dt} = H(t)T(t, t_0)} \tag{14.11}$$

with the initial condition $T(t_0, t_0) = 1$.

The linear operator $H(t)$ is characteristic of the physical system under consideration. We will see that it is analogous to the Hamiltonian function in classical mechanics. This analogy has led to the name *Hamiltonian* operator for $H(t)$, even when the system has no classical counterpart.

We also have

$$|\Psi(t + \varepsilon)\rangle = T(t + \varepsilon, t)|\Psi(t)\rangle$$

or, to first order in ε,

$$|\Psi(t)\rangle + \varepsilon \frac{d}{dt}|\Psi(t)\rangle = \left[1 - \frac{i}{\hbar} \varepsilon H(t)\right]|\Psi(t)\rangle$$

Hence,

$$\boxed{i\hbar \frac{d}{dt}|\Psi(t)\rangle = H(t)|\Psi(t)\rangle} \tag{14.12}$$

Its bra form is

$$-i\hbar \frac{d}{dt}\langle\Psi(t)| = \langle\Psi(t)|H^\dagger(t) \tag{14.13}$$

Equation (14.12) is the ***equation of motion*** for the state vector, giving the general law of motion for any Hamiltonian system.

To specialize to a particular system, we must select an appropriate Hamiltonian operator. The form of (14.12) is reminiscent of the time-dependent Schrödinger equation (3.42). This is no accident, for we have merely reformulated in abstract language those fundamental assumptions that have already proved their worth in wave mechanics. Of course, (14.12) is an equation for the *state vector* rather than for the wave function (14.1), but the distinction is one of generality only. It now becomes clear why Planck's constant was introduced in (14.10). In Chapter 15, we will see how the laws of wave mechanics derive from the general theory.

In wave mechanics, the differential operator H was generally Hermitian, as it must be whenever the Hamiltonian corresponds to the energy operator of the system. Generally, we will assume that H is *Hermitian*. A non-Hermitian Hamiltonian operator, like a complex index of refraction in optics, can be useful for describing the

dynamics of dissipative systems which, through absorption or decay, exchange energy with their environment.

If $H(t)$ is Hermitian, the equation adjoint to (14.11) becomes

$$-i\hbar \frac{dT^\dagger(t, t_0)}{dt} = T^\dagger(t, t_0)H(t) \tag{14.14}$$

and (14.13) becomes

$$-i\hbar \frac{d}{dt} \langle \Psi(t)| = \langle \Psi(t)|H(t) \tag{14.15}$$

By multiplying (14.11) on the left by $T^\dagger(t, t_0)$ and (14.14) on the right by $T(t, t_0)$ and subtracting the two equations, we get

$$\frac{d}{dt} [T^\dagger(t, t_0)T(t, t_0)] = 0$$

Since, by (14.6), the product of the two operators equals the identity at $t = t_0$, it follows that the time development operator is unitary:

$$T^\dagger(t, t_0)T(t, t_0) = 1 \tag{14.16}$$

Hence, the norm of any state vector remains unaltered during the motion. If $|\Psi(t_0)\rangle$ is normalized to unity, such that $\langle \Psi(t_0)|\Psi(t_0)\rangle = 1$, then the normalization will be preserved in the course of time, and we have from (14.7) and (14.16) that

$$\langle \Psi(t)|\Psi(t)\rangle = 1 \quad \text{for all times } t$$

consistent with the assumption that $|\langle L_j|\Psi(t)\rangle|^2$ is the probability of finding the observable L to have the value L_j at time t.

Often H does not depend on the time, and then T can be obtained for finite time intervals by applying the rule (14.8) repeatedly to n intervals, each of length $\varepsilon = (t - t_0)/n$. Hence, by (14.10) we have, with the initial condition $T(t_0, t_0) = 1$,

$$T(t, t_0) = \lim_{\substack{\varepsilon \to 0 \\ n \to \infty}} \left(1 - \frac{i}{\hbar} \varepsilon H\right)^n = \lim_{n \to \infty} \left[1 - \frac{i}{\hbar} \frac{(t - t_0)}{n} H\right]^n$$

In the limit we get by the definition of the exponential function,

$$T(t, t_0) = \exp\left[-\frac{i}{\hbar} (t - t_0)H\right] \tag{14.17}$$

It is obvious that T is unitary if H is Hermitian.

Quantum dynamics is a general framework and contains no unambiguous prescription for the construction of the operator H whose existence it asserts. The Hamiltonian operator must be found on the basis of experience, using the clues provided by the classical description, if one is available. Physical insight is required to make a judicious choice of operators to be used in the description of the system (such as coordinates, momenta, and spin variables) and to construct the Hamiltonian in terms of these variables.

Contact with measurable quantities and classical concepts can be established if we calculate the time derivative of the expectation value of an operator A, which may itself vary with time:

$$i\hbar \frac{d}{dt} \langle \Psi(t)|A|\Psi(t)\rangle = -\langle \Psi(t)|HA|\Psi(t)\rangle + \langle \Psi(t)|AH|\Psi(t)\rangle + i\hbar \langle \Psi(t)| \frac{\partial A}{\partial t} |\Psi(t)\rangle$$

or

$$\boxed{i\hbar \frac{d}{dt}\langle A\rangle = \langle AH - HA\rangle + i\hbar\left\langle \frac{\partial A}{\partial t}\right\rangle = \langle [A, H]\rangle + i\hbar\left\langle \frac{\partial A}{\partial t}\right\rangle} \tag{14.18}$$

where, as usual, the brackets $\langle \ldots \rangle$ signify expectation values of the operators enclosed. We see that the commutation relations of H with observables play an important role in the theory. If A is independent of time *and* commutes with H, the expectation value of A is constant, and A is said to be a *constant of the motion*.

A special example of a time-dependent operator is the *density operator* for the state $|\Psi(t)\rangle$,

$$\rho = |\Psi(t)\rangle\langle\Psi(t)| \tag{14.19}$$

From the equations of motion (14.12) and (14.15),

$$i\hbar \frac{\partial \rho}{\partial t} = H|\Psi(t)\rangle\langle\Psi(t)| - |\Psi(t)\rangle\langle\Psi(t)|H = [H, \rho] \tag{14.20}$$

Hence, (14.18) gives the simple result

$$\frac{d}{dt}\langle\rho\rangle = 0 \tag{14.21}$$

which is not surprising in view of the conservation of probability. The definition (14.19) implies that

$$\langle [H, \rho]\rangle = \langle\Psi(t)|[H, |\Psi(t)\rangle\langle\Psi(t)|]|\Psi(t)\rangle = 0 \tag{14.22}$$

which, by (14.20), leads to the conclusion:

$$\left\langle \frac{\partial \rho}{\partial t}\right\rangle = 0 \tag{14.23}$$

These results have sometimes led to the oxymoronic proposition that the density operator is a "time-dependent constant of the motion"!

The formal relations derived in this section are all rooted in what we have called the quantum mechanical *principle of causality*, which states that the probability amplitude for finding the value of K_i of the generic observables K at time t can be written as the inner product $\langle K_i|\Psi(t)\rangle$. The observables K and their eigenvectors are regarded as constant. Physically, this implies that a system represented by the *same* state vector at two different times has the same statistical distribution at the two times with regard to all observables of the system. In other words, $|\Psi(t)\rangle$ completely characterizes the state of a system at time t, which was the fundamental assumption made in Chapter 9. We will now discuss other, equivalent, formulations of quantum dynamics.

2. ***The Pictures of Quantum Dynamics.*** The mathematical formulation of quantum dynamics given in the last section is not unique. There we showed that the state vector of a system alone may be made responsible for the time evolution, but this is not the only way of dealing with dynamics in the framework of vector spaces. State vectors themselves are not observable or measurable quantities. Rather, the eigenvalues of Hermitian operators and probability amplitudes such as (14.3) and (14.4) are the physically significant objects of quantum mechanics. Comparison with

observation is made in terms of the eigenvalues of observables and of expansion coefficients (probability amplitudes), which are inner products in the abstract state vector space. Measuring an observable L at time t means finding one of its eigenvalues L_j, the probability of the particular result being given by $|\langle L_j|\Psi(t)\rangle|^2$ if $|\Psi(t)\rangle$ denotes the state of the system at time t.

It follows that two vector space formulations, or *pictures*, are equivalent and equally acceptable provided that

(a) *the operators corresponding to the observables maintain their eigenvalue spectra in the two formulations,* and

(b) *the inner products of physical states with eigenvectors for the same eigenvalues are equal in the two pictures.*

It follows from (a) and (b) that all expectation values remain the same from one picture to the other.

Starting from the *Schrödinger picture* of quantum dynamics, a new picture is obtained, satisfying conditions (a) and (b), by applying a time-dependent unitary transformation $U(t)$ to all states and operators. All the state vectors change from $|\Psi(t)\rangle$ to $U(t)|\Psi(t)\rangle$, and every operator A is transformed into $U(t)AU^\dagger(t)$. Owing to the unitary property of $U(t)$, all eigenvalues and all inner products remain *invariant*, but the eigenvectors of an observable change from $|A'\rangle$ to $U(t)|A'\rangle$. Expectation values remain unchanged.

The simplest choice for the transformation $U(t)$ is

$$U(t) = T(0, t) = T^\dagger(t, 0) \tag{14.24}$$

Since

$$U(t)|\Psi(t)\rangle = T(0, t)|\Psi(t)\rangle = |\Psi(0)\rangle \tag{14.25}$$

this transformation has the effect of referring all state vectors back to their values at the initial time 0. In this formulation, called the *Heisenberg* picture, the state vectors are constant in time, and we denote them as

$$\boxed{|\Psi(0)\rangle = |\overline{\Psi}\rangle} \tag{14.26}$$

The observables, which in the Schrödinger picture were represented by operators fixed in time (unless they happened to have an explicit time dependence), are represented in the Heisenberg picture by time-dependent operators,

$$\overline{L}(t) = U(t)LU^\dagger(t) = T^\dagger(t, 0)LT(t, 0) \tag{14.27}$$

At the initial time $t = 0$,

$$\overline{L}(0) = L \tag{14.28}$$

These equations exhibit the evolution of the Heisenberg operators. To identify the Heisenberg picture, the bar notation, as in $|\overline{\Psi}\rangle$ and $\overline{L}(t)$, will be used only in this chapter. Elsewhere in this book, when there is no ambiguity, we will simply use $|\Psi\rangle$ and $L(t)$ for Heisenberg states and operators.

As a special case, note that if H is constant and energy is conserved, then

$$T(t, 0) = \exp\left(-\frac{i}{\hbar}Ht\right) \tag{14.29}$$

and $\overline{H}(t) = H$. If H is constant, the Hamiltonian operator does not change in time, even in the Heisenberg picture.

By differentiating (14.27) with respect to t and using the equation of motion (14.11) for the time development operator, we obtain the Heisenberg equation of motion for an observable:

$$\boxed{i\hbar \frac{d}{dt}\overline{L}(t) = [\overline{L}(t), \overline{H}(t)] + i\hbar \frac{\partial \overline{L}(t)}{\partial t}} \tag{14.30}$$

The last term arises from the definition

$$\frac{\partial \overline{L}(t)}{\partial t} = T^{\dagger}(t, 0) \frac{\partial L}{\partial t} T(t, 0) \tag{14.31}$$

in the event that L is explicitly time-dependent. To emphasize that $\partial \overline{L}(t)/\partial t$ is the Heisenberg form of the derivative of the explicitly time-dependent operator L, strictly speaking we should write this operator as

$$\overline{\frac{\partial L}{\partial t}} = \frac{\partial \overline{L}(t)}{\partial t} \tag{14.32}$$

The expectation value of the operator equation (14.30) in the Heisenberg picture, where the states are constant in time, is the counterpart of Eq. (14.18) in the Schrödinger picture.

Eigenvectors of observables, corresponding to the same eigenvalues, differ in the Schrödinger and Heisenberg pictures by the unitary transformation $U(t)$:

$$\overline{|L_j,t\rangle} = T(0, t)|L_j\rangle \tag{14.33}$$

or

$$|L_j\rangle = T(t, 0)\overline{|L_j\ t\rangle} \tag{14.34}$$

Differentiating this last equation with respect to t, we obtain

$$\frac{dT(t, 0)}{dt}\overline{|L_j, t\rangle} + T(t, 0)\frac{d}{dt}\overline{|L_j, t\rangle} = 0$$

or, using (14.11),

$$i\hbar \frac{d}{dt}\overline{|L_j, t\rangle} = -H\overline{|L_j, t\rangle} \tag{14.35}$$

which is very similar to the equation of motion (14.2) in the Schrödinger picture, except for the all-important minus sign. Its appearance shows that if in the Schrödinger picture we regard the state vectors as ''rotating'' in a certain direction in abstract vector space and the operators with their eigenvectors as fixed, then in the Heisenberg picture the state vectors stand still and the operators with the eigenvectors ''rotate'' in the opposite direction. But the mutual relation between state vectors and operators is the same in the two pictures. They are related to each other in much the same way as the two kinematic descriptions of the rotation of a rigid body with respect to a reference frame. We can consider the body moving in a fixed frame, or the body as being at rest, with the reference frame rotating backward. Since the two pictures are connected by a unitary transformation, the probability amplitudes (inner products) are equal:

$$\langle L_j|\Psi(t)\rangle = \overline{\langle L_j,t|\Psi\rangle} \tag{14.36}$$

Exercise 14.2. Find the equation of motion for the Heisenberg bra $\overline{\langle L_j, t|}$.

Exercise 14.3. Show that in the Heisenberg picture the density operator for state $|\Psi\rangle$,

$$\bar{\rho} = |\Psi(0)\rangle\langle\Psi(0)| = |\overline{\Psi}\rangle\langle\overline{\Psi}| \tag{14.37}$$

satisfies the equations

$$i\hbar \frac{\partial \bar{\rho}}{\partial t} = i\hbar T^\dagger(t, 0) \frac{\partial \rho}{\partial t} T(t, 0) = [\bar{H}, \bar{\rho}] \qquad \text{and} \qquad \frac{d\bar{\rho}}{dt} = 0 \tag{14.38}$$

and that the expectation value of $\bar{\rho}$ is constant as in (14.21).

Instead of attributing the time evolution to either the state vectors or the operators, as in the Schrödinger and Heisenberg pictures, it is obviously possible to partition the time development operator in infinitely many different ways. We can arrange to let both the state vectors and the observables carry a complementary share of the time development. If the system and its Hamiltonian are complicated, it is often sensible to choose for $U(t)$ the adjoint of the time development operator for a suitably chosen *model Hamiltonian* H_0, which is the solution of

$$i\hbar \frac{dU(t)}{dt} = -U(t)H_0 \tag{14.39}$$

subject to the initial condition $U(0) = 1$. Generally, H_0 may be time-dependent.

The observables are then transformed into new operators

$$\tilde{L}(t) = U(t)LU^\dagger(t) \tag{14.40}$$

and these satisfy the equation of motion

$$\boxed{i\hbar \frac{d}{dt} \tilde{L}(t) = [\tilde{L}(t), \tilde{H}_0] + i\hbar \frac{\partial \tilde{L}(t)}{\partial t}} \tag{14.41}$$

More important is the equation for the transformed state vector,

$$|\tilde{\Psi}(t)\rangle = U(t)|\Psi(t)\rangle \tag{14.42}$$

By differentiating this equation with respect to t, and using (14.12), we obtain

$$i\hbar \frac{d}{dt} |\tilde{\Psi}(t)\rangle = -\tilde{H}_0|\tilde{\Psi}(t)\rangle + U(t)HU^\dagger(t)|\tilde{\Psi}(t)\rangle = U(t)(H - H_0)U^\dagger(t)|\tilde{\Psi}(t)\rangle$$

We define an interaction term V as the difference between the Hamiltonian H and the model Hamiltonian H_0 by

$$H = H_0 + V \tag{14.43}$$

giving us the simple-looking formula

$$\boxed{i\hbar \frac{d}{dt} |\tilde{\Psi}(t)\rangle = \tilde{V}(t)|\tilde{\Psi}(t)\rangle} \tag{14.44}$$

where

$$\tilde{V}(t) = U(t)VU^\dagger(t) \tag{14.45}$$

is the transformed version of V. The resulting formulation of quantum dynamics is known as the *interaction* (or *Dirac*) *picture*. Note that if the model Hamiltonian H_0 is time-independent and thus conservative,

$$U(t) = \exp\left(\frac{i}{\hbar} H_0 t\right) \tag{14.46}$$

and

$$\tilde{H}_0(t) = H_0 \tag{14.47}$$

If H_0 is chosen to be the full Hamiltonian, $H_0 = H$, the interaction picture coalesces with the Heisenberg picture. If $H_0 = 0$, the Schrödinger picture is recovered. The interaction picture will be found useful when we consider time-dependent perturbation theory in Chapter 19.

To demonstrate the total equivalence of the various pictures of quantum dynamics, let us suppose that at time t_1 the system has the definite sharp value A' for the observable A. We ask: "What is the probability that at time t_2 the system will have value B'' if the observable B is measured?" The answer is that the required probability is the square of the absolute value of the amplitude

$$\langle \tilde{B}'', t_2 | \tilde{T}(t_2, t_1) | \tilde{A}', t_1 \rangle$$

where $\tilde{T}(t_2, t_1)$ is the time development operator for the state vector in the interaction (or really a generic) picture as defined by

$$|\tilde{\Psi}(t)\rangle = \tilde{T}(t, t_0) |\tilde{\Psi}(t_0)\rangle \tag{14.48}$$

From this it follows, with (14.42), that

$$\tilde{T}(t_2, t_1) = U(t_2) T(t_2, t_1) U^\dagger(t_1) \tag{14.49}$$

Since we infer from (14.42) that

$$|\tilde{A}', t\rangle = U(t)|A'\rangle \quad \text{and} \quad \langle \tilde{B}'', t| = \langle B''| U^\dagger(t) \tag{14.50}$$

we see that the transition amplitude can be expressed equivalently in all pictures as

$$\underbrace{\langle \tilde{B}'', t_2 | \tilde{T}(t_2, t_1) | \tilde{A}', t_1 \rangle}_{\text{Interaction}} = \underbrace{\langle B'' | T(t_2, t_1) | A' \rangle}_{\text{Schrödinger}} = \underbrace{\langle \overline{B'', t_2} | | \overline{A', t_1} \rangle = \langle \overline{B'', t_2} | \overline{A', t_1} \rangle}_{\text{Heisenberg picture}} \tag{14.51}$$

The distinctiveness of the Schrödinger and Heisenberg pictures is manifested by the important fact that the Hamiltonian H (energy) is the same in both pictures.

Exercise 14.4. Show that the expression

$$\langle \tilde{B}'', t_2 | \tilde{T}(t_2, t_1) | \tilde{A}', t_1 \rangle$$

for the transition amplitude is quite general and gives the correct answer if the Schrödinger ($H_0 = 0$) or Heisenberg ($H_0 = H$) pictures are employed.

3. ***Quantization Postulates for a Particle.*** Let us now apply the general equations of quantum dynamics to the special case of a point particle with a mass m. We are concerned with the quantum behavior of this system, but it does have a classical analogue—Newtonian mechanics, or its more sophisticated Lagrangian or Hamil-

tonian forms. The position operators x, y, z are assumed to form a complete set of commuting observables for this physical system.

For the purposes of this chapter it is worthwhile to distinguish in the notation between the *classical* observables, x, y, z, which represent numbers, and the corresponding *quantum* observables, x, y, z, which stand for operators.

Setting $A = x$ in (14.18), we obtain

$$\frac{d\langle x\rangle}{dt} = \frac{\langle xH - Hx\rangle}{i\hbar} \tag{14.52}$$

On the left-hand side there is a velocity, but if we wish to compare this equation with the classical equation for velocities we cannot simply let the operators go over into their classical analogues, because classical observables commute and we would have zero on the right-hand side. Hence, we must let $\hbar \to 0$ at the same time. Thus, we formulate a heuristic version of the *correspondence principle* as follows:

If a quantum system has a classical analogue, expectation values of operators behave, in the limit $\hbar \to 0$, like the corresponding classical quantities.

This principle provides us with a test that the quantum theory of a system with a classical analogue must meet, but it does not give us an unambiguous prescription of how to construct the quantum form of any given classical theory. Certainly, we cannot expect that every valid classical equation can be turned into a correct quantum equation merely by replacing classical variables by expectation values of operators. For example, $\mathrm{xp_x} = m\mathrm{x}\,(d\mathrm{x}/dt) = (1/2)m(d\mathrm{x}^2/dt)$ is a valid classical equation if not a particularly useful one; yet, for operators $\langle xp\rangle = (1/2)m(d/dt)\langle x^2\rangle$ is generally wrong, although

$$\langle x\rangle\langle p_x\rangle = m\langle x\rangle \frac{d\langle x\rangle}{dt} \quad \text{and} \quad \left\langle \frac{1}{2}(xp_x + p_x x)\right\rangle = \frac{1}{2}\, m\, \frac{d}{dt}\langle x^2\rangle$$

are both correct. The trouble comes from the noncommunitivity of x and p_x.

To make the conversion from classical to quantum mechanics, the correspondence principle must be supplemented by a set of *quantization rules*. These rules have to be consistent with the correspondence principle, but their ultimate test lies in a comparison between the theoretical predictions and the experimental data.

We expect that (14.52) is the quantum analogue of one of Hamilton's equations,

$$\frac{d\mathrm{x}}{dt} = \frac{\partial \mathrm{H}}{\partial \mathrm{p_x}} \tag{14.53}$$

where H is the classical Hamiltonian function of x, y, z, $\mathrm{p_x}$, $\mathrm{p_y}$, $\mathrm{p_z}$, which characterize the system. The correspondence principle requires that

$$\lim_{\hbar\to 0} \frac{\langle xH - Hx\rangle}{i\hbar} = \frac{\partial \mathrm{H}}{\partial \mathrm{p_x}} \tag{14.54}$$

Similarly, for $A = p_x$, we have in quantum mechanics

$$\frac{d\langle p_x\rangle}{dt} = \frac{\langle p_x H - Hp_x\rangle}{i\hbar} \tag{14.55}$$

and classically,

$$\frac{d\mathrm{p_x}}{dt} = -\frac{\partial \mathrm{H}}{\partial \mathrm{x}} \tag{14.56}$$

The correspondence principle requires that

$$\lim_{\hbar \to 0} \frac{\langle p_x H - H p_x \rangle}{i\hbar} = -\frac{\partial \mathrm{H}}{\partial \mathrm{x}} \tag{14.57}$$

Similar equations follow for y and z and their conjugate momenta.

All these conditions can be satisfied if we do the following:

1. *Let H be a Hermitian operator identical in form with* H *but replace all coordinates and momenta by their corresponding operators.*

2. *Postulate the fundamental commutation relations between the Hermitian operators representing coordinates and momenta:*

$$[x, p_x] = [y, p_y] = [z, p_z] = i\hbar 1 \tag{14.58}$$

$$[x, p_y] = [x, p_z] = [y, p_z] = [y, p_x] = [z, p_x] = [z, p_y] = 0 \tag{14.59}$$

The coordinates, x, y, z, commute with each other; the three momenta p_x, p_y, p_z, also commute.

Prescription (1) must be applied with care if H contains terms such as $\mathrm{xp_x}$, because x and p_x are noncommuting and would upon translation into operator language give rise to a non-Hermitian H. The symmetrized operator $(xp_x + p_x x)/2$ can then be used instead. It is Hermitian and leads to the correct classical limit. Sometimes there may be several different ways of symmetrizing a term to make it Hermitian. Thus the Hermitian operators $x^2p_x^2 + p_x^2x^2$ and $(xp_x + p_x x)^2/2$ both have the classical limit $2\mathrm{x^2p_x^2}$, but they are not identical. In practice, it is usually possible to avoid such ambiguities.

Exercise 14.5. Show that the operators $x^2p_x^2 + p_x^2x^2$ and $(xp_x + p_x x)^2/2$ differ only by terms of order $\hbar^2$.

The consistency of conditions (1) and (2) and their agreement with (14.54) and (14.57) can be verified for any H that can be expanded in powers of the coordinates and momenta. For instance, the commutation relation $[x, p_x] = i\hbar 1$ agrees with (14.54) and (14.57), as can be seen if we choose $H = p_x$ and $H = x$, respectively. The consistency proof can be continued by letting $H = x^n$. Then

$$\frac{1}{i\hbar}[p_x, x^n] = \frac{x^{n-1}}{i\hbar}[p_x, x] + \frac{1}{i\hbar}[p_x, x^{n-1}]x = -nx^{n-1}$$

by virtue of repeated invocation of the quantum conditions (14.58). This is in agreement with the classical limit (14.57) because for $\mathrm{H} = \mathrm{x}^n$,

$$-\frac{\partial \mathrm{x}^n}{\partial \mathrm{x}} = -n\mathrm{x}^{n-1}$$

More generally, we can continue this type of reasoning to prove that for any two functions, F and G, of the coordinates and momenta, which can be expanded in a power series, the relation

$$\lim_{\hbar \to 0} \frac{\langle GF - FG \rangle}{i\hbar} = \frac{\partial \mathrm{G}}{\partial \mathrm{x}}\frac{\partial \mathrm{F}}{\partial \mathrm{p_x}} - \frac{\partial \mathrm{F}}{\partial \mathrm{x}}\frac{\partial \mathrm{G}}{\partial \mathrm{p_x}} + \frac{\partial \mathrm{G}}{\partial \mathrm{y}}\frac{\partial \mathrm{F}}{\partial \mathrm{p_y}} - \frac{\partial \mathrm{F}}{\partial \mathrm{y}}\frac{\partial \mathrm{G}}{\partial \mathrm{p_y}} + \frac{\partial \mathrm{G}}{\partial \mathrm{z}}\frac{\partial \mathrm{F}}{\partial \mathrm{p_z}} - \frac{\partial \mathrm{F}}{\partial \mathrm{z}}\frac{\partial \mathrm{G}}{\partial \mathrm{p_z}} \tag{14.60}$$

holds, where F and G are the same functions of the ordinary variables as F and G are of the corresponding operators. Equation (14.60) is assumed to be valid for any smooth functions of coordinates and momenta, even if a power series expansion cannot be made. Equations (14.54) and (14.57) are special cases of (14.60).

In classical mechanics, the expression on the right-hand side of (14.60) is abbreviated as $[G, F]_{P.B.}$ and is known as the *Poisson bracket*[1] (*P.B.*) of F and G. Dirac discovered that this is the classical analogue of the commutator $[G, F]/i\hbar$.

Exercise 14.6. Illustrate the validity of Eq. (14.60) by letting $G = x^2$ and $F = p_x^2$, and evaluating both the operator expression on the left, in the limit $\hbar \to 0$, and the corresponding Poisson bracket on the right.

All of these arguments can be summarized in the proposition that the classical equation of motion for an arbitrary function $A(x, p_x, t)$,

$$\frac{dA}{dt} = [A, H]_{P.B.} + \frac{\partial A}{\partial t} \tag{14.61}$$

is the correspondence limit of the quantum equation of motion (14.18),

$$\frac{d\langle A\rangle}{dt} = \left\langle \frac{[A, H]}{i\hbar} \right\rangle + \left\langle \frac{\partial A}{\partial t} \right\rangle \tag{14.62}$$

for the operator $A(x, p_x, t)$, which is the same function of its arguments as $A(x, p_x, t)$.

4. ***Canonical Quantization and Constants of the Motion.*** So far we have considered only descriptions of the physical system in terms of Cartesian coordinates for a point particle. Yet, the connection between classical and quantum mechanics was established by the use of Hamilton's equations of classical mechanics, which are by no means restricted to Cartesian coordinates. Rather, these equations are well known to have the same general form for a large class of variables called *canonical coordinates and momenta*, and denoted by the symbols q and p. Since in the Hamiltonian form of mechanics the Cartesian coordinates do not occupy a unique position, we ask whether the quantization procedure of Section 14.3 could not equally well have been applied to more general canonical variables. Could we replace x by q and p_x by p (assuming for convenience only one degree of freedom), satisfying the more general commutation relations

$$\boxed{qp - pq = i\hbar 1} \tag{14.63}$$

instead of (14.58), and could we still apply the same quantization rules?

To show that we are indeed at liberty to use canonical variables other than the Cartesian ones, we must prove that the same form of quantum mechanics results whether we use x, p_x, or q, p to make the transition to quantum mechanics. To prove that we can pass from the upper left to the lower right corner of Figure 14.1 equivalently by routes 1, 2 or 3, 4 we first consider an infinitesimal canonical transfor-

[1]Goldstein (1980), p. 397.

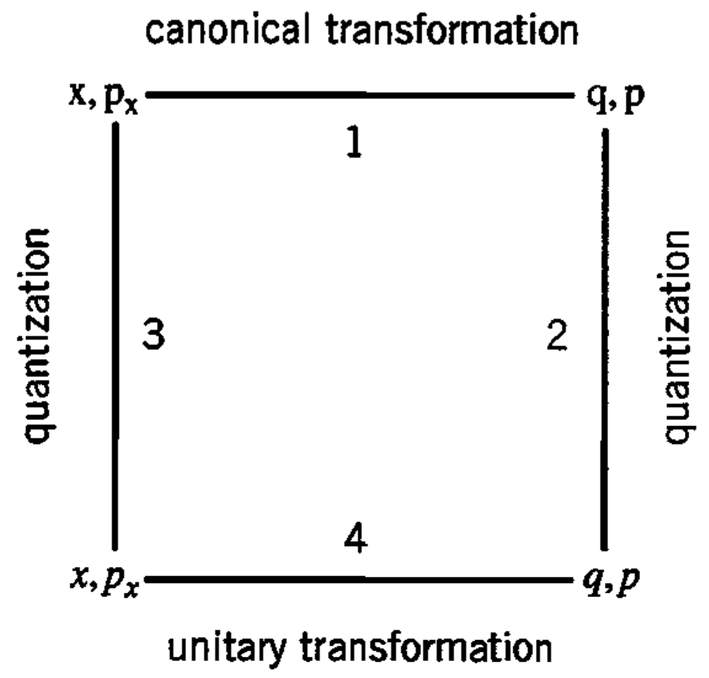


Figure 14.1. Classical canonical transformations and quantum mechanical unitary transformations.

mation (step 1 in the figure), i.e., a transformation that is generated by an infinitesimal function $\varepsilon G(x, p_x)$ from the relations

$$q = x + \varepsilon \frac{\partial G}{\partial p_x}, \qquad p = p_x - \varepsilon \frac{\partial G}{\partial x} \tag{14.64}$$

The new Hamiltonian is

$$\begin{aligned} H'(q, p) = H(x, p_x) = H(q, p) - \varepsilon \frac{\partial H}{\partial x}\frac{\partial G}{\partial p_x} \\ + \varepsilon \frac{\partial H}{\partial p_x}\frac{\partial G}{\partial x} = H(q, p) + \varepsilon[G, H]_{P.B.} \end{aligned} \tag{14.65}$$

This canonical transformation is paralleled in quantum theory by step 4. Agreement with (14.64) and (14.65) in the correspondence limit is assured if we define the Hermitian operators

$$q = x + \frac{\varepsilon}{i\hbar}[x, G], \qquad p = p_x + \frac{\varepsilon}{i\hbar}[p_x, G] \tag{14.66}$$

More generally, for an arbitrary function F(x, p) we find to first order,

$$F(q, p) = F(x, p_x) + \varepsilon \frac{\partial F}{\partial x}\frac{\partial G}{\partial p_x} - \varepsilon \frac{\partial F}{\partial p_x}\frac{\partial G}{\partial x} = F(x, p_x) + \varepsilon[F, G]_{P.B.} \tag{14.67}$$

The corresponding operators satisfy the equation

$$F(q, p) = F(x, p_x) + \frac{\varepsilon}{i\hbar}[F, G] \tag{14.68}$$

again to first order in ε. In quantum mechanics, the new Hamiltonian is constructed by correspondence with (14.65) as

$$H'(q, p) = H(x, p_x) = H(q, p) + \frac{\varepsilon}{i\hbar}[G, H] \tag{14.69}$$

The Hermitian operator $G(x, p_x)$ is constructed from $G(x, p_x)$ by letting x and p_x become operators. The commutators are evaluated by applying the quantization rules of the last section for Cartesian coordinates.

To first order in ε, (14.68) may be written as

$$F(q, p) = \left(1 + \frac{i\varepsilon}{\hbar} G\right) F(x, p_x) \left(1 - \frac{i\varepsilon}{\hbar} G\right) \tag{14.70}$$

showing that the new operators are obtained from the old Cartesian ones by an infinitesimal *unitary* transformation (step 4):

$$F(q, p) = U_\varepsilon F(x, p_x) U_\varepsilon^\dagger \tag{14.71}$$

with

$$U_\varepsilon = 1 + \frac{i\varepsilon}{\hbar} G \tag{14.72}$$

The Hermitian operator G is the *generator* of this infinitesimal transformation.

In terms of the new variables, the quantum analogue of the classical Hamiltonian, (14.69), becomes

$$H'(q, p) = H(x, p_x) = H(U_\varepsilon^\dagger q U_\varepsilon, U_\varepsilon^\dagger q U_\varepsilon) = U_\varepsilon^\dagger H(q, p) U_\varepsilon \tag{14.73}$$

[See Exercise 9.16 for the conditions under which the last equality in (14.73) holds.]

The commutation relations are invariant under unitary transformations because

$$[q, p] = U_\varepsilon x U_\varepsilon^\dagger U_\varepsilon p_x H_\varepsilon^\dagger - U_\varepsilon p_x U_\varepsilon^\dagger U_\varepsilon x U_\varepsilon^\dagger = U_\varepsilon [x, p_x] U_\varepsilon^\dagger = i\hbar 1 \tag{14.74}$$

and we have arrived at (14.63). This completes the proof that the quantization rules of 14.3 can be extended to new canonical variables that are infinitesimally close to Cartesian.

The quantization procedure based on rules (1) and (2) of Section 14.3 can now be immediately generalized to all those canonical variables that can be obtained from the Cartesian ones by a succession of infinitesimal canonical transformations. This is true because two classical canonical transformations made in succession can be replaced by a single direct one. Similarly, in quantum mechanical transformations, successive application of unitary operators is equivalent to the application of a single unitary operator. If we let $\varepsilon = \lambda/N$ (where λ is a finite parameter and N is an integer), and apply the same unitary operator N times, we obtain the in limit $N \to \infty$ the unitary operator,

$$U = \lim_{N \to \infty} \left(1 + \frac{i\lambda}{N\hbar} G\right)^N = \exp\left(\frac{i\lambda}{\hbar} G\right) \tag{14.75}$$

This finite unitary transformation changes the Cartesian variables into

$$q = UxU^\dagger, \qquad p = Up_xU^\dagger \tag{14.76}$$

The commutation relations are also invariant under the finite transformations.

We note that if (14.76) holds, the eigenvalue spectra of x and q are the same, as

$$UxU^\dagger U|x'\rangle = qU|x'\rangle = x'U|x'\rangle$$

Hence,

$$q' = x' \quad \text{and} \quad |q'\rangle = U|x'\rangle$$

We see that the quantization of the system can be carried through by the use of the general commutation relations (14.63) for any pair of canonical variables that

can be obtained from x, p_x by a continuous succession of infinitesimal transformations. For more general canonical transformations than these, the standard quantization procedure may or may not be valid. Clearly, it will be valid whenever the new operators can be obtained from the old ones by a unitary transformation.

A simple example of a failure of the standard quantization is provided by the transition to spherical polar coordinates, r, φ, θ. The transformation to these from Cartesian coordinates is canonical, but it cannot be generated by a succession of infinitesimal transformations, because of the discontinuity of the spherical polar coordinates. Nor does a unitary transformation between x, y, z and r, φ, θ exist, for if it did the eigenvalues of the latter operators would have to be the same as those of x, y, z, and range from $-\infty$ to $+\infty$, contrary to the definition of r, φ, θ. The general procedure for expressing the Hamiltonian operator and the Schrödinger equation in terms of curvilinear coordinates will be given in the Appendix, Section 3.

Because of its close connection with the classical canonical formalism, the quantization procedure described here is referred to as *canonical quantization*. The correspondence between canonical transformations and unitary operators has led to the frequent designation of unitary operators as quantum mechanical *canonical transformations*. This terminology has asserted itself, even though some unitary transformations have no classical analogue, and vice versa.

Exercise 14.7. Show that the transformation

$$\begin{aligned} UaxU^\dagger &= ax\cos\Theta + bp_x\sin\Theta \\ Ubp_xU^\dagger &= -ax\sin\Theta + bp_x\cos\Theta \end{aligned} \tag{14.77}$$

is canonical, if a and b are real-valued constants, and Θ is a real-valued angle parameter. Construct the unitary operator U that effects this transformation. For the special case $\Theta = \pi/2$, calculate the matrix elements of U in the coordinate representation. Noting that this transformation leaves the operator $a^2x^2 + b^2p_x^2$ invariant, rederive the result of Exercises 3.8 and 3.21.

Exercise 14.8. Show that the reflection operator, defined by the relation $U|x'\rangle = |-x'\rangle$, gives rise to a unitary transformation which takes x into $-x$ and p_x into $-p_x$.

An important application of canonical transformations concerns the finding of constants of the motion, which are observables that commute with the Hamiltonian H (see Section 14.1).

A useful way of obtaining constants of the motion for a time-independent Hamiltonian operator $H(q, p)$ consists in noting that if a (finite or infinitesimal) canonical transformation to new variables q', p' is made, the new Hamiltonian H' is related to the old one by the equation

$$H(q, p) = H'(q', p') = UH'(q, p)U^\dagger \tag{14.78}$$

which is just an extension of (14.73) to finite canonical transformations. If the canonical transformation leaves the Hamiltonian *invariant*, so that the new Hamiltonian H' is the *same* function of the canonical variables as the old one,

$$H'(q, p) = H(q, p) \tag{14.79}$$

then

$$UH - HU = [U, H] = 0$$

Hence,

$$\frac{d}{dt}\langle U\rangle = \frac{1}{i\hbar}\langle UH - HU\rangle = 0$$

and thus the unitary operator U is a constant of the motion *if the transformation leaves H invariant*. If, in particular, H is invariant under an infinitesimal transformation

$$U_\varepsilon = 1 + \frac{i\varepsilon}{\hbar} G$$

then the (Hermitian) generator G of the transformation commutes with H,

$$[G, H] = 0 \tag{14.80}$$

and thus G is a constant of the motion. In this way, physical observables are obtained which are constants of the motion.

As an example, consider a free particle whose Hamiltonian is

$$H = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) \tag{14.81}$$

According to (14.66), the generator of infinitesimal translations ε,

$$\varepsilon G = -\varepsilon_x p_x - \varepsilon_y p_y - \varepsilon_z p_z = -\boldsymbol{\varepsilon}\cdot\mathbf{p}$$

produces no change in the momenta

$$\mathbf{p}' = \mathbf{p} \tag{14.82}$$

but, owing to the fundamental commutation relations, changes the coordinates to

$$\mathbf{r}' = \mathbf{r} - \frac{1}{i\hbar}[\mathbf{r}, \boldsymbol{\varepsilon}\cdot\mathbf{p}] = \mathbf{r} - \boldsymbol{\varepsilon} \tag{14.83}$$

Thus, the transformation describes a coordinate translation, as expected from the connection between momentum and displacement operators.

Evidently, any H that does not depend on the coordinates is invariant under the transformation (14.82), (14.83). Hence, $\varepsilon G = -\boldsymbol{\varepsilon}\cdot\mathbf{p}$ (for arbitrary $\boldsymbol{\varepsilon}$) and $\mathbf{p}$ itself are constants of the motion, and we conclude that *linear momentum is conserved* under these conditions.

Similarly, for infinitesimal rotations we recall the generator

$$\varepsilon G = -\delta\phi\,\hat{\mathbf{n}}\cdot\mathbf{L} \tag{14.84}$$

from Section 11.1. For a vector operator $\mathbf{A}$, such a rotation induces the change

$$\delta\mathbf{A} = \frac{\varepsilon}{i\hbar}[\mathbf{A}, G] = -\frac{\delta\phi}{i\hbar}[\mathbf{A}, \hat{\mathbf{n}}\cdot\mathbf{L}] = -\delta\phi\,\hat{\mathbf{n}}\times\mathbf{A} \tag{14.85}$$

by (11.19). The operator $\mathbf{A}\cdot\mathbf{A} = \mathbf{A}^2$ is a rotationally invariant scalar operator, as was shown in Section 11.1.

The Hamiltonian for a particle in a central-force field,

$$H = \frac{p^2}{2m} + V(r) \tag{14.86}$$

is invariant under rotations, since p^2 and r are scalar operators. Hence, $\varepsilon G = -\delta\phi\,\hat{\mathbf{n}}\cdot\mathbf{L}$ (for arbitrary vectors $\hat{\mathbf{n}}$ and rotation angles $\delta\phi$) and $\mathbf{L}$ itself are

constants of the motion, and *orbital angular momentum is conserved* for a system with spherical symmetry.

Exercise 14.9. Show that if both A and B are constants of the motion, they either commute or the commutator $i[A, B]$ is also a constant of the motion. Prove that if the entire spectrum of H is nondegenerate, then A and B must commute. If the constants of the motion A and B do not commute, there must be degenerate energy eigenvalues. Illustrate this theorem by constructing an example Hamiltonian for which $A = L_x$ and $B = L_y$ are constants of the motion.[2]

5. ***Canonical Quantization in the Heisenberg Picture.*** The canonical quantization procedure can be formulated in any of the pictures of quantum dynamics, since they are all related to the Schrödinger picture by a generally time-dependent, unitary transformation $U(t)$. Such a transformation leaves every algebraic relation between operators

$$f(A, B, C, \ldots) = 0$$

formally unchanged (see Exercise 9.16):

$$f(\tilde{A}, \tilde{B}, \tilde{C}, \ldots) = f(UAU^\dagger, UBU^\dagger, UCU^\dagger, \ldots) = Uf(A, B, C, \ldots)U^\dagger = 0$$

Hence, the canonical commutation relations for conjugate variables

$$[q, p] = qp - pq = i\hbar 1$$

become the same in any picture:

$$[\tilde{q}(t), \tilde{p}(t)] = (UqU^\dagger)(UpU^\dagger) - (UpU^\dagger)(UqU^\dagger) = i\hbar 1$$

The dynamical law,

$$i\hbar \frac{d\langle \tilde{A} \rangle}{dt} = \langle [\tilde{A}, \tilde{H}] \rangle \tag{14.87}$$

for any operator A that does not depend on time explicitly also has the same form in all pictures.

Exercise 14.10. Using Eqs. (14.41) and (14.44), verify that the equation of motion (14.87) for expectation values of observables holds in any picture of quantum dynamics.

The equations of motion for the canonical variables are derived from (14.41):

$$\begin{aligned} i\hbar \frac{d\tilde{q}(t)}{dt} &= [\tilde{q}(t), H_0] \\ i\hbar \frac{d\tilde{p}(t)}{dt} &= [\tilde{p}(t), H_0] \end{aligned} \tag{14.88}$$

and depend on the choice of H_0, the model Hamiltonian for the particular dynamical picture.

[2]See Fallieros (1995).

In the Heisenberg picture ($H_0 = H$), the equation of motion can be expressed for any dynamical variable

$$\overline{F}(\overline{q}, \overline{p}, t) \equiv F(\overline{q}(t), \overline{p}(t), t)$$

according to (14.30) in summary form as

$$\boxed{\frac{d}{dt}\overline{F}(\overline{q}, \overline{p}, t) = \frac{1}{i\hbar}[\overline{F}(\overline{q}, \overline{p}, t), H] + \frac{\partial}{\partial t}\overline{F}(\overline{q}, \overline{p}, t)} \tag{14.89}$$

By (14.61), this is the quantum analogue of the classical equation of motion,

$$\frac{d}{dt}\mathrm{F(q, p, }t) = [\mathrm{F(q, p, }t), \mathrm{H}]_{P.B.} + \frac{\partial}{\partial t}\mathrm{F(q, p, }t) \tag{14.90}$$

The formal similarity between the classical equation of motion for canonical variables and the quantum equations for the corresponding *operators*, and not merely their expectation values, confers a measure of distinction on the Heisenberg picture of quantum dynamics, which otherwise is just one of infinitely many (unitarily) equivalent pictures. In the Heisenberg picture, the transition from classical to quantum theory for a system that has a classical analogue is made simply by replacing the canonical variables by operators that can change in time, subject to the commutation relations

$$[\overline{q}_k(t), \overline{p}_\ell(t)] = i\hbar\delta_{k\ell} \tag{14.91}$$

$$[\overline{q}_k(t), \overline{q}_\ell(t)] = [\overline{p}_k(t), \overline{p}_\ell(t)] = 0 \tag{14.92}$$

and by postulating that if the classical equations of motion are expressed in terms of Poisson brackets, the correspondence

$$\boxed{[\mathrm{F, G}]_{P.B.} \to \frac{i}{\hbar}[\overline{F}, \overline{G}]} \tag{14.93}$$

is to be made—barring, as usual, complications that may arise from ambiguities in the ordering of operators.

The simple commutation relations (14.91) and (14.92) are not valid if the operators are taken at two different times. Thus, generally, $\overline{q}(t)$ and $\overline{q}(0)$ do not commute, nor is the commutator $[\overline{q}(0), \overline{p}(t)]$ equal to $i\hbar 1$. For example, if the system is a free particle in one dimension with $H = p^2/2m$, we have

$$\overline{p}(t) = \overline{p}(0) \quad \text{and} \quad \overline{q}(t) = \overline{q}(0) + \frac{\overline{p}(0)}{m}t \tag{14.94}$$

hence,

$$[\overline{q}(t), \overline{q}(0)] = -\frac{i\hbar}{m}t \tag{14.95}$$

If this commutation relation is applied to (10.54), we find the uncertainty relation

$$(\Delta q)_t(\Delta q)_0 = \Delta\overline{q}(t)\Delta\overline{q}(0) \geq \frac{\hbar}{2m}|t| \tag{14.96}$$

which shows that if the particle moves freely, its wave packet in the coordinate representation must spread in the long run as $|t| \to \infty$. In Eq. (14.96), the notation

on the left-hand side emphasizes that the variances, being expectation values, are independent of the choice of the quantum dynamical picture. Except for denoting the coordinate by q rather than x, this inequality provides a precise formulation of the statements about wave packets contracting long before the present time and spreading in the distant future, that were made in Section 2.4. If the particle is free, initially narrow wave packets spread more rapidly than those that initially are broad.

If the initial wave packet at $t = 0$ is the minimum uncertainty wave packet (10.66), with

$$\Delta\bar{q}(0)\Delta\bar{p} = \frac{\hbar}{2}$$

then, using (14.96),

$$\Delta\bar{q}(t) \geq \frac{|t|\Delta\bar{p}}{m}$$

This inequality is consistent with the result of an explicit calculation of the time dependence of the variance of q for a wave packet that has ''minimum uncertainty'' at $t = 0$:

$$(\Delta q)_t^2 = (\Delta q)_0^2 + \frac{(\Delta p)^2 t^2}{m^2} \tag{14.97}$$

Exercise 14.11. For a free particle in one dimension and an arbitrary initial wave packet, calculate the time development of $(\Delta q)_t^2$ and show (as in Problem 2 in Chapter 3) that

$$(\Delta q)_t^2 = (\Delta q)_0^2 + \frac{t}{m}[\langle qp + pq\rangle_0 - 2\langle q\rangle_0\langle p\rangle] + \frac{(\Delta p)^2 t^2}{m^2} \tag{14.98}$$

Verify that for the minimum uncertainty wave packet this result agrees with (14.97). Also compare with the value of the variance $(\Delta q)^2$ as a function of time for a beam of freely moving classical particles whose initial positions and momenta have distributions with variances $(\Delta q)_0^2$ and $(\Delta p)^2$.

If the close correspondence between the classical theory and quantum dynamics gives the Heisenberg picture a certain preferred status, the Schrödinger picture is perhaps a more intuitive form of quantum mechanics. The Schrödinger picture is particularly suitable for a discussion of scattering processes, which are more naturally described by moving wave packets, albeit complex-valued ones, than by operators changing in time. From our present general point of view, however, it is a matter of taste and convenience whether we navigate in the Heisenberg or the Schrödinger picture, or any other picture. Once again, the harmonic oscillator offers the simplest nontrivial illustration.

In the Heisenberg picture, the oscillator Hamiltonian may, similar to the Schrödinger picture form, be written as

$$H = \hbar\omega\left(\bar{a}^\dagger(t)\bar{a}(t) + \frac{1}{2}\right) \tag{14.99}$$

Applying (14.30), we obtain the equation of motion for the lowering (annihilation) operator $\bar{a}(t)$,

$$i\hbar \frac{d\bar{a}(t)}{dt} = [\bar{a}(t), H] = \hbar\omega[\bar{a}(t), \bar{a}(t)^\dagger]\bar{a}(t)$$

Using the commutation relation (10.72), transcribed into the Heisenberg picture, we find the simple differential equation

$$\frac{d\bar{a}(t)}{dt} + i\omega\, \bar{a}(t) = 0 \tag{14.100}$$

Although the operator equations of motion are difficult to solve directly in most problems, necessitating passage to a representation in which these equations become conventional systems of linear differential and integral equations, the present example is an important exception. Equation (14.100) can be solved immediately, even though $\bar{a}(t)$ is an operator:

$$\bar{a}(t) = \bar{a}(t_0)\, e^{-i\omega(t-t_0)} \tag{14.101}$$

Similarly,

$$\bar{a}^+(t) = \bar{a}^\dagger(t_0)\, e^{i\omega(t-t_0)} \tag{14.102}$$

If we choose $t_0 = 0$, the initial values of the operators $\bar{a}(0) = a$ and $\bar{a}^\dagger(0) = a^\dagger$ are two mutually adjoint operators that satisfy the commutation relation (10.72). They are the raising (creation) and lowering (annihilation) operators in the Schrödinger picture, which coalesces with the Heisenberg picture at $t = 0$; they serve as the constants of integration of the dynamical problem. The canonical variables p and q, and any function of these, can be expressed in terms of a and $a^\dagger$. Thus, in principle the equations of motion have all been integrated.

Exercise 14.12. Work out directly from (14.27),

$$\bar{a}(t) = \exp\left(\frac{i}{\hbar} Ht\right) a \exp\left(-\frac{i}{\hbar} Ht\right)$$

as an application of the identity (3.59). Then determine $\bar{q}(t)$ and $\bar{p}(t)$.

We now choose a basis and introduce a fixed representation. The most convenient one is the same as in Section 10.6: A basis supported by the eigenvectors of H. This may be called the *energy representation*. The matrices representing $\bar{a}(t)$ and $\bar{a}^\dagger(t)$ are obtained by multiplying the matrices (10.90a) and (10.90b) representing a and $a^\dagger$ by $e^{-i\omega t}$ and $e^{i\omega t}$, respectively. In the energy representation of the Heisenberg picture, the coordinate operator of the harmonic oscillator is explicitly represented by the matrix:

$$\begin{aligned} \bar{q}(t) &= \sqrt{\frac{\hbar}{2m\omega}}\, [\bar{a}(t) + \bar{a}^\dagger(t)] \\ &= \sqrt{\frac{\hbar}{2m\omega}} \begin{pmatrix} 0 & \sqrt{1}e^{-i\omega t} & 0 & \cdot & \cdot \\ \sqrt{1}e^{i\omega t} & 0 & \sqrt{2}e^{-i\omega t} & \cdot & \cdot \\ 0 & \sqrt{2}e^{i\omega t} & 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & & & \cdot & \cdot \end{pmatrix} \end{aligned} \tag{14.103}$$

This matrix is evidently Hermitian, as it should be. Its elements are harmonic functions of time—a general property of the matrix representing any Heisenberg operator in an energy representation that is supported by the fixed eigenvectors of H:

$$H|E\rangle = E|E\rangle$$

Indeed, for any operator, which is not explicitly time-dependent,

$$\bar{A}(t) = \exp\left(\frac{i}{\hbar} Ht\right) A \exp\left(-\frac{i}{\hbar} Ht\right)$$

giving us the matrix element

$$\langle E'|\bar{A}(t)|E''\rangle = \langle E'|A|E''\rangle \exp\left[\frac{i}{\hbar}(E' - E'')t\right] \tag{14.104}$$

The special feature of the harmonic oscillator as a perfect periodic system is that all matrix elements oscillate with integral multiples of the same frequency, ω.

Exercise 14.13. In either the Heisenberg or Schrödinger picture, show that if at $t = 0$ a linear harmonic oscillator is in a coherent state, with eigenvalue α, it will remain in a coherent state, with eigenvalue $\alpha e^{-i\omega t}$, at time t.

6. ***The Forced Harmonic Oscillator.*** For many applications, especially in many-body and field theory, it is desirable to consider the dynamical effects produced by the addition of a time-dependent interaction that is linear in q to the Hamiltonian of the harmonic oscillator:

$$H = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 q^2 - qQ(t)$$

where $Q(t)$ is a real-valued function of t. This perturbation corresponds to an external time-dependent force that does not depend on the coordinate q (*dipole interaction*). With no additional effort, we may generalize the Hamiltonian even further by introducing a velocity-dependent term:

$$H = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 q^2 - qQ(t) - pP(t) \tag{14.105}$$

where $P(t)$ is also a real function of t.

With the substitutions (10.69) and (10.71), the Hamiltonian (14.105) may be cast in the form

$$\begin{aligned} H &= \hbar\omega\left(a^\dagger a + \frac{1}{2}\right) + f(t)a + f^*(t)a^\dagger \\ &= \hbar\omega\left(\bar{a}^\dagger(t)\bar{a}(t) + \frac{1}{2}\right) + f(t)\bar{a}(t) + f^*(t)\bar{a}^\dagger(t) \end{aligned} \tag{14.106}$$

in either the Schrödinger or Heisenberg picture, provided that we define the complex-valued function $f(t)$:

$$f(t) = -\sqrt{\frac{\hbar}{2m\omega}}\, Q(t) + i\sqrt{\frac{\hbar m\omega}{2}}\, P(t) \tag{14.107}$$

In most applications, we are interested in the lasting rather than the transient changes produced by the time-dependent forces in an initially unperturbed harmonic oscillator. It is therefore reasonable to assume that the disturbance $f(t) \neq 0$ acts only during the finite time interval $T_1 < t < T_2$ and that before T_1 and after T_2 the Hamiltonian is that of a free oscillator. The time development of the system is conveniently studied in the Heisenberg picture, in which the state vector $|\Psi\rangle$ is constant, and the operators are subject to a unitary transformation as they change from the free oscillation regime before T_1 to a free oscillation regime after T_2.

Using the equal-time commutation relation,

$$[\bar{a}(t), \bar{a}^\dagger(t)] = 1 \tag{14.108}$$

we derive the equation of motion

$$i\hbar \frac{d\bar{a}(t)}{dt} = [\bar{a}(t), H] = \hbar\omega\, \bar{a}(t) + f^*(t)$$

or

$$\frac{d\bar{a}(t)}{dt} + i\omega\, \bar{a}(t) = -\frac{i}{\hbar} f^*(t) \tag{14.109}$$

This inhomogeneous differential equation is easily solved by standard methods. For instance, it can be multiplied by $e^{i\omega t}$ and cast in the form

$$\frac{d}{dt}\{\bar{a}(t)e^{i\omega t}\} = -\frac{i}{\hbar} f^*(t)e^{i\omega t}$$

which can then be integrated to produce the general solution

$$\bar{a}(t) = \bar{a}(t_0)e^{-i\omega(t-t_0)} - \frac{i}{\hbar}\int_{t_0}^{t} e^{-i\omega(t-t')}f^*(t')\, dt' \tag{14.110}$$

If we choose $t_0 = 0$, this equation simplifies to

$$\bar{a}(t) = ae^{-i\omega t} - \frac{i}{\hbar}\int_{0}^{t} e^{-i\omega(t-t')}f^*(t')\, dt' \tag{14.111}$$

Although it calls for unnecessarily heavy artillery, it is instructive to review the solution of (14.109) by the use of Green's functions to illustrate a method that has proved useful in many similar but more difficult problems.

A Green's function appropriate to Eq. (14.109) is a solution of the equation

$$\frac{dG(t-t')}{dt} + i\omega G(t-t') = \delta(t-t') \tag{14.112}$$

because such a function permits us to write a particular solution of Eq. (14.109) as

$$\bar{a}(t) = -\frac{i}{\hbar}\int_{-\infty}^{+\infty} G(t-t')f^*(t')\, dt' \tag{14.113}$$

This is easily verified by substituting (14.113) into (14.109). Obviously, for $t \neq t'$ the Green's function is proportional to $e^{-i\omega(t-t')}$, but at $t = t'$ there is a discontinuity. By integrating (14.112) over an infinitesimal interval from $t' - \varepsilon$ to $t' + \varepsilon$, we derive the condition

$$\lim_{\varepsilon\to 0}[G(+\varepsilon) - G(-\varepsilon)] = 1 \tag{14.114}$$

for $\varepsilon > 0$.

Two particular Green's functions are useful:

$$\boxed{G_R(t - t') = \eta(t - t')e^{-i\omega(t-t')}} \tag{14.115}$$

and

$$\boxed{G_A(t - t') = -\eta(t' - t)e^{-i\omega(t-t')}} \tag{14.116}$$

where the Heaviside step function $\eta(t) = 0$ for $t < 0$ and $\eta(t) = 1$ for $t > 0$ (see Appendix, Section 1). These two particular solutions of (14.112) are known as *retarded* and *advanced* Green's functions, respectively.

We note that

$$\bar{a}(t) = -\frac{i}{\hbar}\int_{-\infty}^{+\infty} G_R(t - t')f^*(t')\,dt'$$

is the particular solution of (14.109), which vanishes for $t < T_1$. Similarly,

$$\bar{a}(t) = -\frac{i}{\hbar}\int_{-\infty}^{+\infty} G_A(t - t')f^*(t')\,dt'$$

is the particular solution of (14.109) which vanishes for $t > T_2$.

If we denote by $\bar{a}_{\text{in}}(t)$ and $\bar{a}_{\text{out}}(t)$ those solutions of the *homogeneous* equation

$$\frac{d\bar{a}(t)}{dt} + i\omega\bar{a}(t) = 0$$

which coincide with the solution $\bar{a}(t)$ of the inhomogeneous equation (14.109) for $t < T_1$ and $t > T_2$ respectively, and if we choose $t_0 = 0$, we can write

$$\begin{aligned}\bar{a}(t) &= \bar{a}_{\text{in}}(t) - \frac{i}{\hbar}\int_{-\infty}^{+\infty} G_R(t - t')f^*(t')\,dt' \\ &= a_{\text{in}}e^{-i\omega t} - \frac{i}{\hbar}\int_{-\infty}^{t} e^{-i\omega(t-t')}f^*(t')\,dt'\end{aligned} \tag{14.117}$$

or, alternatively,

$$\begin{aligned}\bar{a}(t) &= \bar{a}_{\text{out}}(t) - \frac{i}{\hbar}\int_{-\infty}^{+\infty} G_A(t - t')f^*(t')\,dt' \\ &= a_{\text{out}}e^{-i\omega t} + \frac{i}{\hbar}\int_{t}^{+\infty} e^{-i\omega(t-t')}f^*(t')\,dt'\end{aligned} \tag{14.118}$$

Both (14.117) and (14.118) are equivalent to solution (14.111). By equating the right-hand sides of (14.117) and (14.118), we obtain the relation

$$\boxed{a_{\text{out}} = a_{\text{in}} - \frac{i}{\hbar}g^*(\omega)} \tag{14.119}$$

where

$$g(\omega) = \int_{T_1}^{T_2} e^{-i\omega t'}f(t')\,dt' = \int_{-\infty}^{+\infty} e^{-i\omega t'}f(t')\,dt' \tag{14.120}$$

is the Fourier transform of the generalized force $f(t)$.

The solution (14.117) or (14.118) in the Heisenberg picture can be used to answer all questions about the time development of the system. It is nevertheless instructive also to consider the dynamical problem from the point of view of other pictures.

To implement the interaction picture, we regard the Hamiltonian of the forced oscillator as the sum, $H = H_0 + V(t)$, of an unperturbed Hamiltonian

$$H_0 = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right) \tag{14.121}$$

and an explicitly time-dependent "interaction" term,

$$V(t) = f(t)a + f^*(t)a^\dagger \tag{14.122}$$

Time-dependent Hamiltonians require more careful treatment than time-independent ones, because generally the interaction operators at two different times do not commute.

We choose the unperturbed Hamiltonian operator H_0 as the model Hamiltonian to define the interaction picture. According to (14.45), the transformed interaction operator is

$$\tilde{V}(t) = e^{i\omega a^\dagger at}[f(t)a + f^*(t)a^\dagger]e^{-i\omega a^\dagger at} \tag{14.123}$$

The interaction operator can be evaluated by use of the identity (3.59), since $[a^\dagger a, a] = -a$ and $[a^\dagger a, a^\dagger] = a^\dagger$. We thus obtain

$$\tilde{V}(t) = f(t)ae^{-i\omega t} + f^*(t)a^\dagger e^{i\omega t} \tag{14.124}$$

The equation of motion for the state vector in the interaction picture is

$$i\hbar\frac{d}{dt}|\tilde{\Psi}(t)\rangle = \{f(t)ae^{-i\omega t} + f^*(t)a^\dagger e^{i\omega t}\}|\tilde{\Psi}(t)\rangle \tag{14.125}$$

Similarly, the time development operator $\tilde{T}(t_2, t_1)$ in the interaction picture as defined in Eq. (14,49) satisfies the equation of motion:

$$i\hbar\frac{d\tilde{T}(t, t_1)}{dt} = \tilde{V}(t)\tilde{T}(t, t_1) \tag{14.126}$$

Integration of (14.126) over the time interval (t_1, t) and use of the initial condition $\tilde{T}(t_1, t_1) = 1$ produce an integral equation for the time development operator:

$$\tilde{T}(t, t_1) = 1 - \frac{i}{\hbar}\int_{t_1}^{t}\tilde{V}(t')\tilde{T}(t', t_1)\,dt' \tag{14.127}$$

A formal solution of this equation can be constructed by successive iteration:

$$\tilde{T}(t, t_1) = 1 - \frac{i}{\hbar}\int_{t_1}^{t}\tilde{V}(t')\,dt' + \left(\frac{i}{\hbar}\right)^2\int_{t_1}^{t}\tilde{V}(t')\,dt'\int_{t_1}^{t'}\tilde{V}(t'')\,dt'' \pm \ldots \tag{14.128}$$

It is sometimes convenient to write this series expansion in a more symmetric form by using the *time-ordered product* of operators. We define time ordering of two operators as

$$\mathrm{T}[\tilde{V}(t')\tilde{V}(t'')] = \begin{cases}\tilde{V}(t')\tilde{V}(t'') & \text{if } t'' \leq t' \\ \tilde{V}(t'')\tilde{V}(t') & \text{if } t' \leq t''\end{cases} \tag{14.129}$$

This convention is easily generalized to products of any number of time-dependent operators. With it we can prove that if $t > t_1$, the time development operator may be written in the form

$$\tilde{T}(t, t_1) = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar}\right)^n \int_{t_1}^{t} \int_{t_1}^{t} \cdots \int_{t_1}^{t} dt_1' dt_2' \cdots dt_n' \mathrm{T}[\tilde{V}(t_1')\tilde{V}(t_2') \cdots \tilde{V}(t_n')] \tag{14.130}$$

or formally and compactly as

$$\boxed{\tilde{T}(t, t_1) = \mathrm{T} \exp\left[-\frac{i}{\hbar} \int_{t_1}^{t} \tilde{V}(t')\, dt'\right]} \tag{14.131}$$

To prove that (14.130) is a solution of (14.126), it is sufficient to write to first order in ε,

$$\tilde{T}(t + \varepsilon, t_1) = 1 + \sum_{n=1}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar}\right)^n \int_{t_1}^{t+\varepsilon} \int_{t_1}^{t+\varepsilon} \cdots \int_{t_1}^{t+\varepsilon} dt_1' dt_2' \cdots dt_n' \mathrm{T}[\tilde{V}(t_1')\tilde{V}(t_2') \cdots \tilde{V}(t_n')]$$

$$= \tilde{T}(t, t_1) + \varepsilon \sum_{n=1}^{\infty} \frac{1}{n!} \left(-\frac{i}{\hbar}\right)^n n\tilde{V}(t) \int_{t_1}^{t} \int_{t_1}^{t} \cdots \int_{t_1}^{t} dt_1' dt_2' \cdots dt_{n-1}' \mathrm{T}[\tilde{V}(t_1')\tilde{V}(t_2') \cdots \tilde{V}(t_{n-1}')]$$

or

$$\frac{\tilde{T}(t + \varepsilon, t_1) - \tilde{T}(t, t_1)}{\varepsilon} = -\frac{i}{\hbar} \tilde{V}(t)\tilde{T}(t, t_1)$$

which in the limit $\varepsilon \to 0$ reduces to (14.126).

If the formula (14.131) is applied to the forced linear harmonic oscillator, with the interaction potential (14.124), we obtain

$$\tilde{T}(t, t_1) = \mathrm{T} \exp\left\{-\frac{i}{\hbar} \int_{t_1}^{t} [f(t')ae^{-i\omega t'} + f^*(t')a^\dagger e^{i\omega t'}]\, dt'\right\} \tag{14.132}$$

Although this is a compact expression for the time development operator, because of the presence of the unwieldy time ordering operator it is not yet in a form convenient for calculating transition amplitudes.

To derive a more manageable formula, we use the group property (14.8) and write

$$\tilde{T}(t, t_1) = \lim_{N \to \infty} e^{V_N} e^{V_{N-1}} e^{V_{N-2}} \cdots e^{V_2} e^{V_1} \tag{14.133}$$

where, by definition,

$$V_k = -\frac{i}{\hbar} \int_{t_1 + (k-1)\varepsilon}^{t_1 + k\varepsilon} \tilde{V}(t')\, dt' \quad \text{and} \quad N\varepsilon = t - t_1 \tag{14.134}$$

Expression (14.133) is valid, even if the interaction operators at different times do not commute, because the time intervals of length ε are infinitesimally short and are not subject to internal time ordering. The expression (14.133) can be further reduced if the commutators $[\tilde{V}(t'), \tilde{V}(t'')]$ are numbers for all t' and t''. This is indeed the case for the forced harmonic oscillator, since according to (14.124),

$$[\tilde{V}(t'), \tilde{V}(t'')] = f(t')f^*(t'')e^{-i\omega(t'-t'')} - f^*(t')f(t'')e^{i\omega(t'-t'')} \tag{14.135}$$

The identity (3.61) can be applied repeatedly to give

$$\tilde{T}(t, t_1) = \lim_{N\to\infty} \exp\left\{\sum_{k=1}^{\infty}\left(V_k + \frac{1}{2}\left[V_k, \sum_{n=1}^{k} V_n\right]\right)\right\}$$

or, if the limit $N \to \infty$ and $\varepsilon \to 0$ is carried out,

$$\tilde{T}(t, t_1) = \exp\left\{-\frac{i}{\hbar}\int_{t_1}^{t} \tilde{V}(t')\,dt' - \frac{1}{2\hbar^2}\int_{t_1}^{t} dt' \int_{t_1}^{t} dt''[\tilde{V}(t'), \tilde{V}(t'')]\right\} \quad (14.136)$$

For the forced harmonic oscillator, inserting (14.124) and (14.135) into (14.136), we thus obtain the time development operator in the interaction picture in the desired form:

$$\tilde{T}(t, t_1) = e^{i\beta(t,t_1)} \exp\{-\zeta^*(t, t_1)a + \zeta(t, t_1)a^\dagger\} \quad (14.137)$$

where we have defined

$$\zeta(t, t_1) = -\frac{i}{\hbar}\int_{t_1}^{t} e^{i\omega t'} f^*(t')\,dt' \quad (14.138)$$

This expression can be connected with the Fourier integral of the applied force given in (14.120):

$$g(\omega) = -i\hbar\zeta^*(T_2, T_1) = -i\hbar\zeta^*(-\infty, +\infty) \quad (14.139)$$

In (14.137), the real phase β appears as the price we must pay for eliminating the time ordering operator, and it stands for:

$$\beta(t, t_1) = \frac{i}{2\hbar^2}\int_{t_1}^{t} dt' \int_{t_1}^{t'} dt''\{f(t')f^*(t'')e^{-i\omega(t'-t'')} - f^*(t')f(t'')e^{i\omega(t'-t'')}\} \quad (14.140)$$

If the initial state at $t = t_1$ is a coherent oscillator state $|\alpha\rangle$, as defined in Section 10.7, the state at time t is

$$\begin{aligned} |\tilde{\Psi}(t)\rangle = \tilde{T}(t, t_1)|\alpha\rangle &= e^{i\beta(t,t_1)}e^{-\zeta^*(t,t_1)a+\zeta(t,t_1)a^\dagger}e^{\alpha a^\dagger-\alpha^* a}|0\rangle \\ &= e^{i\beta(t,t_1)}e^{(\zeta\alpha^*-\zeta^*\alpha)/2}e^{(\zeta+\alpha)a^\dagger-(\zeta^*+\alpha^*)a}|0\rangle \\ &= e^{i\gamma(t,t_1)}|\alpha + \zeta\rangle \end{aligned} \quad (14.141)$$

where γ, like β, is a numerical phase. We arrive at the intriguing and important conclusion that, under the influence of the (dipole) interaction $f(t)a + f^*(t)a^\dagger$, a coherent state remains coherent at all times, because the time development operator (14.137) is a displacement operator for coherent states, like the operator D in (10.98).

Of particular interest is the limit of the operator $\tilde{T}(t, t_1)$ as $t_1 \to -\infty$ and $t \to +\infty$. This limiting time development operator is known as the S (or *scattering*) *operator* and is defined formally as

$$\boxed{S = \tilde{T}(+\infty, -\infty)} \quad (14.142)$$

For the forced harmonic oscillator with an interaction of finite duration during the interval (T_1, T_2) the S operator is

$$\boxed{S = e^{i\beta}\exp\left[-\frac{i}{\hbar}g(\omega)a - \frac{i}{\hbar}g^*(\omega)a^\dagger\right]} \quad (14.143)$$

where we have denoted

$$\beta = \beta(+\infty, -\infty) \tag{14.144}$$

Substituting the expression for $g(\omega)$ defined in (14.120), we obtain

$$S = e^{i\beta} \exp\left\{-\frac{i}{\hbar}\int_{-\infty}^{+\infty} [f(t)ae^{-i\omega t} + f^*(t)a^\dagger e^{i\omega t}]\, dt\right\} \tag{14.145}$$

As the link between the states of the system before the onset and after the cessation of the interaction, the scattering operator, or the S matrix representing it, was first illustrated in Section 6.3. We will encounter the same concepts again in Chapter 20.

If the oscillator is in the ground state before the start of the interaction, what is the probability that it will be found in the nth excited oscillator energy eigenstate after the interaction terminates? The interaction produces the state $S|0\rangle$, which is a coherent state with eigenvalue $\alpha = -(i/\hbar)g^*(\omega)$. The transition probability of finding the oscillator in the nth eigenstate *after* the interaction is, according to (10.110), the Poisson distribution

$$\begin{aligned} |\langle n|S|0\rangle|^2 &= |\langle n|\alpha = -ig^*(\omega)/\hbar\rangle|^2 = P_n(\alpha) \\ &= \frac{1}{n!}\left|\frac{g(\omega)}{\hbar}\right|^{2n} e^{-|g(\omega)|^2/\hbar^2} = \frac{\langle n\rangle^n}{n!}\, e^{-\langle n\rangle} \end{aligned} \tag{14.146}$$

with expectation value $\langle n\rangle = |g(\omega)|^2/\hbar^2$ for the oscillator quantum number. These results can be interpreted in terms of a system of quanta, n being the number of quanta present. The interaction term in the Hamiltonian is linear in $a^\dagger$ and a and creates or annihilates quanta. The strength of the interaction determines the average number $\langle n\rangle$ of quanta present and characterizes the Poisson distribution, which represents the probability that a dipole interaction pulse incident on the vacuum state of our system of quanta leaves after its passage a net number of n quanta behind. These features of the dynamics of the forced or driven linear harmonic oscillator will help us understand the creation and annihilation of photons in Chapter 23.

Finally, we use the results from the interaction picture to deduce the time development operator in the Schrödinger picture. From Eq. (14.49) we infer that

$$T(t_2, t_1) = e^{-(i/\hbar)H_0 t_2}\tilde{T}(t_2, t_1)e^{(i/\hbar)H_0 t_1} \tag{14.147}$$

If we employ the oscillator Hamiltonian (14.121) for H_0 and the time development operator (14.137) in the interaction picture, we obtain

$$\boxed{T(t_2, t_1) = e^{i\beta(t_2,t_1)} \exp\{-\zeta^*(t_2, t_1)ae^{i\omega t_2} + \zeta(t_2, t_1)a^\dagger e^{-i\omega t_2}\}e^{-(i/\hbar)H_0(t_2-t_1)}} \tag{14.148}$$

Exercise 14.14. If $[A, B] = \gamma B$ (as in Exercise 3.15) prove that

$$e^{\lambda A}f(B)e^{-\lambda A} = f(e^{\lambda\gamma}B) \tag{14.149}$$

Exercise 14.15. Verify the expression (14.148) for the time development operator by applying (14.149) and (3.61).

Exercise 14.16. Show that

$$S^\dagger aS = a - \frac{i}{\hbar}g^*(\omega) = a + a_{\text{out}} - a_{\text{in}} \tag{14.150}$$

where S is the operator defined in Eq. (14.142) and a_{in} and a_{out} are related by (14.119). Noting that the operators a, a_{in}, and a_{out} differ from each other by additive constants, and using the unitarity of S, deduce that

$$S^{\dagger} a_{\text{in}} S = a_{\text{out}} \tag{14.151}$$

Exercise 14.17. For a forced harmonic oscillator with a transient perturbation (14.106), derive the change in the unperturbed energy, if $|\Psi\rangle$ is the initial state of the oscillator in the interaction picture, at asymptotically long times before the onset of the interaction. Show that

$$\Delta E = i\omega g(\omega)\langle a\rangle - i\omega g^{*}(\omega)\langle a^{\dagger}\rangle + \omega |g(\omega)|^2/\hbar \tag{14.152}$$

If $|\Psi\rangle$ is the ground state of the oscillator, verify that ΔE given by (14.152) agrees with a direct calculation based on the Poisson distribution formula (14.146).

Problems

1. A particle of charge q moves in a uniform magnetic field B which is directed along the z axis. Using a gauge in which $A_z = 0$, show that $q = (cp_x - qA_x)/qB$ and $p = (cp_y - qA_y)/c$ may be used as suitable canonically conjugate coordinate and momentum together with the pair z, p_z. Derive the energy spectrum and the eigenfunctions in the q-representation. Discuss the remaining degeneracy. Propose alternative methods for solving this eigenvalue problem.

2. A linear harmonic oscillator is subjected to a spatially uniform external force $F(t) = C\eta(t)e^{-\lambda t}$ where λ is a positive constant and $\eta(t)$ the Heaviside step function (A.23). If the oscillator is in the ground state at $t < 0$, calculate the probability of finding it at time t in an oscillator eigenstate with quantum number n. Assuming $C = (\hbar m\lambda^3)^{1/2}$, examine the variation of the transition probabilities with n and with the ratio λ/ω, ω being the natural frequency of the harmonic oscillator.

3. If the term $V(t)$ in the Hamiltonian changes suddenly (''impulsively'') between time t and $t + \Delta t$, in a time Δt short compared with all relevant periods, and assuming only that $[V(t'), V(t'')] = 0$ during the impulse, show that the time development operator is given by

$$T(t + \Delta t, t) = \exp\left[-\frac{i}{\hbar}\int_{t}^{t+\Delta t} V(t')\,dt'\right]$$

   Note especially that the state vector remains unchanged during a sudden change of V by a finite amount.

4. A linear harmonic oscillator in its ground state is exposed to a spatially constant force which at $t = 0$ is suddenly removed. Compute the transition probabilities to the excited states of the oscillator. Use the generating function for Hermite polynomials to obtain a general formula. How much energy is transferred?

5. In the nuclear beta decay of a tritium atom (^{3}H) in its ground state, an electron is emitted and the nucleus changes into an ^{3}He nucleus. Assume that the change is sudden, and compute the probability that the atom is found in the ground state of the helium ion after the emission. Compute the probability of atomic excitation to the $2S$ and $2P$ states of the helium ion. How probable is excitation to higher levels, including the continuum?

. A linear harmonic oscillator, with energy eigenstates $|n\rangle$, is subjected to a time-dependent interaction between the ground state $|0\rangle$ and the first excited state:

$$V(t) = F(t)|1\rangle\langle 0| + F^*(t)|0\rangle\langle 1|$$

(a) Derive the coupled equations of motion for the probability amplitudes $\langle n|\Psi(t)\rangle$.

(b) If $F(t) = \sqrt{2}\hbar\omega\eta(t)$, obtain the energy eigenvalues and the stationary states for $t > 0$.

(c) If the system is in the ground state of the oscillator before $t = 0$, calculate $\langle n|\Psi(t)\rangle$ for $t > 0$.

7. At $t < 0$ a system is in a coherent state $|\alpha\rangle$ (eigenstate of a) of an oscillator and subjected to an impulsive interaction

$$V(t) = \frac{i\hbar}{2}\zeta(a^2 - a^{\dagger 2})\delta(t)$$

where ζ is a real-valued parameter. Show that the sudden change generates a squeezed state. If the oscillator frequency is ω, derive the time dependence of the variances

$$\left[\Delta\left(\frac{a + a^\dagger}{\sqrt{2}}\right)\right]^2 \quad \text{and} \quad \left[\Delta\left(\frac{a - a^\dagger}{\sqrt{2}i}\right)\right]^2$$

CHAPTER 15

The Quantum Dynamics of a Particle

In this chapter, we develop quantum dynamics beyond the general framework of Chapter 14. Generalizing what was done in Section 9.6 for a particle in one dimension, we summarize the three-dimensional quantum dynamics of a particle of mass m (and charge q) in the coordinate or momentum representations. In the coordinate representation, the time evolution of any amplitude is compactly, and for many applications conveniently, derived from the initial state in integral form in terms of a propagator (Green's function). This view of the dynamics of a system leads to Feynman's path integral method, to which a bare-boned introduction is given. Since most quantum systems are composed of more than just one particle in an external field, we show how the formalism is generalized to interacting compound systems and their entangled states. Finally, in the quantum domain it is often not possible to prepare a system in a definite (pure) state $|\Psi\rangle$. Instead, it is common that our information is less than complete and consists of probabilities for the different pure states that make up a statistical ensemble representing the state of the system. We extend the principles of quantum dynamics to the density operator (or density matrix) as the proper tool for the description of the partial information that is generally available. From information theory we borrow the concept of the entropy as a quantitative measure of information that resides in the probability distribution.

1. ***The Coordinate and Momentum Representations.*** We now apply the quantization procedure of Section 14.3 to the dynamics of a particle in three dimensions. No really new results will be obtained, but the derivation of the familiar equations of Section 9.6 from the general theory will be sketched, with appropriate notational changes. Practice in manipulating the bra-ket formalism can be gained by filling in the detailed proofs, which are omitted here.

We assume that the Hermitian operators $\mathbf{r}$ and $\mathbf{p}$ each form a complete set of commuting canonically conjugate observables for the system, satisfying the commutation relations (14.58) and (14.59), compactly written as

$$\mathbf{rp} - \mathbf{pr} = i\hbar \boldsymbol{I} \tag{15.1}$$

and that the results of the measurement of these quantities can be any real number between $-\infty$ and $+\infty$. In Eq. (15.1), the bold italic $\boldsymbol{I}$ denotes an operator in Hilbert space as well as the identity matrix (dyadic) in three dimensions.

The eigenvalues of $\mathbf{r}$, denoted by $\mathbf{r}'$, form a real three-dimensional continuum. The corresponding eigenkets are denoted by $|\mathbf{r}'\rangle$, and we have

$$\mathbf{r}|\mathbf{r}'\rangle = \mathbf{r}'|\mathbf{r}'\rangle \tag{15.2}$$

with the usual normalization

$$\langle\mathbf{r}''|\mathbf{r}'\rangle = \delta(\mathbf{r}' - \mathbf{r}'') \tag{15.3}$$

in terms of the three-dimensional delta function. The closure relation,

$$\int |\mathbf{r}'\rangle\langle\mathbf{r}'|\, d^3\mathbf{r}' = 1 \tag{15.4}$$

is the main computational device for deriving useful representation formulas.

The (coordinate) wave function $\psi(\mathbf{r}')$ is the probability amplitude for a state $|\Psi\rangle$:

$$\psi(\mathbf{r}') = \langle\mathbf{r}'|\Psi\rangle \tag{15.5}$$

and the momentum wave function is analogously defined as

$$\phi(\mathbf{p}') = \langle\mathbf{p}'|\Psi\rangle \tag{15.6}$$

The action of an operator A on an arbitrary state $|\Psi\rangle$ is expressed by

$$\begin{aligned} A|\Psi\rangle &= \int\int |\mathbf{r}''\rangle\langle\mathbf{r}''|A|\mathbf{r}'\rangle\langle\mathbf{r}'|\Psi\rangle\, d^3r''\, d^3r' = \int |\mathbf{r}''\rangle\langle\mathbf{r}''|A|\mathbf{r}'\rangle\psi(\mathbf{r}')\, d^3r''\, d^3r' \\ &= \int\int |\mathbf{p}''\rangle\langle\mathbf{p}''|A|\mathbf{p}'\rangle\langle\mathbf{p}'|\Psi\rangle\, d^3p''\, d^3p' = \int |\mathbf{p}''\rangle\langle\mathbf{p}''|A|\mathbf{p}'\rangle\phi(\mathbf{p}')\, d^3p''\, d^3p' \end{aligned} \tag{15.7}$$

The *matrix elements* of an operator A in the coordinate or momentum representation are continuous rather than discrete functions of the indices. They can be calculated from the definition (15.2) and the commutation relations (15.1). If $A = F(\mathbf{r}, \mathbf{p})$ is an operator that is a function of position *and* momentum, care must be taken to maintain the ordering of noncommuting operators, such as x and p_x, that may occur. If $F(\mathbf{r}, \mathbf{p})$ can be expressed as a power series in terms of the momentum $\mathbf{p}$, we can transcribe the derivation in Section 9.6 to obtain

$$\langle\mathbf{r}''|F(\mathbf{r}, \mathbf{p})|\mathbf{r}'\rangle = F\left(\mathbf{r}'', \frac{\hbar}{i}\nabla''\right)\delta(\mathbf{r}' - \mathbf{r}'') \tag{15.8}$$

Equations (15.7) and (15.8) provide us with the necessary ingredients for rederiving the formulas in Section 3.2. For example, after substituting (15.8) and performing the r'' integration, we can obtain from (15.7),

$$\langle\Psi_b|F(\mathbf{r}, \mathbf{p})|\Psi_a\rangle = \int \psi_b^*(\mathbf{r}')F\left(\mathbf{r}', \frac{\hbar}{i}\nabla'\right)\psi_a(\mathbf{r}')\, d^3r' \tag{15.9}$$

in agreement with Eq. (3.31).

Once we are at home in the coordinate representation,[1] the transformation $\mathbf{r} \to \mathbf{p}$ and $\mathbf{p} \to -\mathbf{r}$, which leaves the commutation relations (15.1) invariant, may be used to convert any coordinate-representation formula into one that is valid in the momentum representation.

Exercise 15.1. Construct the relations that are analogous to (15.8) and (15.9) in the momentum representation.

[1]The coordinate representation employed here is often also referred to as the *Schrödinger* representation. Since we chose to attach Schrödinger's name to one of the *pictures* of quantum dynamics, we avoid this terminology.

Exercise 15.2. Prove that

$$\begin{aligned}\langle \mathbf{r}'|F(\mathbf{r},\mathbf{p})|\Psi\rangle &= F\left(\mathbf{r}', \frac{\hbar}{i}\nabla_{\mathbf{r}'}\right)\langle \mathbf{r}'|\Psi\rangle = F\left(\mathbf{r}', \frac{\hbar}{i}\nabla_{\mathbf{r}'}\right)\psi(\mathbf{r}') \\ \langle \mathbf{p}'|F(\mathbf{r},\mathbf{p})|\Psi\rangle &= F(i\hbar\nabla_{\mathbf{p}'}, \mathbf{p}')\langle \mathbf{p}'|\Psi\rangle = F(i\hbar\nabla_{\mathbf{p}'}, \mathbf{p}')\phi(\mathbf{p}')\end{aligned} \tag{15.10}$$

Equations (15.10) can be used to translate any algebraic equation of the form

$$F(\mathbf{r},\mathbf{p})|\Psi\rangle = 0 \tag{15.11}$$

in the coordinate and momentum representations as

$$F\left(\mathbf{r}', \frac{\hbar}{i}\nabla_{\mathbf{r}'}\right)\psi(\mathbf{r}') = F(i\hbar\nabla_{\mathbf{p}'}, \mathbf{p}')\phi(\mathbf{p}') = 0 \tag{15.12}$$

As an application, the eigenvalue equation for momentum,

$$\mathbf{p}|\mathbf{p}'\rangle = \mathbf{p}'|\mathbf{p}'\rangle$$

in the coordinate representation takes the differential equation form

$$\frac{\hbar}{i}\nabla_{\mathbf{r}'}\langle \mathbf{r}'|\mathbf{p}'\rangle = \mathbf{p}'\langle \mathbf{r}'|\mathbf{p}'\rangle \tag{15.13}$$

for the momentum-coordinate transformation coefficients $\langle \mathbf{r}'|\mathbf{p}'\rangle$. The solution of this differential equation is

$$\langle \mathbf{r}'|\mathbf{p}'\rangle = g(\mathbf{p}')e^{(i/\hbar)\mathbf{p}'\cdot\mathbf{r}'} \tag{15.14}$$

where $g(\mathbf{p}')$ is to be determined from a normalization condition. In (15.14) there is no restriction on the values of $\mathbf{p}'$ other than that they must be real in order to keep the eigenfunctions from diverging in coordinate space. The eigenvalues are thus continuous, and an appropriate normalization condition requires that

$$\begin{aligned}\langle \mathbf{p}'|\mathbf{p}''\rangle = \delta(\mathbf{p}' - \mathbf{p}'') &= \int \langle \mathbf{p}''|\mathbf{r}'\rangle\langle \mathbf{r}'|\mathbf{p}'\rangle\, d^3r' \\ &= g^*(\mathbf{p}'')g(\mathbf{p}')\int e^{(i/\hbar)(\mathbf{p}'-\mathbf{p}'')\cdot\mathbf{r}'}\, d^3r' = (2\pi\hbar)^3|g(\mathbf{p}')|^2\delta(\mathbf{p}' - \mathbf{p}'')\end{aligned}$$

Hence,

$$g(\mathbf{p}') = \frac{1}{(2\pi\hbar)^{3/2}}e^{i\alpha(\mathbf{p}')}$$

Arbitrarily, but conveniently, the phase factor is chosen to be unity so that finally we arrive at the standard form of the transformation coefficients,

$$\boxed{\langle \mathbf{r}'|\mathbf{p}'\rangle = \frac{1}{(2\pi\hbar)^{3/2}}e^{(i/\hbar)\mathbf{p}'\cdot\mathbf{r}'}} \tag{15.15}$$

These probability amplitudes linking the coordinate and momentum representations allow us to reestablish the connection between coordinate and momentum wave functions for a given state $|\psi\rangle$:

$$\begin{aligned}\phi(\mathbf{p}') = \langle \mathbf{p}'|\Psi\rangle &= \int \langle \mathbf{p}'|\mathbf{r}'\rangle\langle \mathbf{r}'|\Psi\rangle\, d^3r' \\ &= (2\pi\hbar)^{-3/2}\int d^3r'\, e^{-(i/\hbar)\mathbf{p}'\cdot\mathbf{r}'}\psi(\mathbf{r}')\end{aligned} \tag{15.16}$$

and its Fourier inverse,

$$\psi(\mathbf{r}') = (2\pi\hbar)^{-3/2} \int d^2p'\ e^{(i/\hbar)\mathbf{p}'\cdot\mathbf{r}'}\phi(\mathbf{p}') \tag{15.17}$$

These relations are, of course, familiar from Chapter 3.

There is a fundamental arbitrariness in the definition of the basis $|\mathbf{r}'\rangle$ of the coordinate representation, which has consequences for the wave function and for the matrix elements of operators. This ambiguity arises from the option to apply to all basis vectors a unitary transformation that merely changes the phases by an amount that may depend on $\mathbf{r}'$. The new coordinate basis is spanned by the basis vectors

$$|\overline{\mathbf{r}'}\rangle = e^{i\varphi(\mathbf{r})/\hbar}|\mathbf{r}'\rangle = e^{i\varphi(\mathbf{r}')/\hbar}|\mathbf{r}'\rangle \tag{15.18}$$

where $\varphi(\mathbf{r})$ is an arbitrary scalar field. In terms of this new coordinate representation, the wave function for state $|\Psi\rangle$ changes from $\psi(\mathbf{r}') = \langle\mathbf{r}'|\Psi\rangle$ to

$$\langle\overline{\mathbf{r}'}|\Psi\rangle = \overline{\psi}(\mathbf{r}') = e^{-i\varphi(\mathbf{r}')/\hbar}\psi(\mathbf{r}') \tag{15.19}$$

In the new basis, the matrix elements of an operator $F(\mathbf{r}, \mathbf{p})$ are

$$\begin{aligned}\langle\overline{\mathbf{r}''}|F(\mathbf{r}, \mathbf{p})|\overline{\mathbf{r}'}\rangle &= \langle\mathbf{r}''|e^{-i\varphi(\mathbf{r})/\hbar}F(\mathbf{r}, \mathbf{p})e^{i\varphi(\mathbf{r})/\hbar}|\mathbf{r}'\rangle \\ &= \langle\mathbf{r}''|F(\mathbf{r}, \mathbf{p} - \nabla\varphi|\mathbf{r}'\rangle = F\left(\mathbf{r}'', \frac{\hbar}{i}\nabla'' - \nabla''\varphi\right)\delta(\mathbf{r}' - \mathbf{r}'')\end{aligned} \tag{15.20}$$

If, when calculating matrix elements in the new basis, we are to keep the rule (15.8), we must make the transformation

$$F(\mathbf{r}, \mathbf{p}) \to F(\mathbf{r}, \mathbf{p} + \nabla\varphi) \tag{15.21}$$

in order to cancel the undesirable extra term on the right-hand side of (15.20). The replacement (15.21) is consistent with the commutation relations for $\mathbf{r}$ and $\mathbf{p}$, since the addition to the momentum $\mathbf{p}$ of an arbitrary irrotational vector field, $\mathbf{V}(\mathbf{r}) = \nabla\varphi(\mathbf{r})$, leaves the conditions (15.1) invariant. The ambiguity in the choice of coordinate basis is seen to be a manifestation of the gauge invariance of the theory, as discussed in Section 4.6 (except that here we have chosen the passive viewpoint).

Exercise 15.3. Prove that the commutation relations (15.1) remain valid if the substitution $\mathbf{p} \to \mathbf{p} + \mathbf{V}(\mathbf{r})$ is made, provided that the vector field $\mathbf{V}$ is irrotational, i.e., $\nabla \times \mathbf{V} = 0$.

When we apply the rules (15.12) to the equation of motion (14.12), we obtain in the coordinate representation the differential equation

$$i\hbar\frac{\partial}{\partial t}\langle\mathbf{r}'|\Psi(t)\rangle = \langle\mathbf{r}'|H|\Psi(t)\rangle = H\left(\mathbf{r}', \frac{\hbar}{i}\nabla'\right)\langle\mathbf{r}'|\Psi(t)\rangle \tag{15.22}$$

for the time-dependent wave function

$$\psi(\mathbf{r}', t) = \langle\mathbf{r}'|\Psi(t)\rangle \tag{15.23}$$

and similarly in the momentum representation,

$$i\hbar\frac{\partial}{\partial t}\langle\mathbf{p}'|\Psi(t)\rangle = \langle\mathbf{p}'|H|\Psi(t)\rangle = H(i\hbar\nabla_{\mathbf{p}'}, \mathbf{p}')\langle\mathbf{p}'|\Psi(t)\rangle \tag{15.24}$$

for the momentum wave function

$$\phi(\mathbf{p}', t) = \langle \mathbf{p}' | \Psi(t) \rangle \tag{15.25}$$

If the Hamiltonian operator H has the common form

$$H = \frac{p^2}{2m} + V(\mathbf{r}) \tag{15.26}$$

we obtain from (15.22), unsurprisingly, the time-dependent Schrödinger or wave equation, (3.1):

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}', t) = \left[-\frac{\hbar^2}{2m} \nabla'^2 + V(\mathbf{r}') \right] \psi(\mathbf{r}', t) \tag{15.27}$$

In the momentum representation, we obtain in a similar fashion

$$i\hbar \frac{\partial}{\partial t} \phi(\mathbf{p}', t) = \frac{p'^2}{2m} \phi(\mathbf{p}', t) + \int \langle \mathbf{p}' | V(\mathbf{r}) | \mathbf{p}'' \rangle \, d^3p'' \, \phi(\mathbf{p}'', t) \tag{15.28}$$

The matrix element $\langle \mathbf{p}' | V(\mathbf{r}) | \mathbf{p}'' \rangle$ is a Fourier integral of the interaction potential.

Exercise 15.4. By expressing $\langle \mathbf{p}' | V(\mathbf{r}) | \mathbf{p}'' \rangle$ in the coordinate representation, verify the equivalence of (15.28) and (3.21).

We have come full circle and have found wave mechanics in its coordinate or momentum version to be a realization of the general theory of quantum dynamics formulated in abstract vector space. The equivalence of these various forms of quantum mechanics shows again that the constant $\hbar$ was correctly introduced into the general theory in Chapter 14. Planck's constant has the same value in all cases.

2. ***The Propagator in the Coordinate Representation.*** If the system is a particle in an external electromagnetic (gauge vector) field described by the potentials $\mathbf{A}(\mathbf{r}, t)$ and $\phi(\mathbf{r}, t)$, the Hamiltonian operator

$$H = \frac{1}{2m}\left(\mathbf{p} - \frac{q}{c} \mathbf{A} \right)^2 + V \tag{15.29}$$

must be used for the transcription of the equation of motion into the coordinate (or momentum) representation. Here, the potential energy V includes the fourth component of the electromagnetic potential. Generally, the Hamiltonian is time-dependent. Hamiltonians for more complex systems must be constructed appropriately when the need arises.

Since we usually choose to work in one particular dynamical picture and one specific representation, the cumbersome notation that distinguishes between different pictures and between operators and their eigenvalues (by the use of primes) can be dispensed with. In the Schrödinger picture and the coordinate representation, the equation of motion is

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = \left[-\frac{\hbar^2}{2m} \left(\nabla - \frac{iq}{\hbar c} \mathbf{A} \right)^2 + V \right] \psi(\mathbf{r}, t) \tag{15.30}$$

In order to transcribe the equation of motion (14.11) for the time development operator in the coordinate representation, we define the *propagator* as the transition amplitude,

$$\boxed{K(\mathbf{r}, \mathbf{r}'; t, t') = \langle \mathbf{r} | T(t, t') | \mathbf{r}' \rangle} \tag{15.31}$$

the propagator satisfies the initial condition (for equal time t):

$$K(\mathbf{r}, \mathbf{r}'; t, t) = \delta(\mathbf{r} - \mathbf{r}') \tag{15.32}$$

and is a nonnormalizable solution of the time-dependent Schrödinger equation

$$\boxed{i\hbar \frac{\partial}{\partial t} K(\mathbf{r}, \mathbf{r}'; t, t') = \left[-\frac{\hbar^2}{2m} \left(\nabla - \frac{iq}{\hbar c} \mathbf{A} \right)^2 + V \right] K(\mathbf{r}, \mathbf{r}'; t, t')} \tag{15.33}$$

The propagator can also be identified as a (spatial) *Green's function* for the time-dependent Schrödinger equation.

Exercise 15.5. From the definition (15.31) of the propagator and the Hermitian character of the Hamiltonian show that

$$K(\mathbf{r}', \mathbf{r}; t', t) = K^*(\mathbf{r}, \mathbf{r}'; t, t') \tag{15.34}$$

linking the transition amplitude that reverses the dynamical development between spacetime points $(\mathbf{r}, t)$ and $(\mathbf{r}', t')$ to the original propagator.

From its definition (15.31) and the composition rule (14.3) we see that the propagator relates the Schrödinger wave functions at two different times:

$$\boxed{\psi(\mathbf{r}, t) = \int K(\mathbf{r}, \mathbf{r}'; t, t') \psi(\mathbf{r}', t') \, d^3r'} \tag{15.35}$$

This relation justifies the name *propagator* for K. Equation (15.35) can also be read as an integral equation for the wave function; K is its *kernel*.

Exercise 15.6. If a gauge transformation is performed, as in Section 4.6, what happens to the propagator? Derive its transformation property.

The *retarded* Green's function, defined as $G_R = 0$ for $t < t'$, is related to the propagator by

$$G_R(\mathbf{r}, t; \mathbf{r}', t') = -\frac{i}{\hbar} \eta(t - t') K(\mathbf{r}, \mathbf{r}'; t, t') \tag{15.36}$$

and satisfies the inhomogeneous equation

$$i\hbar \frac{\partial G_R}{\partial t} = H(t) G_R + \delta(\mathbf{r} - \mathbf{r}') \delta(t - t') \tag{15.37}$$

The propagator and the retarded Green's function are at the center of Feynman's formulation of quantum mechanics, to which a brief introduction will be given in Section 15.3.

Exercise 15.7. Prove Eq. (15.37).

The corresponding *advanced* Green's function is defined as

$$G_A(\mathbf{r}, t; \mathbf{r}', t') = \frac{i}{\hbar}\,\eta(t' - t)K(\mathbf{r}, \mathbf{r}'; t, t') \tag{15.38}$$

It satisfies the same equation (15.37) as the retarded Green's function. Since for the Heaviside step function:

$$\eta(t) + \eta(-t) = 1 \tag{15.39}$$

we note that the two Green's functions, which solve the inhomogeneous equation (15.37), are connected with the propagator (which is a solution of the Schrödinger equation without the inhomogeneous term) by

$$G_R - G_A = -\frac{i}{\hbar}\,K \tag{15.40}$$

If the Hamiltonian is time-independent, the time development operator is, explicitly,

$$T(t, t') = \exp\left[-\frac{i}{\hbar}\,H(t - t')\right] \tag{15.41}$$

and the propagator (15.31) then depends on the times t and t' only through their difference $t - t'$,

$$K(\mathbf{r}, \mathbf{r}'; t - t') = \sum_n \langle \mathbf{r}|\, \exp\left[-\frac{i}{\hbar}\,H(t - t')\right] |E_n\rangle\langle E_n\,|\mathbf{r}'\rangle$$

where

$$H|E_n\rangle = E_n\,|E_n\rangle \tag{15.42}$$

Hence, in terms of the energy eigenfunctions, explicitly,

$$\boxed{K(\mathbf{r}, \mathbf{r}'; t - t') = \sum_n \psi_n^*(\mathbf{r}')\psi_n(\mathbf{r}) \exp\left[-\frac{i}{\hbar}\,E_n(t - t')\right]} \tag{15.43}$$

The sum must be extended over the complete set of stationary states, including the degenerate ones.

If in the Hamiltonian (15.29) the vector potential $\mathbf{A} = 0$ and $V(\mathbf{r})$ is static (or an even function of t), the equation of motion is invariant under time reversal, as discussed in Section 3.5. In this case $K^*(\mathbf{r}, \mathbf{r}'; t' - t)$ is also a solution of the time-dependent Schrödinger equation (15.33). Since the initial condition (15.32) selects a unique solution, we must have, owing to time reversal symmetry,

$$K^*(\mathbf{r}, \mathbf{r}'; t' - t) = K(\mathbf{r}, \mathbf{r}'; t - t') \tag{15.44}$$

Comparison of this result with the general property (15.34) of the propagator shows that under these quite common circumstances the propagator is *symmetric* with respect to the space coordinates:

$$\boxed{K(\mathbf{r}, \mathbf{r}'; t - t') = K(\mathbf{r}', \mathbf{r}; t - t')} \tag{15.45}$$

The simplest and most important example of a propagator is the one for the free particle. In one spatial dimension,

$$T(t, t') = \exp\left[-\frac{i}{\hbar} \frac{p_x^2}{2m} (t - t') \right] \tag{15.46}$$

and, applying (15.43), in its integral form suitable for the continuous momentum variable,

$$K(x, x'; t - t') = \int_{-\infty}^{+\infty} e^{-(i/\hbar)p_x^2(t-t')/2m} \, e^{-(i/\hbar)p_x(x-x')} \, dp_x \tag{15.47}$$

This Fourier integral can be performed explicitly:

$$\boxed{K(x, x'; t - t') = \sqrt{\frac{m}{2\pi i\hbar(t - t')}} \exp\left\{ -\frac{m(x - x')^2}{2i\hbar(t - t')} \right\}} \tag{15.48}$$

Exercise 15.8. Verify that (15.48) solves the time-dependent Schrödinger equation and agrees with the initial condition (15.32). If the initial ($t' = 0$) state of a free particle is represented by $\psi(x, 0) = e^{ik_0 x}$, verify that (15.35) produces the usual plane wave.

If the system is initially (at $t' = 0$) represented by the minimum uncertainty wave packet [see (10.66)]

$$\psi(x, 0) = [2\pi(\Delta x)_0^2]^{-1/4} \exp\left\{ -\frac{x^2}{4(\Delta x)_0^2} + ik_0 x \right\} \tag{15.49}$$

we have, by substituting (15.48) and (15.49) into (15.35),

$$\psi(x, t) = \sqrt{\frac{m}{2\pi i\hbar t}} [2\pi(\Delta x)_0^2]^{-1/4} \times \int_{-\infty}^{+\infty} \exp\left\{ -\frac{m(x - x')^2}{2i\hbar t} - \frac{[x' - 2ik_0(\Delta x)_0^2]^2}{4(\Delta x)_0^2} - k_0^2(\Delta x)_0^2 \right\} dx'$$

If the integration is carried out, this expression reduces to

$$\psi(x, t) = [2\pi(\Delta x)_0^2]^{-1/4} \left[1 + \frac{i\hbar t}{2m(\Delta x)_0^2} \right]^{-1/2} \exp\left[\frac{-\dfrac{x^2}{4(\Delta x)_0^2} + ik_0 x - ik_0^2 \dfrac{\hbar t}{2m}}{1 + \dfrac{i\hbar t}{2m(\Delta x)_0^2}} \right] \tag{15.50}$$

Exercise 15.9. Calculate $|\psi(x, t)|^2$ from (15.50) and show that the wave packet moves uniformly and at the same time spreads so that

$$(\Delta x)_t^2 = (\Delta x)_0^2 \left[1 + \frac{\hbar^2 t^2}{4m^2(\Delta x)_0^4} \right] \tag{15.51}$$

All these results for free-particle dynamics are in agreement with the preliminary conclusions reached in Section 2.4, in Problem 1 in Chapter 2 and Problem 2 in Chapter 3, and in Eq. (14.98). The generalization of the propagator formalism to

the motion of a free particle in three dimensions is straightforward, as long as Cartesian coordinates are used.

The linear harmonic oscillator is a second example. Setting $t' = 0$ for convenience, the Green's function solution of the equation

$$\frac{\partial K(x, x'; t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2 K(x, x'; t)}{\partial x^2} + \frac{1}{2}m\omega^2 K(x, x'; t) \tag{15.52}$$

which satisfies the initial condition (15.32), can be calculated by many different routes.

For instance, we may go back to the method developed in Sections 2.5 and 3.1 and write the propagator in terms of a function $S(x, x'; t)$:

$$K(x, x'; t) = e^{iS(x,x';t)/\hbar} \tag{15.53}$$

such that S satisfies the quantum mechanical Hamilton-Jacobi equation (3.2),

$$\frac{\partial S}{\partial t} + \frac{1}{2m}\left(\frac{\partial S}{\partial x}\right)^2 - \frac{i\hbar}{2m}\frac{\partial^2 S}{\partial x^2} + V(x) = 0 \tag{15.54}$$

where for the linear harmonic oscillator $V = m\omega^2 x^2/2$. Since this potential is an even function of the coordinate and since initially,

$$K(-x, -x'; 0) = K(x, x'; 0) \tag{15.55}$$

it follows that for all t,

$$S(x, x'; t) = S(-x,-x'; t) = S(x', x; t) = -S^*(x, x'; -t) \tag{15.56}$$

where the symmetry relations (15.34) and (15.45) have been employed. Combining all this information and attempting to solve (15.54) by a series in powers of x and x', we conjecture that S must have the form

$$S(x, x'; t) = a(t)x^2 + b(t)xx' + a(t)x'^2 + c(t) \tag{15.57}$$

Substitution of this *Ansatz* into (15.54) yields the coupled ordinary differential equations,

$$\dot{a} = -\frac{2}{m}a^2 - \frac{1}{2m\omega^2} = -\frac{b^2}{2m}, \qquad \dot{b} = -\frac{2}{m}ab, \qquad \dot{c} = \frac{i\hbar}{m}a \tag{15.58}$$

All of these requirements, including especially the initial condition (15.32), can be satisfied only if (15.58) is solved by

$$a(t) = \frac{1}{2}m\omega\cot\omega t, \qquad b(t) = -\frac{m\omega}{\sin\omega t}, \qquad c(t) = \frac{i\hbar}{2}\log\left(\frac{2\pi i\hbar}{m}\sin\omega t\right)$$

giving finally the result

$$\boxed{K(x, x'; t) = \left(\frac{m\omega}{2\pi i\hbar\sin\omega t}\right)^{1/2}\exp\left\{-\frac{m\omega}{2i\hbar\sin\omega t}(x^2\cos\omega t - 2xx' + x'^2\cos\omega t\right\}} \tag{15.59}$$

In the limit $\omega \to 0$, the propagator (15.59) for the oscillator reduces to the free-particle propagator (15.48).

Exercise 15.10. Applying Mehler's formula (5.46) to the stationary state expansion (15.43) of the propagator, verify the result (15.59) for the linear harmonic oscillator. Conversely, show that (15.59) has the form

$$e^{-i\omega t/2} f(e^{-i\omega t})$$

and deduce the energy eigenvalues of the oscillator by comparing with (15.43).

Exercise 15.11. Show that if the initial state of a harmonic oscillator is represented by the displaced ground state wave function

$$\psi(x, 0) = N \exp\left\{ -\frac{m\omega}{2\hbar} (x - x_0)^2 \right\} \tag{15.60}$$

the state at time t is

$$\psi(x, t) = N \exp\left\{ -\frac{m\omega}{2\hbar} (x - x_0 \cos \omega t)^2 - i\left(\frac{\omega}{2} t + \frac{m\omega}{\hbar} x_0 x \sin \omega t - \frac{m\omega}{4\hbar} x_0^2 \sin 2\omega t \right) \right\} \tag{15.61}$$

Show that $|\psi(x, t)|$ oscillates without any change of shape.

Although, in principle, the propagator (15.59) answers all questions about the dynamics of the harmonic oscillator, for many applications, especially in quantum optics, it is desirable to express the time development of the oscillator in terms of coherent states. This was done in Section 14.6 in the interaction picture for the forced or driven harmonic oscillator. Here we revert to the Schrödinger picture.

We know from Eq. (10.122) that the displaced oscillator ground state wave function (15.60) is the coordinate representative of a coherent state $|\alpha\rangle$, with $x_0 = \sqrt{2\hbar/m\omega}\ \alpha$. An initial state

$$|\Psi(0)\rangle = |\alpha\rangle \tag{15.62}$$

develops in time under the action of the free, unforced oscillator Hamiltonian in the Schrödinger picture (see Exercise 14.13) as

$$|\Psi(t)\rangle = e^{-i\omega(a^\dagger a + 1/2)t} |\alpha\rangle = e^{-i\omega t/2} |\alpha e^{-i\omega t}\rangle \tag{15.63}$$

In words: If we represent a coherent state $|\alpha\rangle$ by its complex eigenvalue α as a vector in the complex α plane (Figure 10.1), the time development of the oscillator is represented by a uniform clockwise rotation of the vector with angular velocity ω. Since

$$\alpha = \sqrt{\frac{m\omega}{2\hbar}} \left(\langle x \rangle + \frac{i}{m\omega} \langle p \rangle \right) \tag{15.64}$$

the complex α plane can be interpreted as a quantum mechanical analogue of classical phase space. The expectation values $\langle x \rangle$ and $\langle p \rangle$ perform harmonic oscillations, as dictated by classical mechanics.

Exercise 15.12. Using (10.122), show that except for a normalization factor the amplitude $\langle x|\Psi(t)\rangle$ calculated from (15.63) again yields the oscillating wave packet (15.61). Relate the normalization factors for the two expressions.

Exercise 15.13. For the harmonic oscillator, derive $\psi(x, t)$ directly from $\psi(x, 0)$ by expanding the initial wave function, which represents a displaced ground

state as in (15.60), in terms of stationary states. Use the generating function (5.33) to obtain the expansion coefficients and again to sum the expansion. Rederive (15.61).

Exercise 15.14. From (15.64) and its time development, derive the expectation values $\langle x \rangle_t$ and $\langle p \rangle_t$ in terms of their initial values.

The forced or driven harmonic oscillator represents the next stage in complexity of an important dynamical system. If the Hamiltonian has the form

$$H = \frac{p^2}{2m} + \frac{1}{2} m\omega^2 x^2 - xF(t) \tag{15.65}$$

where $F(t)$ stands for an arbitrary driving force, the propagator may be evaluated from the time development operator derived, in the interaction picture, in Section 14.6. Alternatively, it may be obtained by extending the solution of the quantum mechanical Hamilton-Jacobi equation to the forced oscillator, adding the time-dependent interaction term $-xF(t)$ to the potential energy $V(x)$ in (15.54). It is then possible (but tedious) to show that when the propagator is expressed in terms of the function $S(x, t)$ as in (15.53), the result (15.57) for the free oscillator can still be used, but it must be augmented by an interaction-dependent correction that is linear in the coordinates and has the form:

$$S_{\text{int}}(x, x'; t, t') = f(t, t')x - f(t', t)x' + g(t, t') \tag{15.66}$$

We only quote the results of the calculation:[2]

$$f(t, t') = \frac{1}{\sin \omega(t - t')} \int_{t'}^{t} F(t'') \sin \omega(t'' - t')\, dt'' \tag{15.67}$$

and

$$g(t, t') = \int_{t'}^{t} \frac{dt''}{\sin^2 \omega(t'' - t')} \int_{t'}^{t''} dt_1\, F(t_1) \int_{t'}^{t''} dt_2\, F(t_2) \sin \omega(t_1 - t') \sin \omega(t_2 - t') \tag{15.68}$$

All the quantum mechanical functions S that we have calculated so far are essentially the same as Hamilton's Principal Function in classical mechanics, except for a purely time-dependent term $c(t)$ in (15.57) and (15.58). This latter term arises from the presence of the term proportional to $\hbar$ (occasionally called the *quantum potential*) in the quantum mechanical Hamilton-Jacobi equation. It is responsible for the correct normalization of the propagator. This very close connection between classical and quantum mechanics is contingent on the simple form of the interaction potential as a polynomial of second degree in x, as in the generalized parametric and driven harmonic oscillator. The addition of anharmonic terms to the interaction complicates matters and makes solving the Hamilton-Jacobi equation more difficult. The resulting $S(x, t)$ will generally exhibit more distinctive quantum effects.

The propagator formalism and its expression in terms of the action function S provides a natural entrée to Feynman's path integral formulation of quantum dynamics.

[2]See Feynman and Hibbs (1965), Chapter 3.

3. ***Feynman's Path Integral Formulation of Quantum Dynamics.*** We saw in the last section that if the potential energy depends on the coordinate x only through terms that are linear or at most quadratic in x, the x and x' dependence of the propagator for a transition from spacetime (x', t') to (x, t) is entirely contained in a real-valued phase function, which is Hamilton's (classical) Principal Function S for the motion between these two spacetime points. Here we use S (in roman font) to denote the classical function to distinguish it from its quantum mechanical counterpart, S (in italics). In this section, we limit ourselves to the motion of a particle in one dimension.

From classical mechanics[3] it is known that $\mathrm{S}(x, x'; t, t')$ is the stationary value of the classical *action* function

$$I(x, x'; t, t') = \int_{t'}^{t} L(x(t''), \dot{x}(t''), t'')\, dt'' \tag{15.69}$$

where

$$L(x(t), \dot{x}(t), t) = \frac{1}{2} m\dot{x}^2 - V(x, t) \tag{15.70}$$

is the classical Lagrangian function for the simple one-dimensional one-particle system that we are considering. Hamilton's Principle for the variation of the action,

$$\delta I(x, x'; t, t') = \delta \int_{t'}^{t} L(x(t''), \dot{x}(t''), t')\, dt'' = 0 \tag{15.71}$$

singles out the motion $\mathrm{x}(t)$ that takes the particle from the initial spacetime point (x', t') to the final destination (x, t). Thus, Hamilton's Principal Function is

$$\mathrm{S}(x, x'; t, t') = \int_{t'}^{t} L(\mathrm{x}(t''), \dot{\mathrm{x}}(t''), t'')\, dt'' \tag{15.72}$$

where it is now understood that $\mathrm{x}(t)$ is the correct classical motion connecting the two endpoints.

Exercise 15.15. For a particle moving from spacetime point (x', t') to (x, t) with the classical Lagrangian $L = (1/2)m\dot{x}^2$, show that S, derived from Hamilton's Principle, reproduces the exponent in the free particle propagator (15.48).

Exercise 15.16. For a particle moving from spacetime point (x', t') to (x, t) with the classical Lagrangian $L = (1/2)m\dot{x}^2 - (1/2)m\omega^2 x^2$, show that S, derived from Hamilton's Principle, reproduces the exponent in the harmonic-oscillator propagator (15.59).

To derive the Feynman path integral expression for the propagator in quantum mechanics, we first observe that owing to the fundamental group property (14.8) of the time development, the propagator satisfies the composition rule:

$$\begin{aligned} K(x, x'; t, t') &= \int K(x, x''; t, t'')K(x'', x'; t'', t')\, dx'' \\ &= \langle x|T(t, t')|x'\rangle = \int \langle x|T(t, t'')|x''\rangle\langle x''|T(t'', t')|x'\rangle\, dx'' \end{aligned} \tag{15.73}$$

[3]Goldstein (1980), Section 10.1.

for any value of the time t''. In order to utilize a simple approximation, we partition the time interval (t', t) into N infinitesimally short intervals of duration ε. When this is done, the composition rule generalizes to

$$K(x, x'; t, t') = \int \ldots \int dx_{N-1} \ldots dx_1 \; K(x, x_{N-1}; t, t - \varepsilon) \ldots \\ K(x_2, x_1, t' + 2\varepsilon, t' + \varepsilon) K(x_1, x'; t' + \varepsilon, t') \tag{15.74}$$

The construction of this expression implies that the x integrations are to be performed as soon as any two successive propagators are multiplied. Equivalently, however, we may first multiply the N propagators, leaving the integrations to the end. Since each of the coordinates x_i ranges from $-\infty$ to $+\infty$, the composition of the N propagators may then be construed as a sum over infinitely many different *paths* from the initial spacetime point (x', t') to the final spacetime point (x, t), as indicated schematically in Figure 15.1. The propagator for each infinitesimal time interval is now approximated by assuming that the motion of the particle from (x_{n-1}, t_{n-1}) to (x_n, t_n) is governed by a potential that is at most a second-degree polynomial in x. From Section 15.2, we know that in this approximation the propagator that takes us from (x_{n-1}, t_{n-1}) to (x_n, t_n) is in the form

$$K(x_n, x_{n-1}; \varepsilon) = C(t_{n-1}, t_n) e^{(i/\hbar)S(x_n x_{n-1}; t_n, t_{n-1})} \tag{15.75}$$

If we multiply the N elementary propagators for a particular "path" in spacetime together in the integrand of (15.74) and take the limit $\varepsilon \to 0$ and $N \to \infty$, the additivity of the action function shows that each path contributes, in units of Planck's constant $\hbar$, a real phase

$$S[x(t)] \equiv \int_{t'}^{t} L(x(t''), \dot{x}(t''), t'') \, dt'' \tag{15.76}$$

Although all quantities in (15.76) are classical functions of coordinates and of time, the path $x(t)$ that takes the particle from the initial spacetime point (x', t') to the final destination (x, t) now is generally not the actual classical motion $x(t)$ that Hamilton's Principle selects.

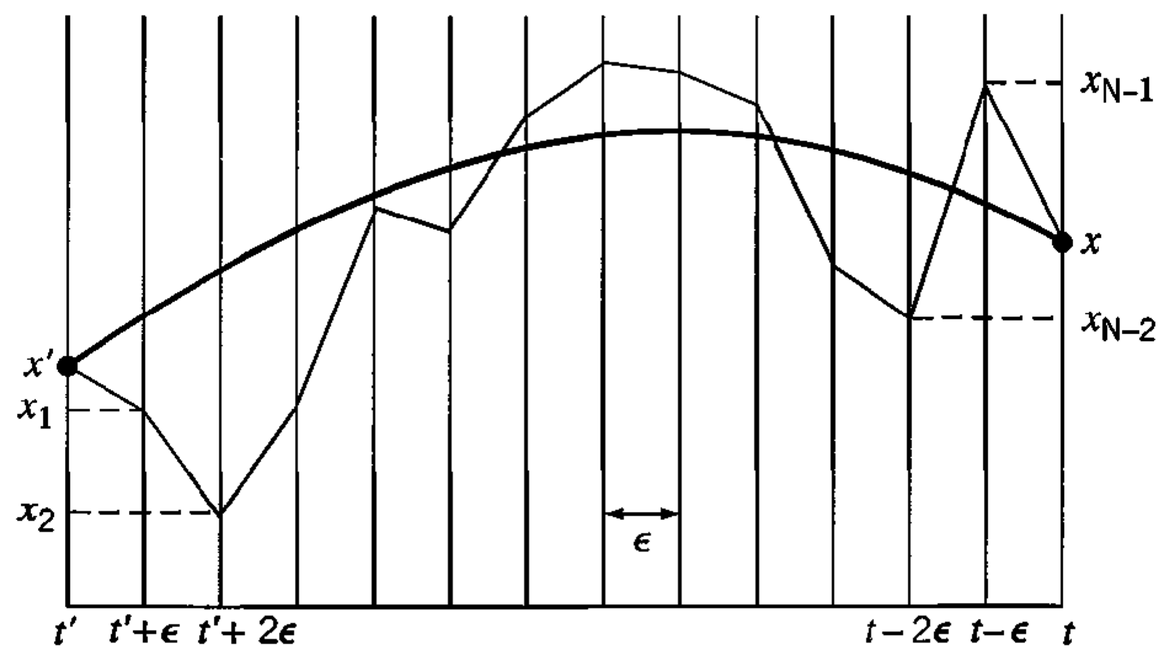


Figure 15.1 Paths linking the initial spacetime point (x', t') to the final spacetime point (x,t). The smooth curve represents the classical path $x(t)$ for the particle motion in the potential $V(x)$. The curve composed of broken straight segments is a typical path that makes a contribution $\exp\left\{\frac{i}{\hbar}S[x(t)]\right\}$ to the Feynman path integral.

Substituting the results from (15.75) and (15.76) into the composition rule (15.74), we finally arrive at *Feynman's path integral* formula for the propagator,

$$\boxed{\langle \overline{x,\ t} | \overline{x',\ t'} \rangle = K(x,\ x';\ t,\ t') = C \sum_{\text{all paths}} e^{(i/\hbar)S[x(t)]} = C \int e^{(i/\hbar)S[x(t)]} D[x(t)]} \quad (15.77)$$

where the factor C, which is independent of the coordinates, arises from the product of the time-dependent factors $C(t_{n-1},\ t_n)$ in the propagators (15.75) and is attributable to the term proportional to $i\hbar$ in the quantum mechanical Hamilton-Jacobi equation. In the last expression on the right-hand side of (15.77), the differential $D[x(t)]$ is intended to remind us that the propagator is a *functional integral*, in which the variable of integration is the function $x(t)$. To evaluate such an integral, which is the limit of the sum over all paths sketched in Figure 15.1, it is obviously necessary to extend the concepts of mathematical analysis beyond the standard repertoire and define an appropriate measure and a suitable parametrization in the space of possible paths.[4]

The derivation of (15.77) given here is a bit cavalier, but it captures the essence of the argument and produces correct results. To prove this, one can show that those contorted paths that are not accurately represented by the approximation (15.75) for the individual path segments contribute negligibly to the sum over all paths in (15.77), due to destructive interference caused by extremely rapid phase variations between neighboring paths.

Although it is in general a difficult mathematical problem, the integration over paths reduces in many applications effectively to the sum of contributions from only a few isolated paths. The *stationary phase method*, which in effect was already used in Chapter 2 for obtaining approximate wave functions, is a useful tool for evaluating the propagator by the Feynman path integral method. The actual classical spacetime path x(t) that connects the initial and final spacetime points, (x', t') and (x, t), corresponds, according to Hamilton's Principle, to the stationary phase in the path integral (15.77). The neighboring paths add constructively, and a first (semiclassical) approximation for the propagator is therefore

$$\boxed{K(x,\ x';\ t,\ t') \approx C(t,\ t')\ e^{(i/\hbar)S[x(t)]}} \quad (15.78)$$

We saw in Section 15.2 that this formula is not just an approximation, but is exact for a large class of problems, including the free particle and the harmonic oscillator, even with an arbitrary linear driving term.

In this chapter we have confined ourselves to describing the path integral formulation of quantum dynamics for the simple case of a nonrelativistic particle in one dimension, but the Feynman method is quite general.[5] For all systems that can be quantized by either method, it is equivalent to the canonical form of quantum mechanics, developed in Chapter 14, but the path integral approach offers a road to quantum mechanics for systems that are not readily accessible via Hamiltonian mechanics.

[4]Schulman (1981).

[5]For an excellent discussion of interference and diffraction of particle states in relation to path integrals, see Townsend (1992), Chapter 8.

4. ***Quantum Dynamics in Direct Product Spaces and Multiparticle Systems.*** Often the state vector space of a system can be regarded as the *direct*, *outer*, or *tensor* product of vector spaces for simpler subsystems. The direct product space is formed from two independent unrelated vector spaces that are respectively spanned by the basis vectors $|A_1'\rangle$ and $|B_2'\rangle$ by constructing the basis vectors

$$|A_1'B_2'\rangle = |A_1'\rangle \otimes |B_2'\rangle \equiv |A_1'\rangle|B_2'\rangle \tag{15.79}$$

Although the symbol $\otimes$ is the accepted mathematical notation for the direct product of state vectors, it is usually dispensed with in the physics literature, and we adopt this practice when it is unlikely to lead to misunderstandings. If n_1 and n_2 are the dimensions of the two factor spaces, the product space has dimension $n_1 \times n_2$. This idea is easily extended to the construction of direct product spaces from three or more simple spaces.

The most immediate example of a direct product space is the state vector space for a particle that is characterized by its position $\mathbf{r}(x, y, z)$. The basis vector $|\mathbf{r}\rangle = |x, y, z\rangle$ may be expressed as the direct product $|x\rangle \otimes |y\rangle \otimes |z\rangle \equiv |x\rangle|y\rangle|z\rangle$, since the three Cartesian coordinates can be chosen independently to specify the location of the particle. (On the other hand, the Euclidean three-space with basis vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, $\hat{\mathbf{z}}$, is the *sum* and not the product of the three one-dimensional spaces supported by $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$.)

Any operator that pertains to only one of the factor spaces is regarded as acting as an identity operator with respect to the other factor spaces. More generally, if M_1 and N_2 are two linear operators belonging to the vector spaces 1 and 2 such that

$$M_1|A_1'\rangle = \sum_{A_1''} |A_1''\rangle\langle A_1''|M_1|A_1'\rangle$$

and

$$N_2|B_2'\rangle = \sum_{B_2''} |B_2''\rangle\langle B_2''|N_2|B_2'\rangle$$

we define the *direct* or *tensor product operator* $M_1 \otimes N_2$ by the equation

$$M_1 \otimes N_2|A_1'B_2'\rangle = \sum_{A_1'', B_2''} |A_1''B_2''\rangle\langle A_1''|M_1|A_1'\rangle\langle B_2''|N_2|B_2'\rangle \tag{15.80}$$

Hence, $M_1 \otimes N_2$ is represented by a matrix that is said to be the *direct product* of the two matrices representing M_1 and N_2 separately and that is defined by

$$\boxed{\langle A_1''B_2''|M_1 \otimes N_2|A_1'B_2'\rangle = \langle A_1''|M_1|A_1'\rangle\langle B_2''|N_2|B_2'\rangle} \tag{15.81}$$

Exercise 15.17. If M_1 and P_1 are operators in space 1 and N_2 and Q_2 are operators in space 2, prove the identity

$$M_1P_1 \otimes N_2Q_2 = (M_1 \otimes N_2)(P_1 \otimes Q_2) \tag{15.82}$$

Check this identity for the corresponding matrices.

We are now prepared to generalize the formalism of one-particle quantum mechanics unambiguously to systems composed of several particles. If the particles are identical, very important peculiarities require consideration. Since Chapter 21 deals with these exclusively, we confine ourselves in this section to the quantum mechanics of systems with *distinguishable* particles. Furthermore, to make things clear, it is sufficient to restrict the discussion to systems containing just two particles. Ex-

amples are the ordinary hydrogenic atom or the muonium atom, including the dynamics of the nucleus, the deuteron composed of a proton and a neutron, and the positronium (electron and positron) atom.

We denote the two particles by the subscripts 1 and 2. As long as the spin can be ignored, six spatial coordinates are used to define the basis $|\mathbf{r}_1\mathbf{r}_2\rangle = |\mathbf{r}_1\rangle|\mathbf{r}_2\rangle$ for the two-particle system. In analogy to (15.5), we introduce the two-particle wave function

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \langle \mathbf{r}_1\mathbf{r}_2|\Psi\rangle \tag{15.83}$$

The interpretation of this probability amplitude is the usual one: $|\psi(\mathbf{r}_1, \mathbf{r}_2)|^2\, d^3r_1\, d^3r_2$ is proportional to the probability that particle 1 is found in volume element d^3r_1 centered at $\mathbf{r}_1$ and simultaneously particle 2 in volume element d^3r_2 centered at $\mathbf{r}_2$. If $\psi(\mathbf{r}_1, \mathbf{r}_2)$ is quadratically integrable, we usually assume the normalization

$$\int\int |\psi(\mathbf{r}_1, \mathbf{r}_2)|^2\, d^3r_1\, d^3r_2 = 1 \tag{15.84}$$

Since ψ is now a function of two different points in space, it can no longer be pictured as a wave in the naïve sense that we found so fruitful in the early chapters. Instead, ψ for two particles may be regarded as a wave in a six-dimensional *configuration space* of the coordinates $\mathbf{r}_1$ and $\mathbf{r}_2$.

The Hamiltonian of the two-particle system (without spin and without external forces) is taken over from classical mechanics and has the general form

$$H = \frac{\mathbf{p}_1^2}{2m_1} + \frac{\mathbf{p}_2^2}{2m_2} + V(|\mathbf{r}_1 - \mathbf{r}_2|) \tag{15.85}$$

In the coordinate representation, this leads to the Schrödinger equation

$$\left[-\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 + V(|\mathbf{r}_1 - \mathbf{r}_2|)\right]\psi(\mathbf{r}_1, \mathbf{r}_2) = E\psi(\mathbf{r}_1, \mathbf{r}_2) \tag{15.86}$$

in configuration space. It is easily verified that the substitutions

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \qquad \mathbf{R} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2} \tag{15.87}$$

transform the Schrödinger equation to

$$-\frac{\hbar^2}{2M}\left(\frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2} + \frac{\partial^2}{\partial Z^2}\right)\psi - \frac{\hbar}{2m_r}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\psi + V(r)\psi = E\psi \tag{15.88}$$

where now $\psi = \psi(\mathbf{r}, \mathbf{R})$ is a function of the ***relative coordinate*** $\mathbf{r}(x, y, z)$ and the coordinate of the center of mass $\mathbf{R}(X, Y, Z)$. In this equation, $M = m_1 + m_2$ is the total mass, and $m_r = m_1m_2/(m_1 + m_2)$ is the ***reduced mass*** of the system.

The new Hamiltonian is a sum

$$H = H_R + H_r \tag{15.89}$$

and each of the two sub-Hamiltonians possesses a complete set of eigenfunctions. Hence, all the eigenfunctions of (15.88) can be obtained by assuming that ψ is a product

$$\psi(\mathbf{r}, \mathbf{R}) = \psi(\mathbf{r})\varphi(\mathbf{R}) \tag{15.90}$$

and the energy a sum, $E = E_R + E_r$, such that

$$\boxed{-\frac{\hbar^2}{2M}\nabla_R^2\varphi(\mathbf{R}) = E_R\varphi(\mathbf{R})} \tag{15.91}$$

and

$$\boxed{-\frac{\hbar^2}{2m_r}\nabla_r^2\psi(\mathbf{r}) + V(r)\psi(\mathbf{r}) = E_r\psi(\mathbf{r})} \tag{15.92}$$

As anticipated, the relative motion of a system of two particles subject to central forces can be treated like a one-particle problem if the reduced mass is used. This justifies the simple reduced one-particle treatment of the diatomic molecule (Section 8.6) and the hydrogen atom (Chapter 12). As pointed out earlier, the most conspicuous manifestation of the reduced mass is the shift that is observed in a comparison of the spectral lines of hydrogen, deuterium, positronium, muonium, and so on. Equation (15.91), whose solutions are plane waves, represents the quantum form of Newton's first law: the total momentum of an isolated system is constant.

The canonical transformation (15.87) could equally well have been made before the quantization. We note that the linear momenta are transformed according to

$$\mathbf{p} = m_r(\mathbf{v}_1 - \mathbf{v}_2) = \frac{m_2}{M}\mathbf{p}_1 - \frac{m_1}{M}\mathbf{p}_2 \qquad \mathbf{P} = \mathbf{p}_1 + \mathbf{p}_2 \tag{15.93}$$

The kinetic energy takes the form

$$T = \frac{\mathbf{p}^2}{2m_r} + \frac{\mathbf{P}^2}{2M} \tag{15.94}$$

and the orbital angular momentum of the system becomes

$$\mathbf{L} = \mathbf{r}_1 \times \mathbf{p}_1 + \mathbf{r}_2 \times \mathbf{p}_2 = \mathbf{r} \times \mathbf{p} + \mathbf{R} \times \mathbf{P} \tag{15.95}$$

If the Hamiltonian is expressed as

$$H = \frac{\mathbf{P}^2}{2M} + \frac{\mathbf{p}^2}{2m_r} + V(r) \tag{15.96}$$

subsequent quantization and use of the coordinate representation lead again to (15.88).

Exercise 15.18. Prove that $\mathbf{r}$, $\mathbf{p}$ and $\mathbf{R}$, $\mathbf{P}$ defined in (15.87) and (15.93) satisfy the commutation relations for conjugate canonical variables. Also show that the Jacobian of the transformation from coordinates $\mathbf{r}_1$, $\mathbf{r}_2$ and $\mathbf{r}$, $\mathbf{R}$ is unity.

It is interesting to ask whether the wave function for a two-particle system is factorable, or separable, and can be written as the product of a function that depends only on the coordinates of particle 1 and a function that depends only on particle 2, such that

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2) \tag{15.97}$$

Obviously, such states are particularly simple to interpret, since we can say that in these cases the two particles are described by their own independent probability

amplitudes. This is sometimes expressed by saying that states like (15.97) do not exhibit *correlations* between the two particles. The fundamental coordinate basis states $|\mathbf{r}_1\mathbf{r}_2\rangle = |\mathbf{r}_2\rangle|\mathbf{r}_2\rangle$ have this special character.

Exercise 15.19. Show that the state of two particles with sharp momenta $\mathbf{p}_1$ and $\mathbf{p}_2$, corresponding to the plane wave function

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \exp\left(\frac{i}{\hbar}\,\mathbf{p}_1 \cdot \mathbf{r}_1\right) \exp\left(\frac{i}{\hbar}\,\mathbf{p}_2 \cdot \mathbf{r}_2\right) \tag{15.98}$$

is also separable when it is transformed by use of (15.87) and (15.93) into $\psi(\mathbf{r}, \mathbf{R})$.

Most two-particle states are not factorable like (15.97). Except for the special case (15.98), wave functions of the type (15.90), which are factorable in relative coordinates, are generally *not* separable with regard to the two particles and are said to be *correlated*. An extreme example of a correlated wave function is afforded by the simple model of two particles confined to staying on the x axis and represented by an idealized amplitude:

$$\psi(x_1, x_2) = \psi(x, X) = \psi(x)\varphi(X) = (2\pi\hbar)^{-1/2}\,\delta(x_1 - x_2 - a) = (2\pi\hbar)^{-1/2}\,\delta(x - a) \tag{15.99}$$

Here a is a positive constant, which may be chosen as large as we please. The state represented by (15.99) corresponds to the two particles being separated precisely and invariably by the distance a, but the probability of finding one of the particles, say particle 1, anywhere regardless of the position of particle 2 is constant and independent of the position x_1. Once a measurement shows particle 2 to be located at coordinate x_2, then particle 1 is certain to be found at position $x_1 = x_2 + a$, and nowhere else. The wave function $\psi(x) = \delta(x - a)$ describes the relative motion of the two particles, and $\varphi(X) = \sqrt{1/2\pi\hbar}$ is a momentum eigenstate of the center-of-mass motion, corresponding to zero total momentum.

Since for any function $\psi(x_1 - x_2)$,

$$\frac{\hbar}{i}\left(\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_2}\right)\psi(x_1 - x_2) = 0$$

we see that (15.99) is an eigenstate of the total momentum, corresponding to eigenvalue zero. Hence, if the momentum of particle 2 is measured and found to have the value p_2, then particle 1 is certain to be found to have the sharp momentum value $p_1 = -p_2$.

Thus, depending on whether the coordinate or the momentum of particle 2 is measured, we are led to conclude that particle 1 after such a measurement is represented by a one-particle state of sharp position (delta function in coordinate space, centered at $x_1 = x_2 + a$) or sharp momentum (plane wave with momentum $-p_2$).

In their famous 1935 paper, Einstein, Podolsky, and Rosen articulated the distress that many physicists felt—and occasionally still feel—about these unequivocal conclusions of quantum mechanics. If we assume that the quantum mechanical amplitude gives a complete (statistical) account of the behavior of a *single* system, it appears that, even when the two particles are arbitrarily far apart (large a), what can be known about the state of particle 1, after a measurement on particle 2 is undertaken, depends on the choice of measurement made on particle 2, such as a coordinate or a momentum measurement. These mysterious long-range correlations between the two widely separated particles and the strange dependence of the expected

behavior of particle 1 on the subjective fickleness of a distant human experimenter, who has no means of interacting with particle 1, seemed to Einstein to signal a violation of the innate sense that the world of physics is governed by *local realism*.

Einstein tried to resolve this conflict by suggesting that quantum mechanical amplitudes pertain only to ensembles of systems, rather than single systems, and provide a correct but incomplete description of physical reality. In principle, a more complete theory, consistent with quantum mechanics, might thus be eventually discovered. But John S. Bell showed that such a program cannot be carried out, as long as the theory is required to be *local*, that is, not afflicted with unaccountable actions-at-a-distance between measuring devices. Any theory built on strict local realism fails to reproduce some predictions of multiparticle quantum mechanics—predictions that have been verified experimentally to a high degree of accuracy. The quest for a return to local realism in physics must thus remain unfulfilled, and we have to accept the existence of quantum correlations between widely separated subsystems. Furthermore, we persist in interpreting the formalism of quantum mechanics as providing complete statistical predictions of the behavior of single systems.

Exercise 15.20. An alternative representation of two-particle states is given in terms of the "mixed" basis states, $|\mathbf{r}, \mathbf{P}\rangle$, where $\mathbf{r}$ is the relative coordinate vector and $\mathbf{P}$ the total momentum. By using the intermediate coordinate basis $|\mathbf{r}, \mathbf{R}\rangle$, derive the transformation coefficients $\langle \mathbf{r}_1\mathbf{r}_2|\mathbf{r}, \mathbf{P}\rangle$. For the correlated state $|\Psi\rangle$ represented by (15.99), show that the wave function in the mixed relative representation is

$$\langle x, P|\Psi\rangle = \delta(x - a)\delta(P) \tag{15.100}$$

Correlated amplitudes like (15.99) or (15.100), which cannot be factored with regard to the two subsystems 1 and 2, are sometimes called *entangled*, a term coined by Schrödinger and illustrated dramatically in his famous cat allegory. Using the basis states (15.79), it is not difficult to construct examples of entangled states for a system composed of two independent subsystems. A general state may be expanded as

$$|\Psi\rangle = \sum_{k,\ell} |A_1'(k)\rangle \otimes |B_2'(\ell)\rangle\langle A_1'(k)B_2'(\ell)|\Psi\rangle \tag{15.101}$$

where for typographic clarity the quantum number labels, k and ℓ, for the basis states of the separate subsystems have been placed as arguments in parentheses instead of as the usual subscripts. The necessary and sufficient condition for the state of the composite system to be factorable with respect to particles 1 and 2 is that the $n_1 \times n_2$ dimensional rectangular matrix of the amplitudes $\langle A_1'(k)B_2'(\ell)|\Psi\rangle$ be expressible in terms of $n_1 + n_2$ complex numbers as

$$\begin{pmatrix} a(1) \\ a(2) \\ \vdots \\ a(n_1) \end{pmatrix} \otimes (b(1)\ b(2)\ .\ .\ b(n_2)) \tag{15.102}$$

The state is entangled if and only if the amplitudes cannot be expressed in the form (15.102).[6]

[6]For a reprint compilation of the key historical papers on entangled states and the puzzling questions they have raised, see Wheeler and Zurek (1983).

Exercise 15.21. Check that the amplitude (15.99) is entangled by making a (Fourier) expansion in terms of momentum eigenfunctions or any other complete set of one-particle basis functions.

5. ***The Density Operator, the Density Matrix, Measurement, and Information.*** The density operator for a state $|\Psi(t)\rangle$ was defined in Eq. (14.19), and its time development was considered in Sections 14.1 and 14.2. As a projection operator for a state $|\Psi\rangle$, the density operator

$$\rho = |\Psi\rangle\langle\Psi| \tag{15.103}$$

contains all relevant information about the state. The density operator is idempotent, since owing to normalization,

$$\rho^2 = |\Psi\rangle\langle\Psi|\Psi\rangle\langle\Psi| = |\Psi\rangle\langle\Psi| = \rho \tag{15.104}$$

Except for an irrelevant phase factor, the state $|\Psi\rangle$ can be recovered from the density operator as the eigenvector of ρ which corresponds to eigenvalue 1. All expectation values can be expressed in terms of the density operator, as can be seen from

$$\langle\Psi|A|\Psi\rangle = \sum_i \langle\Psi|K_i\rangle\langle K_i|A|\Psi\rangle = \sum_i \langle K_i|A|\Psi\rangle\langle\Psi|K_i\rangle = \text{trace}(A\rho) \tag{15.105}$$

For $A = I$, this formula is the normalization condition

$$\langle\Psi|\Psi\rangle = \text{trace } \rho = 1 \tag{15.106}$$

If (15.105) is applied to the projection operator $A = |\Phi\rangle\langle\Phi| = P_\Phi$, an expression for a probability in terms of the density operator results:

$$p_\Phi = |\langle\Phi|\Psi\rangle|^2 = \langle\Psi|\Phi\rangle\langle\Phi|\Psi\rangle = \langle\Psi|P_\Phi|\Psi\rangle = \text{trace}(|\Phi\rangle\langle\Phi|\rho) = \text{trace}(P_\Phi\rho) \tag{15.107}$$

Since this can also be written as

$$p_\Phi = |\langle\Phi|\Psi\rangle|^2 = \langle\Phi|\Psi\rangle\langle\Psi|\Phi\rangle = \langle\Phi|\rho|\Phi\rangle \tag{15.108}$$

we infer that ρ is a positive Hermitian operator and, in particular, that the diagonal elements of any density matrix are nonnegative.

The probabilities p_Φ and $1 - p_\Phi$ are associated with the outcomes of measuring the positive operators P_Φ and $I - P_\Phi$. Somewhat imprecisely we say that p_Φ is the probability of finding the system in state $|\Phi\rangle$, and $1 - p_\Phi$ is the probability of finding the system *not* to be in state $|\Phi\rangle$.

Building on the foundations laid in chapters 4, 9, and 10—especially Sections 4.2 and 10.4—we characterize a *complete orthonormal* quantum measurement by considering a set of n mutually orthogonal, and hence commuting, rank-one (or one-dimensional) projection operators $P_1, P_2, \ldots P_n$ (where n is the dimensionality of the state vector, or Hilbert, space of the system). The completeness is expressed by the closure relation (10.44):

$$\sum_{\ell=1}^{n} P_\ell = I \tag{15.109}$$

Each projection operator corresponds to a different outcome of the proposed measurement. In a specific application of the formalism, $P_\ell = |K_\ell\rangle\langle K_\ell|$ may be the projection operator corresponding to the eigenvalue K_ℓ of an observable (or, more generally, a complete commuting set of observables) symbolized by K. If this is the

case, we regard the values K_ℓ as the possible outcomes of the measurement. The probabilities of finding these outcomes are

$$p_\ell = \text{trace}(P_\ell \rho) \tag{15.110}$$

Equation (15.109) guarantees that the probabilities add up to unity. If the state of the system happens to be an eigenstate $|K_\ell\rangle$, the variance (or uncertainty) of K is zero, and K'_ℓ can be regarded as the sharp value of this observable, akin to a classical observable. Somewhat casually, we call this procedure a "measurement of the set of observable operators K."

Although the complete orthonormal measurements just described stand out prominently, it is possible to generalize the notion of a quantum measurement to include *nonorthogonal* operators. Thus, we assume the existence of a set of r positive Hermitian operators,

$$A_j = w_j P_j \tag{15.111}$$

which are positive multiples of rank-one projection operators P_j. The projection operators in (15.111) are not necessarily mutually orthogonal, nor do they generally commute. With their weights, w_j $(0 \leq w_j \leq 1)$, they are subject to the completeness, or overcompleteness, relation

$$\sum_{j=1}^{r} A_j = \sum_{j=1}^{r} w_j P_j = I \tag{15.112}$$

If the measurement is to be implemented on the system in any arbitrary state, completeness requires that the number of terms r in this sum must be at least as large as n, the dimensionality of the system's state vector space. The probability that the measurement described by the operators A_j yields the jth outcome is given by the formula

$$p_j = \langle \Psi | A_j | \Psi \rangle = \text{trace}(A_j \rho) \tag{15.113}$$

The condition (15.112) ensures that these probabilities add up to unity.

Exercise 15.22. Using the Schwarz inequality, prove that $p_j \leq w_j$.

In this brief discussion of the generalized quantum measurement defined by the set of r operators A_j, and technically referred to as a *probability-operator-valued measure* (or POM), we only emphasize the significance of nonorthogonal terms in (15.113).[7] Suppose that the state of the system is an eigenstate $|\Phi_k\rangle$ of the element $A_k = w_k P_k$ of the POM, corresponding to eigenvalue w_k. The probability of the jth outcome of the measurement defined by the POM is

$$p_j = \langle \Phi_k | A_j | \Phi_k \rangle = w_j \langle \Phi_k | P_j | \Phi_k \rangle = w_j \,\text{trace}(P_j P_k) \tag{15.114}$$

This shows that the probability of the kth outcome $(j = k)$ may be less than unity, and that of the other outcomes $(j \neq k)$ may not be zero (as would be expected if the states k and j were orthogonal). There is a quantum mechanical *fuzziness* inherent in nonorthogonal measurements. In these, unlike complete orthogonal measure-

[7]See Peres (1995) for more detail and many references to books and articles on quantum measurement and quantum information. For a clear discussion of entropy in the context of coding theory, see Schumacher (1995).

ments, p_k can generally not be interpreted as the probability of finding the system to be *in* state $|\Phi_k\rangle$, and this state cannot be associated unambiguously with just one of the possible outcomes of such a measurement. Nonorthogonal quantum measurements exhibit more peculiarly quantal features than orthogonal measurements of observables, whose eigenstates can be unambiguously associated with sharp values of the corresponding physical quantities, in a manner reminiscent of classical physics.

As an illustration of a POM for a one-dimensional system with an infinite-dimensional Hilbert space, we draw attention to the closure relation (10.111) for the coherent states of a harmonic oscillator:

$$\frac{1}{\pi}\int |\alpha\rangle\, d^2\alpha\, \langle\alpha| = I \tag{15.115}$$

which is precisely in the form (15.112), applied to a continuously variable outcome, identified by the complex number α. If the system is known to be in the kth energy eigenstate of the harmonic oscillator, the probability density (per unit area in the complex α plane of Figure 10.1) for outcome α in a measurement of the nonorthogonal POM defined by the coherent states, $A_\alpha = (1/\pi)|\alpha\rangle\langle\alpha|$, is

$$p_\alpha(k) = \frac{1}{\pi k!}\,|\alpha|^{2k} e^{-|\alpha|^2} \tag{15.116}$$

Except for the factor $w_\alpha = 1/\pi$, this is the same as (10.110).

So far, in this section, the density operator has merely served as an alternative to describing a quantum state by a vector in Hilbert space. It would be possible to formulate all of quantum mechanics exclusively in terms of density operators and their matrix elements, but the required mathematical techniques are generally less familiar than those pertaining to amplitudes and wave functions. (However, effective approximation schemes for complex many-particle quantum systems have been invented using density operator and density matrix methods.[8])

In Chapter 16, we will illustrate the use of the density operator and its representative, the density matrix, for the simple case of a spin one-half system whose spatial coordinate degrees of freedom are irrelevant and can be suppressed. We will find that the full benefit of using the density matrix accrues only when it is applied to a *statistical ensemble* of imaginary replicas of the system in the tradition of Gibbs, thereby creating a *mixture* of different quantum states. A mixture can be visualized as the set of *probabilities*, or relative frequencies, p_i, with which N different quantum states $|\Psi_i\rangle$ or the corresponding density operators $\rho_i = |\Psi_i\rangle\langle\Psi_i|$ occur in the ensemble denoted by $\mathscr{E}$. We must require that

$$\sum_{i=1}^{N} p_i = 1 \tag{15.117}$$

but the states $|\Psi_i\rangle$ generally need not be orthonormal. Equations (15.105), (15.107), and (15.108) show that probabilities and expectation values for quantum states depend on the density operator linearly, whereas they depend on the state $|\Psi\rangle$ quad-

[8]See, for example, Parr and Yang (1989).

ratically. It follows that all statistical predictions for the *ensemble* can be expressed in terms of the generalized *density operator* of the system,

$$\boxed{\rho = \sum_{i=1}^{N} p_i \rho_i = \sum_{i=1}^{N} p_i |\Psi_i\rangle\langle\Psi_i|} \tag{15.118}$$

by the universal formula for the average value of an operator A

$$\boxed{\langle A\rangle = \sum_{i=1}^{N} p_i \langle\Psi_i|A|\Psi_i\rangle = \sum_{i=1}^{N} p_i \operatorname{trace}(A\rho_i) = \operatorname{trace}(A\rho)} \tag{15.119}$$

The density operator ρ, like its constituents, ρ_i, is a positive Hermitian operator. If all probabilities p_i except one vanish, the density operator (15.118) reduces to the idempotent operator (15.103). It is then said to describe a *pure state*. Otherwise it represents a *mixed state*. Since in applications one usually employs a particular basis to represent states and operators, the same symbol ρ is often also used to denote the corresponding *density matrix*.

All density operator relations that we derived for pure states at the beginning of this section carry over to the generalized density operator for a mixture, except for (15.104), which is quadratic in ρ and characterizes a pure state or one-dimensional (rank one) projection operator. Instead, owing to the positive definiteness of the density operator, we have in general,

$$\boxed{0 \leq \operatorname{trace}\rho^2 \leq (\operatorname{trace}\rho)^2 = 1 \quad \text{and} \quad \rho_{ii}\rho_{jj} \geq |\rho_{ij}|^2} \tag{15.120}$$

We have constructed the density operator ρ for a mixture from the assumed *a priori* knowledge of the N pure states ρ_i representing the ensemble $\mathscr{E}$ and the corresponding probabilities p_i. It is not possible to reverse this procedure and to infer the composition of a mixture uniquely. A given density operator ρ is compatible with many (generally, infinitely many) different ways of mixing pure states. We will presently quantify the information loss that is incurred in the mixing process. However, an exceptionally useful decomposition is always provided by a complete set of orthonormal eigenstates $|\bar{p}_j\rangle$ of the Hermitian density operator ρ and its eigenvalues $\bar{p}_j$

$$\rho = \sum_{j=1}^{n} \bar{p}_j |\bar{p}_j\rangle\langle\bar{p}_j| \tag{15.121}$$

where n is the dimensionality of the Hilbert space, and

$$\sum_{j=1}^{n} |\bar{p}_j\rangle\langle\bar{p}_j| = I \tag{15.122}$$

Some of the eigenvalues $\bar{p}_j$ may be zero, and there is a certain amount of arbitrariness in the choice of eigenvectors, if eigenvalues are repeated. In particular, if all eigenvalues of ρ are equal to $1/n$, the density operator is proportional to the identity, $\rho = (1/n)I$, and the mixture is as random as possible. Borrowing a term from the physics of spatially orientable systems, a completely mixed state for which $\rho = (1/n)I$ is said to be *unpolarized*. Any POM composed of positive operators

$A_k = w_k P_k$, which resolves the identity according to (15.112), can be employed to represent an unpolarized ensemble by writing

$$\rho = \frac{1}{n} \sum_{j=1}^{r} w_j P_j \tag{15.123}$$

Exercise 15.23. Prove the inequalities (15.120). *Hint:* Trace inequalities are most easily proved by using the eigenstates of the density operator as a basis. For the second inequality, maximize the probability of finding the system in a superposition state

$$|\Psi\rangle = c_i|\Psi_i\rangle + c_j|\Psi_j\rangle$$

Exercise 15.24. If the state of a quantum system is given by a density operator

$$\rho = p_1|\Psi_1\rangle\langle\Psi_1| + p_2|\Psi_2\rangle\langle\Psi_2| \tag{15.124}$$

where $|\Psi_{1,2}\rangle$ are two nonorthogonal normalized state vectors, show that the eigenvalues of the density operator are

$$\bar{p}_{1,2} = \frac{1 \pm \sqrt{1 - 4p_1p_2(1 - |\langle\Psi_1|\Psi_2\rangle|^2)}}{2} \tag{15.125}$$

If a mixed state with density operator ρ is defined by a given probability distribution of N known pure states $\rho_1, \rho_2, \ldots \rho_N$ with probabilities $p_1, p_2, \ldots p_N$, our incomplete knowledge of the state can be quantified in terms of the information and entropy concepts that are introduced in Section 2 of the Appendix. The *Shannon mixing entropy* (A.43) for this ensemble $\mathscr{E}$, denoted by $H(\mathscr{E})$, is

$$H(\mathscr{E}) = -\sum_{i=1}^{N} p_i \ln p_i \tag{15.126}$$

We have chosen to express the entropy in terms of the *natural* logarithm, so that the *nat* is the unit of $H(\mathscr{E})$.

The quantity $H(\mathscr{E})$ is a measure of our ignorance of the state. A large mixing entropy $H(\mathscr{E})$ implies a highly randomized ensemble. If the state of the system is pure ($\rho^2 = \rho$), the information is maximal and the mixing entropy is $H(\mathscr{E}) = 0$. The information about the state is complete. If, on the other hand, all ρ_i are equally probable,

$$p_1 = p_2 = \ldots = p_N = \frac{1}{N} \tag{15.127}$$

and the mixing entropy is $H(\mathscr{E}) = \ln N$ nats.

In quantum information theory one investigates how, given a set of *a priori* probabilities about a quantum state, our ignorance and the entropy can be reduced, or information gained, by performing measurements on an ensemble $\mathscr{E}$. The decomposition (15.121) of a given density operator in terms of its complete set of orthonormal pure eigenstates occupies a special place among the probability distributions compatible with ρ. Its mixing entropy is denoted by $S(\rho)$ and defined as

$$\boxed{S(\rho) = -\sum_{i=1}^{n} \bar{p}_i \ln \bar{p}_i = -\sum_{i=1}^{n} \langle\bar{p}_i|\rho \ln \rho|\bar{p}_i\rangle = -\text{trace}(\rho \ln \rho)} \tag{15.128}$$

Here, the function $\ln \rho$ of the density operator is understood to be defined as in Eq. (10.30). Among all the different entropies that can be usefully defined, $S(\rho)$ is singled out and referred to as the *von Neumann entropy*. It can be shown to be the smallest of all mixing entropies (15.126) for a given density operator:

$$H(\mathscr{E}) = -\sum_{i=1}^{N} p_i \ln p_i \geq -\sum_{j=1}^{n} \bar{p}_j \ln \bar{p}_j = S(\rho) \tag{15.129}$$

Thus, the ensemble composed of the orthonormal eigenstates of the density operator, the *eigen-ensemble*, is the least random of all the possible decompositions of ρ. In Chapter 16, this extremal property of the von Neumann entropy will be further demonstrated by several examples.

Exercise 15.25. If an ensemble $\mathscr{E}$ consists of an equal-probability mixture of two nonorthogonal (but normalized) states $|\Psi_1\rangle$ and $|\Psi_2\rangle$ with overlap $C = \langle\Psi_1|\Psi_2\rangle$, evaluate the Shannon mixing entropy $H(\mathscr{E})$ and the von Neumann entropy, $S(\rho)$. Compare the latter with the former as $|C|$ varies between 0 and 1. What happens as $C \to 0$?

Exercise 15.26. A given ensemble $\mathscr{E}$ consists of a mixture of two equiprobable orthonormal states $|\Psi_1\rangle$ and $|\Psi_2\rangle$ and a third normalized state $|\Psi_3\rangle$, which is itself a superposition (not a mixture!) $|\Psi_3\rangle = c_1|\Psi_1\rangle + c_2|\Psi_2\rangle$, so that the density operator is

$$\rho = p|\Psi_1\rangle\langle\Psi_1| + p|\Psi_2\rangle\langle\Psi_2| + (1-2p)|\Psi_3\rangle\langle\Psi_3| \qquad (0 \leq p \leq 1/2) \tag{15.130}$$

Work out the eigenprobabilities of ρ and the Shannon and von Neumann entropies. Discuss their dependence on the mixing probability p and on the amplitudes $c_{1,2}$.

Entropy can be defined for any probability distribution. To gauge the predictability of the outcome of a measurement of an observable K on a system with density operator ρ, we define the *outcome entropy*:

$$H(K) = -\sum_{j=1}^{n} p(K_j) \ln p(K_j) \tag{15.131}$$

Since the probabilities

$$p(K_j) = \text{trace}(|K_j\rangle\langle K_j|\rho) \tag{15.132}$$

can be calculated directly from the density operator, the value of the outcome entropy, $H(K)$, is independent of the particular ensemble $\mathscr{E}$ which represents ρ. Again, the von Neumann entropy stands out, because one can prove that

$$H(K) = -\sum_{j=1}^{n} p(K_j) \ln p(K_j) \geq -\sum_{i=1}^{n} \bar{p}_i \ln \bar{p}_i = S(\rho) \tag{15.133}$$

For the special case of a pure quantum state $|\Psi\rangle$, or $\rho = |\Psi\rangle\langle\Psi|$ and $S(\rho) = 0$, the relation (15.133) makes the trite but true statement that there is generally an inevitable loss of information, if we know only the probabilities $|\langle K_j|\Psi\rangle|^2$ for measuring the observable K. We are missing the valuable information stored in the relative phases of the amplitudes.

The fundamental significance of the von Neumann entropy $S(\rho)$ should now be apparent. Thermodynamic considerations show that, multiplied by the Boltzmann

constant, k, the von Neumann entropy is also the form of the entropy whose maximum, subject to certain constraints, yields, according to the second law of thermodynamics, the equilibrium distributions for quantum statistical mechanics. We will implement this principle in Section 22.5, after an introduction to the quantum physics of identical particles.

To complete the discussion of the density operator, we must give an account of its time evolution. If the density operator (15.118) for the system, with a Hermitian Hamiltonian H, is given at some initial time t_0, each constituent pure-state density operator ρ_i develops according to the equation of motion (14.20). Owing to the linear dependence of ρ on the components ρ_i, and the linearity and homogeneity of (14.20), the density operator ρ develops in the Schrödinger picture according to the dynamical equation

$$\boxed{i\hbar \frac{\partial \rho}{\partial t} = [H, \rho]} \tag{15.134}$$

If we assume that

$$i\hbar \frac{\partial \rho^{n-1}}{\partial t} = [H, \rho^{n-1}]$$

for any positive integer n, it follows by induction that

$$i\hbar \frac{\partial \rho^n}{\partial t} = i\hbar \frac{\partial \rho}{\partial t} \rho^{n-1} + i\hbar \rho \frac{\partial \rho^{n-1}}{\partial t} = [H, \rho]\rho^{n-1} + \rho[H, \rho^{n-1}] = [H, \rho^n]$$

Hence, (15.134) can be generalized for any analytic function $f(\rho)$ of the density operator:

$$i\hbar \frac{\partial f(\rho)}{\partial t} = [H, f(\rho)] \tag{15.135}$$

The equation of motion (14.18) for the expectation value of an operator A, which may be time-dependent, can be equally well applied to a mixed state:

$$\boxed{i\hbar \frac{d\langle A \rangle}{dt} = \langle [A, H] \rangle + i\hbar \left\langle \frac{\partial A}{\partial t} \right\rangle = \text{trace}(\rho[A, H]) + i\hbar \,\text{trace}\left(\rho \frac{\partial A}{\partial t} \right)} \tag{15.136}$$

By substituting $A = f(\rho)$ in (15.136), and using (15.135), if follows that the expectation value of any function of the density operator is constant in time:

$$\boxed{\frac{d}{dt} \langle f(\rho) \rangle = 0} \tag{15.137}$$

Exercise 15.27. Give a direct proof that for a general mixed state, $\langle f(\rho) \rangle = \text{trace}(\rho f)$ is constant in time, by noting that the density operator evolves in time by a unitary transformation, $\rho(t) = T\rho(t_0)T^\dagger$.

Exercise 15.28. Prove that

$$\text{trace}(\rho[f(\rho), H]) = \text{trace}(f(\rho)[H, \rho]) = 0 \tag{15.138}$$

and that consequently,

$$\left\langle \frac{\partial f(\rho)}{\partial t} \right\rangle = 0 \tag{15.139}$$

As an important corollary, it follows that the von Neumann entropy, which is the mean of the density operator function ln ρ, remains constant as the system evolves in time:

$$\frac{dS(\rho)}{dt} = -\frac{d}{dt}\,\text{trace}(\rho \ln \rho) = 0 \tag{15.140}$$

This exact conclusion is not inconsistent with the familiar property of entropy in statistical thermodynamics as a quantity that increases in time during the irreversible approach to equilibrium, because it holds only under the precise conditions that we have specified, including the idealization that the probability distribution of the statistical mixture representing the ensemble is fixed in time.

The sketchy introduction to the concepts of the quantum theory of measurement and information presented in this section will be supplemented by concrete examples in the next chapter in the context of quantum mechanics in a vector space of only two dimensions. As we apply the results obtained in this section, we should remember that common terms like ''measurement'' and ''information'' are being used here with a specific technical meaning. In particular, this is not the place for a detailed analysis of real experimental measurements and their relation to the theoretical framework. We merely note that, in the information theoretic view of quantum mechanics, the probabilities and the related density operators and entropies, which are employed to assess the properties of quantum states and the outcomes of measurement, provide a coherent and consistent basis for understanding and interpreting the theory.

Problems

1. For a system that is characterized by the coordinate $\mathbf{r}$ and the conjugate momentum $\mathbf{p}$, show that the expectation value of an operator F can be expressed in terms of the Wigner distribution $W(\mathbf{r}', \mathbf{p}')$ as

$$\langle F \rangle = \langle \Psi | F | \Psi \rangle = \int \int F_W(\mathbf{r}', \mathbf{p}') W(\mathbf{r}', \mathbf{p}')\, d^3r'\, d^3p'$$

where

$$F_W(\mathbf{r}', \mathbf{p}') = \int e^{(i/\hbar)\mathbf{p}'\cdot\mathbf{r}''} \left\langle \mathbf{r}' - \frac{\mathbf{r}''}{2} \middle| F \middle| \mathbf{r}' + \frac{\mathbf{r}''}{2} \right\rangle d^3r''$$

and where the function $W(\mathbf{r}', \mathbf{p}')$ is defined in Problem 5 in Chapter 3. Show[9] that for the special cases $F = f(\mathbf{r})$ and $F = g(\mathbf{p})$ these formulas reduce to those obtained in Problems 5 and 6 in Chapter 3, that is, $F_W(\mathbf{r}') = f(\mathbf{r}')$ and $F_W(\mathbf{p}') = g(\mathbf{p}')$.

[9]Recall that in expressions involving the Wigner distribution $\mathbf{r}$ and $\mathbf{p}$ stand for operators, and the primed variables are real-number variables.

2. Show that the probability current density at $\mathbf{r}_0$ is obtained with

$$\mathbf{j}_W(\mathbf{r}_0; \mathbf{r}', \mathbf{p}') = \frac{\mathbf{p}'}{2m}\,\delta(\mathbf{r}' - \mathbf{r}_0)$$

so that the current density at $\mathbf{r}_0$ is

$$\mathbf{j}(\mathbf{r}_0) = \int W(\mathbf{r}_0, \mathbf{p}')\,\frac{\mathbf{p}'}{2m}\,d^3p'$$

3. Derive the Wigner distribution function for an isotropic harmonic oscillator in the ground state.

4. Prove that for a pure state the density operator $|\Psi\rangle\langle\Psi|$ is represented in the Wigner distribution formalism by

$$\rho_W(\mathbf{r}', \mathbf{p}') = (2\pi\hbar)^3 W(\mathbf{r}', \mathbf{p}')$$

Check that this simple result is in accord with the normalization condition $\langle\rho\rangle = 1$ for the density operator.

5. For a free particle, derive the equation of motion for the Wigner distribution

$$\frac{\partial W(\mathbf{r}', \mathbf{p}', t)}{\partial t} + \frac{\mathbf{p}'}{m} \cdot \nabla_{\mathbf{r}'} W(\mathbf{r}', \mathbf{p}', t) = 0$$

from the time-dependent Schrödinger equation. What does the equation of motion for W for a particle in a potential $V(\mathbf{r})$ look like?

6. Two particles of equal mass are constrained to move on a straight line in a common harmonic oscillator potential and are coupled by a force that depends only on the distance between the particles. Construct the Schrödinger equation for the system and transform it into a separable equation by using relative coordinates and the coordinates of the center of mass. Show that the same equation is obtained by first constructing a separable classical Hamiltonian and subjecting it to canonical quantization.

7. Assuming that the two particles of the preceding problem are coupled by an elastic force (proportional to the displacement), obtain the eigenvalues and eigenfunctions of the Schrödinger equation and show that the eigenfunctions are either symmetric or antisymmetric with respect to an interchange of the two particles.

CHAPTER 16

The Spin

The spin (one-half) of a particle or atom or nucleus provides an opportunity to study quantum dynamics in a state vector space with only two dimensions. All laws and equations can be expressed in terms of two components and 2×2 matrices. Moreover, we gain insight into the effect of rotations on quantum states. The lessons learned here are transferable to the general theory of rotations in Chapter 17. Polarization and resonance in static and time-varying fields are characteristic spin features described by the theory and observed in experiments. The spin also lends itself to an explicit and relatively transparent discussion of the interpretation of quantum mechanics and its amplitudes, density matrices, and probabilities. In the quantum mechanics of two-dimensional complex vector spaces, it is possible to concentrate on the intriguing features of the theory, untroubled by mathematical complexities.

1. ***Intrinsic Angular Momentum and the Polarization of ψ Waves.*** In Chapter 15, we were concerned with the quantum description of a particle as a mass point, and it was assumed that the state of the particle can be completely specified by giving the wave function ψ as a function of the spatial coordinates x, y, z, with no other degrees of freedom. The three dynamical variables were postulated to constitute a complete set. Alternatively and equivalently, the linear momentum components p_x, p_y, p_z also form a complete set of dynamical variables, since $\phi(\mathbf{p})$ contains just as much information about the state as $\psi(\mathbf{r})$. The Fourier integral links the two equivalent descriptions and allows us to calculate ϕ from ψ, and vice versa.

It is important to stress here that completeness of a set of dynamical variables is to be understood with reference to a model of the physical situation, but it would be presumptuous and quite unsafe to attribute completeness in any other sense to the mathematical description of a physical system. For no matter how complete the description of a state may seem today, the history of physics teaches us that sooner of later new experimental facts will come to light which will require us to improve and extend the model to give a more detailed and usually more complete description.

Thus, the wave mechanical description of the preceding chapters is complete with reference to the simple model of a point particle in a given external field, and it is remarkable how many fundamental problems of atomic, molecular, and nuclear physics can be solved with such a gross picture. Yet this achievement must not blind us to the fact that this simple model is incapable of accounting for many of the finer details. In particle physics and in many problems in condensed-matter physics, it is inadequate even for a first orientation.

A whole host of quantum properties of matter can be understood on the basis of the discovery that many particles, including electrons, protons, neutrons, quarks, and neutrinos, are not sufficiently described by the model of a point particle whose wave function as a function of position or momentum exhausts its dynamical properties. Rather, all the empirical evidence points to the need for attributing an *angular momentum* or *spin* to these particles in addition to their orbital angular momentum, and, associated with this, a *magnetic moment*. For composite particles like protons

and neutrons, these properties can be understood in terms of their internal quark structure, but leptons like electrons and muons appear to be elementary point-like constituents of matter, yet nevertheless possess intrinsic angular momentum.

What is the most direct evidence for the spin and the intrinsic magnetic moment? Although it was not realized at the time, Stern and Gerlach first measured the intrinsic magnetic moment in experiments[1] whose basic features are interesting here because they illustrate a number of concepts important in interpreting quantum mechanics. The particles, which may be entire atoms or molecules whose magnetic moment $\boldsymbol{\mu}$ is to be measured, are sent through a nonuniform magnetic field $\mathbf{B}$. They are deflected by a force which according to classical physics is given by

$$\mathbf{F} = \nabla(\boldsymbol{\mu} \cdot \mathbf{B}) \tag{16.1}$$

and they precess around the field under the influence of the torque $\boldsymbol{\tau} = \boldsymbol{\mu} \times \mathbf{B}$. The arrangement is such that in the region through which the beam passes the direction of $\mathbf{B}$ varies only slowly, but its magnitude B is strongly dependent on position. Hence, the projection μ_B of $\boldsymbol{\mu}$ in the direction $\mathbf{B}$ remains sensibly unchanged, and we have approximately

$$\mathbf{F} \approx \mu_B \nabla B \tag{16.2}$$

By measuring the deflection, through inspection of the trace that the beam deposits on the screen, we can determine this force, hence the component of the magnetic moment in the direction of $\mathbf{B}$. Figure 16.1 shows the outline of such an experiment.

The results of these experiments were striking. Classically, we would have expected a single continuous trace, corresponding to values of μ_B, ranging from $-\mu$ to $+\mu$. Instead, observations showed a number of *distinct traces*, giving clear proof of the discrete quantum nature of the magnetic moment. Since the vector $\boldsymbol{\mu}$ seemed to be capable of assuming only certain directions in space, it became customary to speak of *space quantization*.

Stern and Gerlach also obtained quantitative results. They found that the values of μ_B appeared to range in equal steps from a minimum, $-\mu$, to a maximum, μ. The value μ of the maximum projection of $\boldsymbol{\mu}$ is conventionally regarded as *the* magnetic moment of a particle.

In order to interpret these results, we recall Ampère's hypothesis that the mag-

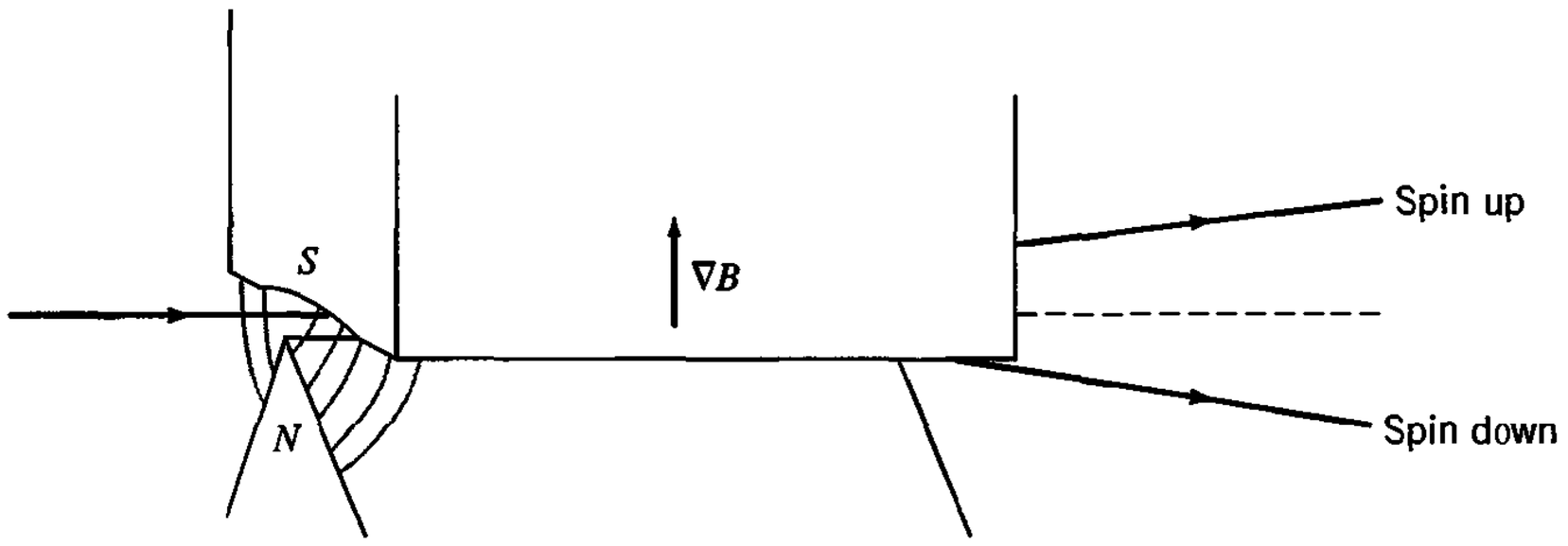


Figure 16.1. Measurement of the vertical component of the magnetic moment of atoms in an inhomogeneous magnetic field (Stern-Gerlach experiment). Silver atoms incident from the left produce two distinct traces corresponding to "spin up" and "spin down."

[1]See Cagnac and Pebay-Peyroula (1971), p. 239.

netic properties of matter are attributable to electric currents of one form or another. Thus, the circulating currents due to electrons (of charge $-e$ and mass m_e) in atoms produce an orbital angular momentum $\mathbf{L}$ and a magnetic moment $\boldsymbol{\mu}$ connected by the classical relation,

$$\boldsymbol{\mu} = -\frac{e}{2m_e c}\mathbf{L} \tag{16.3}$$

which, being a simple proportionality of two vectors, is expected to survive in quantum mechanics also. Since any component of $\mathbf{L}$ has $2\ell + 1$ eigenvalues, we expect the projection of $\boldsymbol{\mu}$ in a fixed direction, such as $\mathbf{B}$, also to possess $2\ell + 1$ distinct eigenvalues and to be expressible as

$$\mu_B = -\frac{e\hbar}{2m_e c}m = -\beta_0 m \tag{16.4}$$

where the magnetic quantum number m can assume the values $-\ell$, $-\ell + 1$, $\ell - 1$, ℓ. The *Bohr magneton* β_0 is defined as

$$\beta_0 = \frac{e\hbar}{2m_e c} \tag{16.5}$$

and has[2] the value 9.27401×10^{-24} J/T $= 9.27401 \times 10^{-21}$ erg/gauss $= 5.78838 \times 10^{-5}$ eV/T.

Since ℓ is an integer, we expect an *odd* number $(2\ell + 1)$ of traces in the Stern-Gerlach experiment. It is well known that the classical experiment with a beam of silver atoms, passing through an inhomogeneous magnetic field, yielded instead *two* traces, i.e., an even number, corresponding to

$$\mu_B = \pm\frac{e\hbar}{2m_e c} = \pm\beta_0$$

We may ask if the semiclassical arguments used above are valid when we contend with quantum phenomena. Equation (16.2) is purely classical, and we may wonder if its application to quantized magnetic moments has not led us astray. The answer to these questions is that like most experiments the Stern-Gerlach experiment has components that are properly and correctly described by the laws of classical physics. For these are the laws that govern the experiences of our senses by which we ultimately, if indirectly, make contact with what happens inside atoms and nuclei. If the particles that the inhomogeneous field in a Stern-Gerlach experiment deflects are sufficiently massive, their motion can be described by wave packets that spread very slowly; hence, this motion can be approximated by a classical description.

The correct interpretation was given to the Stern-Gerlach observations only after Goudsmit and Uhlenbeck were led by a wealth of spectroscopic evidence to hypothesize the existence of an electron spin and intrinsic magnetic moment. If one assumes that the electron is in an S state in the Ag atom, there can be no contribution to the magnetic moment from the orbital motion, and $\mu = e\hbar/2m_e c$ measures the maximum value of a component of the *intrinsic* magnetic moment. Unlike a magnetic moment arising from charged particles moving in spatial orbits, this magnetic moment may be assumed to have only *two* projections, $\mu_B = \pm\beta_0$. According to the

[2]Cohen and Taylor (1996). This useful compilation of fundamental constants is updated and appears annually in the August issue of *Physics Today*.

Goudsmit-Uhlenbeck hypothesis, we envisage the electron to be a point charge with a finite magnetic dipole moment, the projection of which can take on only two discrete values. It is now known that the electron magnetic moment differs very slightly from the Bohr magneton and has the value 1.001 159 652 193 β_0, owing to a small quantum electrodynamic correction. The muon magnetic moment similarly differs by a minute amount from its naïvely expected value $m_e\beta_0/m_{\text{muon}}$.

Goudsmit and Uhlenbeck also postulated that the electron has an intrinsic angular momentum (*spin*), but this quantity is not nearly as easy to measure directly as the magnetic moment. Without appealing to the original justification for the electron spin, which was based on experience with atomic spectra, we can marshal a fundamental argument for the assumption that an electron must have intrinsic angular momentum: From experiment we know that an electron, whether free or bound in an atom, does have a magnetic moment. Unless the atomic electron, moving in the electric field of the nucleus, possesses intrinsic angular momentum, conservation of angular momentum cannot be maintained for an isolated system such as an atom.

To elaborate on this point, we note that, just as a moving charge is subject to a force in a magnetic field, so a moving magnetic moment, such as the intrinsic electron moment is envisaged to be, is also acted on by forces in an electric field. The potential energy associated with these forces is

$$\boldsymbol{\mu} \cdot \frac{\mathbf{v}}{c} \times \mathbf{E}$$

which, for a central field [$\mathbf{E} = f(r)\mathbf{r}$], is proportional to $\boldsymbol{\mu} \cdot \mathbf{v} \times \mathbf{r}$, or to

$$\boldsymbol{\mu} \cdot \mathbf{r} \times \mathbf{v} = \boldsymbol{\mu} \cdot \mathbf{L} \tag{16.6}$$

The factor of proportionality depends only on the radial coordinate r. If the Hamiltonian operator contains, in addition to the central potential, an interaction term like (16.6) proportional to $\boldsymbol{\mu} \cdot \mathbf{L}$, the energy of the electron depends on the relative orientation of the magnetic moment and the orbital angular momentum. It is apparent that $\mathbf{L}$, whose components do not commute, can then no longer be a constant of the motion. Conservation of angular momentum can be restored only if the electron can participate in the transfer of angular momentum by virtue of an intrinsic spin associated with the intrinsic magnetic moment $\boldsymbol{\mu}$.

We conclude that the magnetic moment of a system must always be associated with an angular momentum (see Section 16.4). For leptons with no internal structure, the relativistic Dirac theory of the electron in Chapter 24 will provide us with a deeper understanding of these properties. However, at a comparatively unsophisticated level in describing interactions that are too weak to disturb the internal structure of the particles appreciably, we may treat mass, charge, intrinsic angular momentum, and magnetic moment as given fixed properties.

As the presence of $\hbar$ in the formula $\mu = e\hbar/2mc$ shows, the intrinsic spin and the corresponding magnetic moment are quantum effects signaling an orientation in space, and we must now find an appropriate way of including this in the theory. Wave mechanics was developed in Chapter 2 with relative ease on the basis of the correspondence between the momentum of a particle and its wavelength. This suggests that, in our effort to construct a theory that includes the spin, we should be aided by first determining what *wave* feature corresponds to this physical property.

A scattering experiment can be designed to bring out the directional properties of waves. If a homogeneous beam of particles described by a *scalar* wave function $\psi(x, y, z, t)$, such as alpha particles or pions, is incident on a scatterer, and if the

target is composed of spherically symmetric or randomly oriented scattering centers (atoms or nuclei), as discussed in detail in Chapter 13, we expect the scattered intensity to depend on the scattering angle θ but *not* on the azimuthal angle φ that defines the orientation of the scattering plane with respect to some fixed reference plane. In actual fact, if the beam in such experiments with electrons, protons, neutrons, or muons is suitably prepared, a marked azimuthal asymmetry is observed, including a *right-left asymmetry* between particles scattered at the same angle θ but on opposite sides of the target. It is empirically found that the scattered intensity can be represented by the simple formula

$$\boxed{I = a(\theta) + b(\theta) \cos \varphi} \tag{16.7}$$

provided that a suitable direction is chosen as the origin of the angle measure φ. The simplest explanation of this observation is that ψ representing an electron is not a scalar field, and that *ψ waves can be polarized.* (Here, "electron" is used as a generic term. Polarization experiments are frequently conducted with protons, neutrons, atoms, nuclei, and other particles.)

Figure 16.2 shows the essential features of one particular polarization experiment. A beam I_0 of unpolarized electrons is incident on an unpolarized scatterer A. The particles, scattered at an angle θ_1 from the direction of incidence, are scattered again through the angle θ_2 by a second unpolarized scatterer B, and the intensity of the so-called second scattered particles is measured as a function of the azimuthal angle φ, which is the angle between the first and second planes of scattering. Owing to the axial symmetry with respect to the z axis, the intensities I_1 and I_1' are equal, but $I_R \neq I_L$, and the azimuthal dependence of the second scattered particle beam can be fitted by an expression of the form (16.7).

It is instructive to compare these conclusions with the results of the analogous double scattering experiment for initially unpolarized X rays. With the same basic arrangement as in Figure 16.2, no right-left asymmetry of X rays is observed, but the polarization manifests itself in a $\cos^2 \varphi$ dependence of the second scattered intensity. Since intensities are calculated as squares of amplitudes, such a behavior suggests that electromagnetic waves may be represented by a *vector* field that is transverse and whose projection on the scattering plane, when squared, determines the intensity.

The presence of a $\cos \varphi$, instead of a $\cos^2 \varphi$, term precludes a similar conclusion for the electron waves and shows that, if their polarization can be represented by a vector, the intensity must depend on this vector *linearly* and not quadratically. Hence, the wave function, whose *square* is related to the intensity, is not itself a vectorial quantity, and the polarization vector ($\mathbf{P}$) will have to be calculated from it indirectly.

In summary, the polarization experiments suggest that the wave must be represented by a wave function, which under spatial rotations transforms neither as a scalar nor as a vector, but in a more complicated way. On the other hand, the interpretation of the Stern-Gerlach experiment requires that, in addition to x, y, z, the wave function must depend on at least one other dynamical variable to permit the description of a magnetic moment and intrinsic angular momentum which the electron possesses. Since both the polarization of the waves and the lining up of the particle spins are aspects of a spatial *orientation* of the electron, whether it be wave

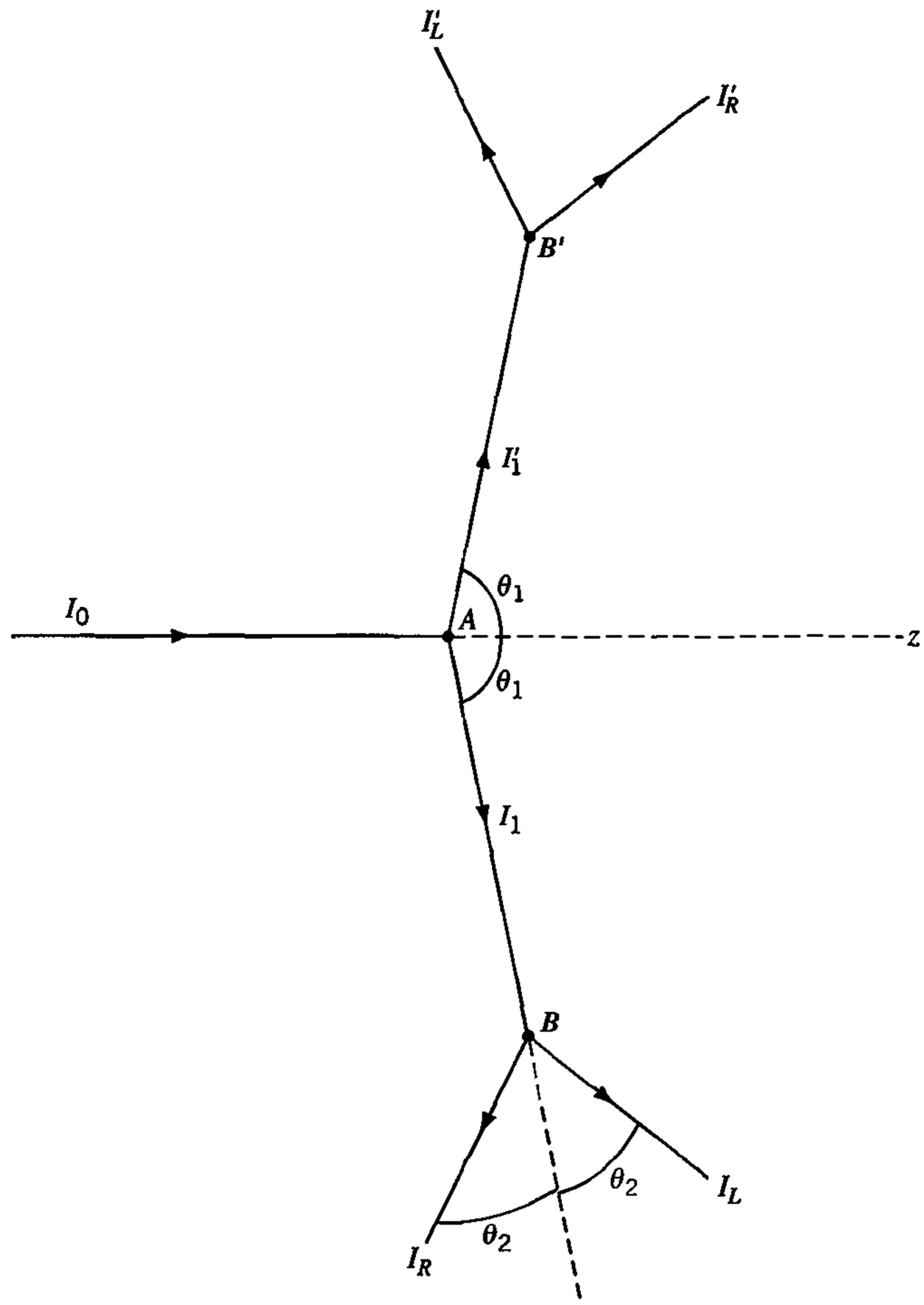


Figure 16.2. Geometry of a double scattering experiment. The first plane of scattering at A is formed by I_0, I, and I'_1, in the plane of the figure. The first scattering polarizes the beam, and the second scattering at B and B' analyzes the degree of polarization. The second plane of scattering, formed by I_1, I_R, and I_L, need not coincide with the first plane of scattering. The angle between the two planes is φ, but is not shown.

or particle, it is not far-fetched to suppose that the same extension of the formalism of wave mechanics may account for both observations.

Similarly, we will see in Chapter 23 that the vector properties of electromagnetic waves are closely related to the intrinsic angular momentum (spin 1) of photons.

2. ***The Quantum Mechanical Description of the Spin.*** Although the formalism of quantum mechanics, which we developed in Chapters 9, 10, and 14, is of great generality, we have so far implemented it only for the nonrelativistic theory of single particles that have zero spin or whose spin is irrelevant under the given physical circumstances. To complement the set of continuously variable fundamental observables x, y, z for an electron, we now add a fourth *discrete* observable that is assumed to be independent of all the coordinate (and momentum) operators and commutes

with them. We denote its (eigen-)values by σ. This *spin variable*, which is capable of taking on only two distinct values, is given a physical meaning by associating the two possible projections of the magnetic moment $\boldsymbol{\mu}$, as measured in the Stern-Gerlach experiment, with two arbitrarily chosen distinct values of σ.

$$\sigma = +1 \quad \text{with } \mu_B = -\frac{e\hbar}{2mc}$$
$$\sigma = -1 \quad \text{with } \mu_B = +\frac{e\hbar}{2mc} \tag{16.8}$$

Often $\sigma = +1$ is referred to as ''spin up'' and $\sigma = -1$ as ''spin down'' (see Figure 16.1). We assume that the basic rules of quantum mechanics apply to the new independent variable in the same way as to the old ones.

In the coordinate representation, the probability amplitude or wave function for an electron now depends on the discrete variable σ in addition to x, y, z, and may be written as $\psi_\sigma(\mathbf{r}, t) = \langle \mathbf{r}, \sigma, t | \Psi \rangle$. This can be regarded as a two-component object composed of the two complex-valued amplitudes, $\psi_+(\mathbf{r}, t) = \langle \mathbf{r}, +1, t | \Psi \rangle$ for ''spin up'' and $\psi_-(\mathbf{r}, t) = \langle \mathbf{r}, -1, t | \Psi \rangle$ for ''spin down.'' Suppressing the time dependence of the wave function, $|\psi_\pm(x, y, z)|^2\, dx\, dy\, dz$ is thus assumed to measure the probability of finding the particle near x, y, z, and of revealing the value $\mu_B = \mp\beta_0$, respectively, for the projection of the magnetic moment in the direction of the field $\mathbf{B}$.

There is no *a priori* reason to expect that such a modest generalization of the theory will be adequate, but the appearance of merely two traces in the Stern-Gerlach experiment, and, as we will see later, the splitting of the spectral lines of one-electron atoms into narrow doublets, make it reasonable to assume that a variable which can take on only two different values—sometimes called a *dichotomic* variable—may be a sufficiently inclusive addition to the theory.

The mathematical apparatus of Chapters 9 and 10 can be applied to extend the formalism of wave mechanics *without* spin to wave mechanics *with* spin. Since space and spin coordinates are assumed to be independent of each other, it is natural to use a two-dimensional matrix representation for the specification of the state:

$$\boxed{|\Psi\rangle \rightarrow \psi = \begin{pmatrix} \psi_+(x, y, z) \\ \psi_-(x, y, z) \end{pmatrix}} \tag{16.9}$$

where the matrix ψ, with one column and two rows, now stands for a two-component spin wave function. Wherever we previously had an integration over the continuously infinitely many values of the position variables, we must now introduce an additional summation over the pairs of values which σ assumes, such as in the normalization integral:

$$\int |\psi_+(x, y, z)|^2\, dx\, dy\, dz + \int |\psi_-(x, y, z)|^2\, dx\, dy\, dz = 1 \tag{16.10}$$

It is instructive to study the behavior of the spin variable separately from the space coordinates and to consider a system whose state is described by ignoring the x, y, z coordinates and determined, at least to good approximation, entirely by two spin amplitudes. We designate such a general spin state as χ and write it as

$$\chi = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \tag{16.11}$$

The complex-valued matrix elements c_1 and c_2 are the amplitudes for "spin up" and "spin down," respectively. The column matrix (16.11), often referred to as a *spinor*, represents a state vector in an abstract two-dimensional complex vector space. Such states are more than mathematical idealizations. In many physical situations, the bodily motion of a particle can be ignored or treated classically, and only its spin degree of freedom need be considered quantum mechanically. The study of nuclear magnetism is an example, since we can discuss many experiments by assuming that the nuclei are at fixed positions and only their spins are subject to change owing to the interaction with a magnetic field.

Study of the spin formalism in isolation from all other degrees of freedom serves as a paradigm for the behavior of any quantum system whose states can be described as linear superpositions of only two independent states. There are innumerable problems in quantum mechanics where such a two-state formalism is applicable to good approximation, but that have nothing to do with spin angular momentum. The analysis of reflection and transmission from a one-dimensional potential in Chapters 6 and 7 has already illustrated the convenience of the two-dimensional matrix formalism. Other examples are the coupling of the $2S$ and $2P$ states of the hydrogen atom through the Stark effect (Chapter 18), the magnetic quenching of the triplet state of positronium (Problem 4 in Chapter 17), the isospin description of a nucleon, the transverse polarization states of a photon (Chapter 23), and the life and death of a neutral kaon (Problem 1 in Chapter 16).

The basis states of the representation defined by the assignments (16.8) are

$$\alpha = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{16.12}$$

Thus, α represents a state with spin "up," and β represents a state with spin "down." In the general state,

$$\chi = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = c_1\alpha + c_2\beta \tag{16.13}$$

$|c_1|^2$ is the probability of finding the particle with spin up, and $|c_2|^2$ is the probability of finding it with spin down. Hence, we must require the normalization

$$|c_1|^2 + |c_2|^2 = (c_1^* \quad c_2^*)\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = 1 \tag{16.14}$$

This can be written as

$$\chi^\dagger\chi = 1 \tag{16.15}$$

if we remember that

$$\chi^\dagger = (c_1^* \quad c_2^*) \tag{16.16}$$

Given two spinors, χ and χ', the (Hermitian) inner (or complex scalar) product is defined as

$$\chi^\dagger\chi' = (c_1^* \quad c_2^*)\begin{pmatrix} c_1' \\ c_2' \end{pmatrix} = c_1^*c_1' + c_2^*c_2' \tag{16.17}$$

Two spinors are *orthogonal* if this product is zero. The two spinors α and β defined in (16.12) are orthogonal and normalized, as $\alpha^\dagger\alpha = \beta^\dagger\beta = 1$. Such pairs of orthonormal spinors span the basis of a representation.

All definitions and manipulations introduced in Chapters 4, 9, and 10 for complex linear vector (Hilbert) spaces of n dimensions can in this simple case, where $n = 2$, be written out explicitly in terms of two-dimensional matrices. If we commit ourselves to a specific fixed representation, all equations and theorems for state vectors and linear operators can be interpreted directly as matrix equations. As long as confusion is unlikely to occur, the same symbol can be used for a state and the spinor that represents it; similarly, the same letter may be used for a physical quantity and the matrix (operator) that represents it. In many ways, the spin formalism is much simpler than wave mechanics with its infinite-dimensional representations. Since the state vector space is two-dimensional, the mathematical complexity of the theory is significantly reduced.

For example, if A is a linear operator (perhaps representing a physical quantity), it appears as

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \tag{16.18}$$

and its action on the spinor χ, which produces the new spinor $\xi = A\chi$, is represented as

$$\begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \tag{16.19}$$

where the components of ξ are denoted by d_1 and d_2. The (Hermitian) adjoint $A^\dagger$ of A is

$$A^\dagger = \begin{pmatrix} A_{11}^* & A_{21}^* \\ A_{12}^* & A_{22}^* \end{pmatrix} \tag{16.20}$$

and the expectation value of A in the state χ is

$$\langle A \rangle = (c_1^* \quad c_2^*) \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \chi^\dagger A \chi \tag{16.21}$$

Exercise 16.1. In the spin matrix formalism, show that if and only if the expectation value of a physical quantity A is real-valued, the matrix A is Hermitian. Prove, by direct calculation, that the eigenvalues of any Hermitian 2×2 matrix are real and its eigenspinors orthogonal if the two eigenvalues are different. What happens if they are the same?

An arbitrary state can be expanded in terms of the orthonormal eigenspinors, u and v, of any Hermitian matrix A:

$$\chi = u(u^\dagger \chi) + v(v^\dagger \chi) \tag{16.22}$$

The expansion coefficients

$$(u^\dagger \chi) = (u_1^* \quad u_2^*) \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \quad \text{and} \quad (v^\dagger \chi) = (v_1^* \quad v_2^*) \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \tag{16.23}$$

are the probability amplitudes of finding the eigenvalues of A corresponding to the eigenspinors u and v, respectively.

To endow this purely mathematical framework with physical content, we must identify the physical quantities associated with the spin of a particle and link them with the corresponding Hermitian matrices. A physical quantity of principal interest is the component of the electron's intrinsic magnetic moment in the direction of the

magnetic field, which motivated the extension of the theory to dichotomic spin variables. Since **B** can be chosen to point in any direction whatever, we first select this to be the z axis of the spatial coordinate system. Then the z component of the intrinsic magnetic moment of an electron is evidently represented by the Hermitian matrix

$$\mu_z = -\beta_0 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{16.24}$$

since the eigenvalues of μ_z are to be $\mp\beta_0 = \mp e\hbar/2m_e c$, and the corresponding states may be represented by the basis spinors α and β.

How are the other components of $\boldsymbol{\mu}$ represented? The magnetic moment $\boldsymbol{\mu}$ has three spatial components, μ_x, μ_y, μ_z, and by choosing a different direction for **B** we can measure any projection, μ_B, of $\boldsymbol{\mu}$. If our two-dimensional formalism is adequate to describe the physical situation, any such projection μ_B must be represented by a Hermitian matrix with the eigenvalues $-\beta_0$ and $+\beta_0$. In order to determine the matrices μ_x and μ_y, we stipulate that the three components of $\langle\boldsymbol{\mu}\rangle$ must under a rotation transform as the components of an ordinary three-*vector*. Since an expectation value, such as $\langle\mu_x\rangle = \chi^\dagger \mu_x \chi$, is calculated from matrices and spinors, we cannot say how the components of $\langle\boldsymbol{\mu}\rangle$ transform unless we establish the transformation properties of a spinor χ under rotation. We will now turn to this task.

3. ***Spin and Rotations.*** Rotations of systems described by wave functions $\psi(x, y, z)$ were already considered in Chapter 11; here we extend the theory to spin states. We first consider a right-handed rotation of the physical system about the z axis, keeping the coordinate axes fixed. This is an *active* rotation, to be distinguished from a *passive* rotation, which leaves the physical system fixed and rotates the coordinate system. As long as we deal solely with the mutual relation between the physical system under consideration and the coordinate system, the distinction between these two kinds of rotations is purely a matter of taste. However, if, as is more commonly the case, the physical system that we describe by the quantum mechanical state vector is not isolated but is embedded in an environment of external fields or interacts with other systems, which we choose to keep fixed as the rotation is performed, the active viewpoint is the appropriate one, and we generally prefer it.

Figure 11.1(b) pictures an active rotation by an angle ϕ about the z axis, which carries an arbitrary spin state χ into a state χ'. The relation between these two spinors may be assumed to be linear. (As will be shown in Section 17.1, this assumption involves no loss of generality, and in any case we will see that a valid linear transformation representing any rotation can be found.) Thus, we suppose that the two spinors are related by

$$\chi' = U\chi \tag{16.25}$$

where U is a matrix whose elements depend on the three independent parameters of the rotation only, e.g., the axis of rotation $\hat{\mathbf{n}}$ and the angle ϕ.

Since the physical content of the theory should be invariant under rotation, we expect that normalization of χ implies the same normalization of χ':

$$\chi'^\dagger \chi' = \chi^\dagger U^\dagger U \chi = 1$$

Since χ is arbitrary, it follows that

$$U^\dagger U = I \tag{16.26}$$

so U must be a *unitary* matrix. From this matrix equation we infer that

$$\det U^\dagger \det U = |\det U|^2 = 1 \tag{16.27}$$

Hence, a unitary matrix has a unique inverse, $U^{-1} = U^\dagger$, and

$$UU^\dagger = I \tag{16.28}$$

The unitary matrix U, which corresponds to the rotation that takes χ into χ', is said to *represent* this rotation. If U_1 represents a rotation R_1 about an axis through the origin, and U_2 represents a second rotation R_2 also about an axis through the origin, then U_2U_1 represents another such rotation R_3, obtained by performing first R_1 and then R_2. In this way χ is first transformed to $\chi' = U_1\chi$, which subsequently is transformed to $\chi'' = U_2\chi' = U_2U_1\chi$. Alternatively, we could, according to Euler's famous theorem, have obtained the same physical state directly from χ by performing a single rotation R_3, represented by U_3. Hence, the unitary rotation matrices are required to have the property

$$U_2U_1 = e^{i\varphi(R_1R_2)}U_3 \tag{16.29}$$

The phase factor has been put in, because all spinors $e^{i\varphi}\chi$ represent the same state.

Our goal, the construction of U corresponding to a given rotation R, will be considerably facilitated if we consider *infinitesimal rotations* first. A small rotation must correspond to a matrix very near the identity matrix, and thus for a small rotation we write to a first approximation the first two terms in a Taylor series:

$$U = I - \frac{i}{\hbar}\,\varepsilon\hat{\mathbf{n}}\cdot\mathbf{J} \tag{16.30}$$

where $\hat{\mathbf{n}}$ is the axis of rotation, ε is the angle of rotation about this axis, and $\mathbf{J}$ represents three constant matrices J_x, J_y, J_z. Their detailed structure is yet to be determined, and they are called the *generators of infinitesimal rotations*. The factor $i/\hbar$ has been introduced so that $\mathbf{J}$ will have certain desirable properties. In particular, the imaginary coefficient ensures that $\mathbf{J}$ must be Hermitian if U is to be unitary, i.e.

$$\mathbf{J}^\dagger = \mathbf{J} \tag{16.31}$$

If the three matrices J_x, J_y, and J_z were known, U for any finite rotation could be constructed from (16.30) by successive application of many infinitesimal rotations, i.e., by integration of (16.30).

This integration is easily accomplished because any rotation can be regarded as the product of $N = \phi/\varepsilon$ successive rotations by a small angle ε about a fixed axis (constant $\hat{\mathbf{n}}$):

$$U_R \approx \left(I - \frac{i\phi}{\hbar N}\,\hat{\mathbf{n}}\cdot\mathbf{J}\right)^N \tag{16.32}$$

or in the limit $N \to \infty$,

$$U_R = \lim_{n\to\infty}\left(I - \frac{i\phi}{\hbar N}\,\hat{\mathbf{n}}\cdot\mathbf{J}\right)^N = \exp\left(-\frac{i}{\hbar}\,\phi\,\hat{\mathbf{n}}\cdot\mathbf{J}\right) \tag{16.33}$$

as in elementary calculus, even though U_R and $\hat{\mathbf{n}}\cdot\mathbf{J}$ are matrices. The exponential function with a matrix in the exponent is defined by (16.33) or by the usual power series expansion. The necessary groundwork for the matrix algebra was laid in Section 3.4, where we may read "matrix" for "operator."

We still have to derive the conditions under which a matrix of the form (16.33) is actually the solution to our problem, i.e., represents the rotation R and satisfies the basic requirement (16.29). The application of the condition (16.29) will lead to severe restrictions on the possible form of the Hermitian matrices J_x, J_y, J_z, which so far have not been specified at all. However, it is convenient not to attack this problem directly, but instead to discuss first the rotational transformation properties of a vector $\langle \mathbf{A} \rangle$, where A_x, A_y, A_z are three matrices (operators) such that the expectation values $\langle A_x \rangle$, $\langle A_y \rangle$, $\langle A_z \rangle$ transform as the components of a vector. As stated at the end of the last section, the components of the magnetic moment $\langle \boldsymbol{\mu} \rangle$ are an example of matrices that must satisfy this condition. Generally, as in Section 11.1, a set of three matrices A_x, A_y, A_z is called a *vector operator* $\mathbf{A}$ if the expectation values of A_x, A_y, A_z transform under rotation like the components of a vector.

It is of decisive importance to note that $\mathbf{J}$ itself is a vector operator. This follows from its definition as the generator of the infinitesimal rotation:

$$\chi' = \left(1 - \frac{i}{\hbar}\,\varepsilon\hat{\mathbf{n}} \cdot \mathbf{J}\right)\chi \tag{16.34}$$

Multiplying on the left by $\chi^\dagger$, we obtain

$$\chi^\dagger\chi' = \chi^\dagger\chi - \frac{i}{\hbar}\,\varepsilon\hat{\mathbf{n}} \cdot \langle \mathbf{J} \rangle \tag{16.35}$$

where the expectation value $\langle \mathbf{J} \rangle$ is taken with respect to the state χ. The inner products are invariant under a unitary transformation that represents an arbitrary finite rotation, applied simultaneously to both χ and χ'. Hence, the scalar product $\hat{\mathbf{n}} \cdot \langle \mathbf{J} \rangle$ is also a *rotational invariant*. Since $\hat{\mathbf{n}}$ is a vector, $\langle \mathbf{J} \rangle$ must also transform like a vector, and thus $\mathbf{J}$ is a vector operator.

The transformation properties of a three-vector $\langle \mathbf{A} \rangle = \chi^\dagger \mathbf{A} \chi$ under an active rotation are characterized by the equation

$$\langle \mathbf{A} \rangle' = \langle \mathbf{A} \rangle + \hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \langle \mathbf{A} \rangle)(1 - \cos\phi) + \hat{\mathbf{n}} \times \langle \mathbf{A} \rangle \sin\phi \tag{16.36}$$

where $\langle \mathbf{A} \rangle' = \chi'^\dagger \mathbf{A} \chi'$ is the expectation value of $\mathbf{A}$ after rotation. In standard 3×3 matrix notation, this equation appears as

$$\begin{pmatrix} \langle A_x \rangle' \\ \langle A_y \rangle' \\ \langle A_z \rangle' \end{pmatrix} = R \begin{pmatrix} \langle A_x \rangle \\ \langle A_y \rangle \\ \langle A_z \rangle \end{pmatrix} \tag{16.37}$$

where R is the usual real orthogonal rotation matrix (with $\det R = 1$) familiar from analytic geometry and corresponding to an active rotation.

Exercise 16.2. Check the transformation (16.36) by visualizing a three-dimensional rotation. Verify it algebraically for a right-handed rotation about the z axis and express it in 3×3 matrix form.

Exercise 16.3. Starting with an infinitesimal rotation about the unit vector $\hat{\mathbf{n}}(n_x, n_y, n_z)$, prove that the rotation matrix R can be represented as

$$R = \exp(-\phi\,\hat{\mathbf{n}} \cdot \mathbf{X}) \tag{16.38}$$

where

$$X_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \qquad X_y = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad X_z = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{16.39}$$

are three antisymmetric matrices.[3] Work out their commutation relations and compare them with the commutation relations for the components of angular momentum.

For an infinitesimal rotation, (16.36) reduces to

$$\langle \mathbf{A} \rangle' = \langle \mathbf{A} \rangle + \hat{\mathbf{n}} \times \langle \mathbf{A} \rangle \varepsilon \tag{16.40}$$

We now substitute the expression (16.30) on the left-hand side of this equation and equate the terms linear in ε on the two sides. Since χ is an arbitrary state, it follows that

$$\frac{i}{\hbar} \hat{\mathbf{n}} \cdot \mathbf{J}\mathbf{A} - \mathbf{A} \frac{i}{\hbar} \hat{\mathbf{n}} \cdot \mathbf{J} = \hat{\mathbf{n}} \times \mathbf{A} \tag{16.41}$$

or

$$\mathbf{A}\hat{\mathbf{n}} \cdot \mathbf{J} - \hat{\mathbf{n}} \cdot \mathbf{J}\mathbf{A} = [\mathbf{A}, \hat{\mathbf{n}} \cdot \mathbf{J}] = i\hbar \hat{\mathbf{n}} \times \mathbf{A} \tag{16.42}$$

which is exactly the same condition as Eq. (11.19) derived in Section 11.1, except for the replacement of the orbital angular momentum operator **L** by **J**. This generalization was already anticipated in Section 11.2, where an algebraic approach to the eigenvalue problem of the angular momentum operator was taken. We can make use of the results derived there, since **J** itself is a vector operator and must satisfy (16.42):

$$[\mathbf{J}, \hat{\mathbf{n}} \cdot \mathbf{J}\,] = i\hbar \hat{\mathbf{n}} \times \mathbf{J} \tag{16.43}$$

or, using the subscripts i, j, k, with values 1, 2, 3 to denote the Cartesian components x, y, z,

$$\boxed{[J_i, J_j] = i\hbar \varepsilon_{ijk} J_k \quad \text{or} \quad \varepsilon_{ijk} J_i J_j = i\hbar J_k} \tag{16.44}$$

The Levi-Civita symbol ε_{ijk} was defined in Section 11.5. Taking care to maintain the order of noncommuting operators, we may combine these commutation relations symbolically in the equation

$$\mathbf{J} \times \mathbf{J} = i\hbar \mathbf{J} \tag{16.45}$$

Exercise 16.4. Employing the techniques developed in Section 3.4, verify that the commutation relations for **A** and **J** assure the validity of the condition (16.36) or, explicitly,

$$\exp\left(\frac{i}{\hbar} \hat{\mathbf{n}} \cdot \mathbf{J}\phi\right) \mathbf{A} \exp\left(-\frac{i}{\hbar} \hat{\mathbf{n}} \cdot \mathbf{J}\phi\right) = \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \mathbf{A}) - \hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times \mathbf{A}) \cos\phi + \hat{\mathbf{n}} \times \mathbf{A} \sin\phi \tag{16.46}$$

for finite rotations.

[3] Biedenharn and Louck (1981). See also Mathews and Walker (1964), p. 430.

Since the trace of a commutator is zero, the commutation relations (16.44) imply that the trace of every component of **J** vanishes. Hence, by (10.29),

$$\det U_R = 1 \tag{16.47}$$

so that the matrices U_R representing a rotation are *unimodular*. If we evaluate the determinant on both sides of Eq. (16.29), we then conclude that $e^{i\varphi(R_1,R_2)} = \pm 1$ and (16.29) takes the more specific form

$$\boxed{U_2 U_1 = \pm U_3} \tag{16.48}$$

Applying successive finite rotations to a vector operator and using Eq. (16.46), it can be shown that the commutation relations for **J** are not only necessary but also sufficient for the unitary operator (16.33) to represent rotations and satisfy the requirement (16.48).[4] (For $n = 2$ a proof will be given in Section 16.4.)

Although they were prompted by our interest in the two-dimensional intrinsic spin of the electron, none of the arguments presented in this section have depended on the dimensionality of the matrices involved. The states χ and χ' connected by the unitary matrix U in (16.25) could have n rows, and all results would have been essentially the same. In particular, the commutation relations (16.43) or (16.44) would then have to be satisfied by three $n \times n$ matrices. That a closed matrix algebra satisfying these commutation relations can be constructed for every nonnegative integer n was already proved in Section 11.2. We will thus be able to use the results of this section in Chapter 17, when we deal with angular momentum in more general terms. In the remainder of this chapter, however, we confine ourselves to the case $n = 2$, and we must now explicitly determine the Hermitian 2×2 matrices **J** which satisfy the commutation relations.

4. ***The Spin Operators, Pauli Matrices, and Spin Angular Momentum.*** Following the usual convention, we supposed in Section 16.2 that the z component of the vector operator $\boldsymbol{\mu}$, the intrinsic magnetic moment, is represented by the diagonal matrix (16.24) and that the components c_1 and c_2 of the spinor χ are the probability amplitudes for finding $\mu_z = -\beta_0$ (spin up) and $+\beta_0$ (spin down), respectively. A rotation about the z axis can have no effect on these probabilities, implying that the matrix

$$U = \exp\left(-\frac{i}{\hbar}\,\phi\, J_z\right)$$

must be diagonal in the representation we have chosen. It follows that J_z must itself be a diagonal matrix.

Exercise 16.5. From the commutation relations, prove that if the z component of some vector operator is represented by a diagonal matrix, J_z must also be diagonal (as must be the z component of any vector operator).

The problem of constructing the matrices **J** in a representation in which J_z is diagonal has already been completely solved in Section 11.2. The basis vectors (or basis spinors or basis kets or basis states) of this representation are the eigenvectors of J_z. The commutation relations (11.24) are identical to (16.44). We now see that

[4] Biedenharn and Louck (1981), Section 3.5.

for the description of the spin of the electron we must use as a basis the two eigenstates of J_z and $\mathbf{J}^2$, which correspond to $j = 1/2$ and $m = \pm 1/2$. From Eqs. (11.42) and (11.43) (Exercise 11.11) we obtain the matrices

$$J_+ = J_x + iJ_y = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad J_- = J_x - iJ_y = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{16.49}$$

It is customary and useful to define a vector operator (matrix) $\boldsymbol{\sigma}$ proportional to the 2×2 matrix $\mathbf{J}$:

$$\mathbf{J} = \frac{\hbar}{2} \boldsymbol{\sigma} \tag{16.50}$$

Hence,

$$\sigma_+ = \sigma_x + i\sigma_y = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \sigma_- = \sigma_x - i\sigma_y = \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix} \tag{16.51}$$

from which we obtain the celebrated ***Pauli spin matrices***,

$$\boxed{\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}} \tag{16.52}$$

Some simple properties of the Pauli matrices are easily derived.

Exercise 16.6. Prove that the Pauli matrices are unitary and that

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1 \tag{16.53}$$

Exercise 16.7. Prove that

$$\sigma_x\sigma_y = i\sigma_z, \qquad \sigma_y\sigma_z = i\sigma_x, \qquad \sigma_z\sigma_x = i\sigma_y \tag{16.54}$$

and that any two different Pauli matrices anticommute: $\sigma_x\sigma_y + \sigma_y\sigma_x = 0$, and so forth.

Exercise 16.8. Prove that the only matrix which commutes with all three Pauli matrices is a multiple of the identity. Also show that no matrix exists which anticommutes with all three Pauli matrices.

The traces of all Pauli matrices vanish:

$$\text{trace } \sigma_x = \text{trace } \sigma_y = \text{trace } \sigma_z = 0 \tag{16.55}$$

which is a reflection of the general property that the trace of any commutator of two matrices vanishes. It follows from the commutation relations (16.42) that the trace of any vector operator is zero. In the two-dimensional case ($n = 2$), this implies that the z-component A_z of every vector operator is proportional to J_z and consequently that all vector operators $\mathbf{A}$ are just multiples of $\mathbf{J}$:

$$\mathbf{A} = k\mathbf{J} = k\frac{\hbar}{2}\boldsymbol{\sigma} \tag{16.56}$$

where k is a constant number. The proportionality of $\mathbf{A}$ and $\mathbf{J}$, which generally holds only for $n = 2$, is the simplest illustration of the Wigner-Eckart theorem which will be derived in Chapter 17.

The four matrices 1, σ_x, σ_y, σ_z are linearly dependent, and any 2×2 matrix can be represented as

$$A = \lambda_0 1 + \lambda_1 \sigma_x + \lambda_2 \sigma_y + \lambda_3 \sigma_z = \lambda_0 1 + \boldsymbol{\lambda} \cdot \boldsymbol{\sigma} \tag{16.57}$$

If A is Hermitian, all coefficients in (16.57) must be real.

Exercise 16.9. Take advantage of the properties (16.54) and (16.55) of the Pauli matrices to work out the eigenvalues and eigenspinors of A in terms of the expansion coefficients λ_0 and $\boldsymbol{\lambda}$. Specialize to the case $\lambda_0 = 0$ and $\boldsymbol{\lambda} = \hat{\mathbf{n}}$, where $\hat{\mathbf{n}}$ is a real-valued arbitrary unit vector.

Exercise 16.10. Show that if U is a unitary 2×2 matrix, it can always be expressed as

$$U = e^{i\gamma}(1 \cos \omega + i\hat{\mathbf{n}} \cdot \boldsymbol{\sigma} \sin \omega) \tag{16.58}$$

where γ and ω are real angles, and $\hat{\mathbf{n}}$ is a real unit vector.

Exercise 16.11. If $\mathbf{A}$ and $\mathbf{B}$ are two vectors that commute with $\boldsymbol{\sigma}$, prove the useful identity

$$\boxed{(\boldsymbol{\sigma} \cdot \mathbf{A})(\boldsymbol{\sigma} \cdot \mathbf{B}) = \mathbf{A} \cdot \mathbf{B} + i\boldsymbol{\sigma} \cdot (\mathbf{A} \times \mathbf{B})} \tag{16.59}$$

Applying the identity (16.59) to the power series expansion of an exponential, we see that (16.58) is the same as

$$\boxed{U = \exp(i\gamma + i\omega\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}) = e^{i\gamma}(1 \cos \omega + i\hat{\mathbf{n}} \cdot \boldsymbol{\sigma} \sin \omega)} \tag{16.60}$$

which is a generalized de Moivre formula. Any unitary 2×2 matrix can be written in this form. In the two-dimensional spin formalism, the rotation matrix (16.33) takes the form

$$U_R = \exp\left(-\frac{i}{2}\phi\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}\right) \tag{16.61}$$

Comparing the last two expressions, we see that every unitary matrix with $\gamma = 0$ represents a rotation. The angle of rotation is $\phi = -2\omega$, and $\hat{\mathbf{n}}$ is the axis of rotation. For $\gamma = 0$ we have $\det U_R = 1$, and the matrix U_R is *unimodular*. The set of all unitary unimodular 2×2 matrices constitutes the *group* $SU(2)$. The connection between this group and three-dimensional rotations will be made precise in Chapter 17.

We may now write the rotation matrix (16.33) in the form

$$U_R = \exp\left(-\frac{i}{2}\phi\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}\right) = 1 \cos\frac{\phi}{2} - i\hat{\mathbf{n}} \cdot \boldsymbol{\sigma} \sin\frac{\phi}{2} \tag{16.62}$$

One simple but profound consequence of this equation is that for $\phi = 2\pi$ we get $U = -1$. A full rotation by 360° about a fixed axis, which is equivalent to zero rotation (or the identity), thus *changes the sign of every spinor component.* The double-valuedness of the spin rotation matrices is sanctioned, although not required, by the relation (16.48). Vectors (and tensors in general) behave differently: they return to their original values upon rotation. However, this sign change of spinors under rotation is no obstacle to their usefulness, since all expectation values and

matrix elements depend bilinearly on spinors, rendering them immune to the sign change.

Exercise 16.12. Using the special properties of the Pauli matrices, prove directly that

$$U_R^\dagger \boldsymbol{\sigma} U_R = \hat{\mathbf{n}}(\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}) - \hat{\mathbf{n}} \times (\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}) \cos \phi + \hat{\mathbf{n}} \times \boldsymbol{\sigma} \sin \phi \tag{16.63}$$

if U_R is given by (16.62)

Since the right-hand side of (16.63) is the expression for the rotated form of a vector, it is evident that if we perform in succession two rotations R_1 and R_2, equivalent to an overall rotation R_3, we can conclude that

$$U_1^\dagger U_2^\dagger \boldsymbol{\sigma} U_2 U_1 = U_3^\dagger \boldsymbol{\sigma} U_3$$

or

$$[U_2 U_1 U_3^\dagger, \boldsymbol{\sigma}] = 0$$

From Exercise 16.8 we thus infer that $U_2 U_1 U_3^\dagger$ must be a multiple of the identity. Since the spin rotation matrices are unimodular (det $U = 1$), we are led back to (16.48), proving that in the case $n = 2$ the commutation relations are not only necessary but also sufficient to ensure the validity of the group property (16.48).

It may be helpful to comment on the use of the term *vector* that is current in quantum mechanics. A vector $\mathbf{V}$ in ordinary space must not be confused with a (state) vector such as χ in a (complex) vector space. In the context of this chapter, the latter is represented by a two-dimensional spinor, but in other situations, such as when describing the intrinsic degree of freedom of a spin-one particle, the state vector is three-dimensional. To contrast the different behavior of spinors and ordinary vectors under rotation, we consider the example of a rotation about the x axis by an angle ϕ. From (16.62), (16.52), and (16.25), we obtain for the spinor components:

$$c_1' = c_1 \cos \frac{\phi}{2} - i c_2 \sin \frac{\phi}{2}$$

$$c_2' = -i c_1 \sin \frac{\phi}{2} + c_2 \cos \frac{\phi}{2}$$

The components of a vector $\mathbf{V}$, on the other hand, transform according to

$$\begin{aligned} V_x' &= V_x \\ V_y' &= V_y \cos \phi - V_z \sin \phi \\ V_z' &= V_y \sin \phi + V_z \cos \phi \end{aligned}$$

The differences between these two transformations are apparent, but they are connected. If $\mathbf{A}$ is a vector operator, the spinor transformation induces the correct transformation among the components of the expectation value $\langle \mathbf{A} \rangle$.

We must now find the matrix representation of the physical observables that are associated with an electron or other spin one-half particle. Since, according to (16.56), the vector operator $\boldsymbol{\sigma}$ is essentially unique, we conclude from (16.24) that the intrinsic magnetic moment of an electron is given by

$$\boldsymbol{\mu} = -\beta_0 \boldsymbol{\sigma} \tag{16.64}$$

thus completing the program of determining the components of $\boldsymbol{\mu}$.

What about the *intrinsic angular momentum* of the electron, its *spin*? It was shown in Section 16.1 that conservation of angular momentum is destroyed unless the electron is endowed with an intrinsic angular momentum, in addition to its orbital angular momentum. The interaction energy (16.6) responsible for compromising the spherical symmetry of the central forces is proportional to $\boldsymbol{\mu} \cdot \mathbf{L}$, which in turn, according to (16.64), is proportional to $\boldsymbol{\sigma} \cdot \mathbf{L}$ for a spin one-half particle. We express the wave function ψ for the state of the particle in the spinor form (16.9),

$$\psi = \begin{pmatrix} \psi_+(x, y, z) \\ \psi_-(x, y, z) \end{pmatrix} \tag{16.65}$$

When the interaction $\boldsymbol{\sigma} \cdot \mathbf{L}$ is applied to ψ, the operator $\mathbf{L}$ acts only on the functions $\psi_\pm(x, y, z)$ of the coordinates, but $\boldsymbol{\sigma}$ couples the two spinor components.

A term of the form $\boldsymbol{\sigma} \cdot \mathbf{L}$ in the Hamiltonian is often referred to as the *spin-orbit interaction*. As was explained in Section 16.1, an interaction of this form arises in atoms as a magnetic and relativistic correction to the electrostatic potential. It produces a fine structure in atomic spectra. In nuclei the spin-orbit interaction has its origin in strong interactions and has very conspicuous effects.

In the presence of a spin-orbit interaction, $\mathbf{L}$ is no longer a constant of the motion. It is our hope that an *intrinsic angular momentum* $\mathbf{S}$ can be defined in such a manner that, when it is added to the *orbital angular momentum* $\mathbf{L}$, the *total angular momentum*,

$$\boxed{\mathbf{J} = \mathbf{L} + \mathbf{S}} \tag{16.66}$$

will again be a constant of motion. Since $\mathbf{S}$, like $\mathbf{L}$, is a vector operator, it must be proportional to $\boldsymbol{\sigma}$. Indeed, the spin angular momentum $\mathbf{S}$ is nothing other than the generator (16.50) of rotations for spinors:

$$\mathbf{S} = \frac{\hbar}{2} \boldsymbol{\sigma} \tag{16.67}$$

since both $\mathbf{S}$ and $\mathbf{L}$ are just different realizations of the generator $\mathbf{J}$, which was introduced in its general form in Section 16.3. The unitary operator that transforms the state (16.65) of a particle *with spin* under an infinitesimal three-dimensional rotation must be given by

$$U_R = 1 - \frac{i}{\hbar} \varepsilon \hat{\mathbf{n}} \cdot \mathbf{S} - \frac{i}{\hbar} \varepsilon \hat{\mathbf{n}} \cdot \mathbf{L} = 1 - \frac{i}{\hbar} \varepsilon \hat{\mathbf{n}} \cdot \mathbf{J} \tag{16.68}$$

The scalar operator $\boldsymbol{\sigma} \cdot \mathbf{L}$ is invariant under this rotation, and (16.66) is the desired constant of the motion.

We can verify this identification by employing the commutation relation (16.42) for a vector operator twice. First, we let $\mathbf{J} = \mathbf{A} = \mathbf{L}$ and replace the vector $\hat{\mathbf{n}}$ by $\mathbf{S}$, which is legitimate because $\mathbf{S}$ commutes with $\mathbf{L}$. This yields the equation

$$[\mathbf{L}, \mathbf{S} \cdot \mathbf{L}] = i\hbar \mathbf{S} \times \mathbf{L}$$

Next, we let $\mathbf{J} = \mathbf{A}$ and $\mathbf{S}$ and replace $\hat{\mathbf{n}}$ by $\mathbf{L}$:

$$[\mathbf{S}, \mathbf{L} \cdot \mathbf{S}] = i\hbar \mathbf{L} \times \mathbf{S}$$

Owing to the commutivity of **L** and **S**, we have $\mathbf{S} \times \mathbf{L} = -\mathbf{L} \times \mathbf{S}$ and $\mathbf{S} \cdot \mathbf{L} = \mathbf{L} \cdot \mathbf{S}$; hence, it follows that $\mathbf{J} = \mathbf{L} + \mathbf{S}$ commutes with the operator $\mathbf{L} \cdot \mathbf{S}$ and is indeed conserved in the presence of a spin-orbit interaction. No other combination of **L** and **S** would have satisfied this requirement.

Exercise 16.13. Show that no operator of the form $\mathbf{L} + a\boldsymbol{\sigma}$, other than $\mathbf{J} = \mathbf{L} + (\hbar/2)\boldsymbol{\sigma}$ commutes with the scalar $\boldsymbol{\sigma} \cdot \mathbf{L}$.

Evidently, any component of the intrinsic angular momentum **S** defined by (16.67) has the two eigenvalues $+\hbar/2$ and $-\hbar/2$. The maximum value of a component of **S** in units of $\hbar$ is 1/2, and we say that *the electron has spin* 1/2. Furthermore, we note that

$$\mathbf{S}^2 = \frac{\hbar^2}{4}\,\boldsymbol{\sigma} \cdot \boldsymbol{\sigma} = \frac{\hbar^2}{4}\,(\sigma_x^2 + \sigma_y^2 + \sigma_z^2) = \frac{3\hbar^2}{4}\,\mathbf{1} \tag{16.69}$$

Hence, any spinor is an eigenspinor of $\mathbf{S}^2$, with eigenvalue $3\hbar^2/4$, corresponding to $s = 1/2$ if we express $\mathbf{S}^2$ as $s(s+1)\hbar^2$.

Thus, we see that when the spin is taken into account, $\mathbf{J} = \mathbf{L} + \mathbf{S}$ is the generator of infinitesimal rotations (multiplied by $\hbar$), and conservation of angular momentum is merely a consequence of the invariance of the Hamiltonian under rotations. This broad viewpoint, which places the emphasis on symmetries, is the subject of Chapter 17.

5. ***Quantum Dynamics of a Spin System.*** The general dynamical theory of Chapter 14 is directly applicable to any physical system with two linearly independent states, such as the spin of a particle in isolation from other degrees of freedom. In the Schrödinger picture, the time development of a two-component state or spinor $\chi(t)$ is governed by the equation of motion,

$$i\hbar\,\frac{d\chi(t)}{dt} = H\chi(t) \tag{16.70}$$

where the Hamiltonian H is in this instance a 2×2 matrix characteristic of the physical system under consideration. The essential feature of the equation of motion is its linearity, which preserves superpositions, but since we want to apply the theory to systems that can decay, we will at this stage not assume that H is necessarily Hermitian.

Obviously, if there are no time-dependent external influences acting and the system is invariant under translation in time, H must be a constant matrix, independent of t. Under these conditions, Eq. (16.70) can be integrated, giving

$$\chi(t) = \exp\left(-\frac{i}{\hbar}\,Ht\right)\chi(0) \tag{16.71}$$

in terms of the initial state $\chi(0)$.

As usual, it is convenient to introduce the eigenvalues of H, which are defined as the roots of the characteristic equation

$$\det(H - \lambda 1) = 0 \tag{16.72}$$

If there are two distinct roots $\lambda = E_1, E_2$ with $E_1 \neq E_2$, we have

$$H\chi_1 = E_1\chi_1 \quad \text{and} \quad H\chi_2 = E_2\chi_2 \tag{16.73}$$

and an arbitrary two-component spinor may be expanded as

$$\chi = c_1\chi_1 + c_2\chi_2 \tag{16.74}$$

If H is not Hermitian, its eigenvalues will generally not be real.

If $f(z)$ is a function of a complex variable, the function $f(H)$ of the matrix H is a new matrix defined by the relation

$$f(H)\chi = c_1 f(H)\chi_1 + c_2 f(H)\chi_2 = c_1 f(E_1)\chi_1 + c_2 f(E_2)\chi_2 \tag{16.75}$$

By substitution into (16.75), the equality

$$\boxed{f(H) = f(E_1)\frac{E_2 1 - H}{E_2 - E_1} + f(E_2)\frac{E_1 1 - H}{E_1 - E_2}} \tag{16.76}$$

is seen to hold. If the characteristic equation has only one distinct root, so that $E_2 = E_1$, the preceding equation degenerates into

$$f(H) = f(E_1)1 + f'(E_1)(H - E_1 1) \tag{16.77}$$

Exercise 16.14. Prove Eq. (16.76), and derive (16.77) from it in the limit $E_2 \to E_1$.

Equation (16.76) may be applied to expand the time development operator $f(H) = \exp\left(-\frac{i}{\hbar}Ht\right)$ in the form

$$\exp\left(-\frac{i}{\hbar}Ht\right) = \frac{1}{E_2 - E_1}(E_2 e^{-iE_1 t/\hbar} - E_1 e^{-iE_2 t/\hbar}) + \frac{H}{E_2 - E_1}(e^{-iE_2 t/\hbar} - e^{-iE_1 t/\hbar}) \tag{16.78}$$

if $E_1 \neq E_2$. A system whose Hamiltonian has exactly two distinct eigenvalues may be called a *two-level system*. The formula (16.78) answers all questions about its time development.

From (16.70) it follows in the usual way that

$$\frac{d}{dt}[\chi^\dagger(t)\chi(t)] = -\frac{i}{\hbar}\chi^\dagger(t)(H - H^\dagger)\chi(t) \tag{16.79}$$

and if H is constant in time this may be integrated to give

$$\chi^\dagger(t)\chi(t) = \chi^\dagger(0)e^{(i/\hbar)H^\dagger t}e^{-(i/\hbar)Ht}\chi(0) \tag{16.80}$$

If the matrix H is Hermitian, $\chi^\dagger\chi$ is constant and probability is conserved. This must certainly happen if H represents the energy. If H is Hermitian, E_1 and E_2 are real numbers and the corresponding eigenspinors are orthogonal.

If the Hamiltonian matrix is not Hermitian, the eigenvalues of H are complex numbers and can be expressed as

$$E_1 = E_{01} - i\frac{\Gamma_1}{2} \quad \text{and} \quad E_2 = E_{02} - i\frac{\Gamma_2}{2} \tag{16.81}$$

where the real parts, E_{01}, E_{02}, are the energy levels. If the imaginary parts, Γ_1 and Γ_2, are positive, the two eigenstates are *decaying* states. The general solution of the dynamical equation (16.70) is the superposition

$$\chi(t) = c_1 e^{-(i/\hbar)E_{01}t} e^{-\Gamma_1 t/2}\, \chi_1 + c_2 e^{-(i/\hbar)E_{02}t} e^{-\Gamma_2 t/2} \chi_2 \tag{16.82}$$

Unless the two decay rates are equal, the state does not generally follow a pure exponential decay law.

As an application, consider the example of the probability amplitude for a transition from an initial state α (''spin up'') to a state β (''spin down''). One gets immediately

$$\langle \beta | T(t, 0) | \alpha \rangle = \beta^\dagger e^{-(i/\hbar)Ht} \alpha = (0 \quad 1) e^{-(i/\hbar)Ht} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{\langle \beta | H | \alpha \rangle}{E_2 - E_1} (e^{-iE_2 t/\hbar} - e^{-iE_1 t/\hbar}) \tag{16.83}$$

The probability obtained from this expression exhibits an interference term.

As was mentioned in Section 16.3, the dynamics of two-state systems, with or without decay, is applicable in many different areas of physics, and the spin formalism can be adapted to all such systems. Often a two-state system is prepared or created in a state other than an eigenstate of the Hamiltonian H, and its time development is subsequently observed, displaying intriguing oscillations in time, due to interference between the eigenstates of H. Qualitatively similar effects occur in other few-state systems, but the analysis is more complicated. We confine ourselves to the case of two-state systems and use the ordinary electron or nuclear spin 1/2 in a magnetic field as the standard example of the theory.

Exercise 16.15. In many applications, conservation laws and selection rules cause a decaying two-level system to be prepared in an eigenstate of σ_z, say $\alpha = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, and governed by the simple normal Hamiltonian matrix

$$H = aI + b\sigma_x \tag{16.84}$$

where a and b are generally complex constants. In terms of the energy difference $\Delta E = E_{02} - E_{01}$ and the decay rates Γ_1 and Γ_2, calculate the probabilities of finding the system at time t in state α or state β, respectively.

Exercise 16.16. If the Hermitian matrix $\Gamma = i(H - H^\dagger)$ is positive definite, show that Γ_1 and Γ_2 defined by (16.81) are positive. Conversely, if $\Gamma_{1,2} > 0$ and if the two decaying eigenstates, χ_1 and χ_2, of H are orthogonal (implying that H is a normal matrix), show that the time rate of change of the total probability $\chi^\dagger(t)\chi(t)$ is negative for all states χ at all times. Verify this conclusion using the results of Exercise 16.15 as an example.

6. ***Density Matrix and Spin Polarization.*** In discussing two-level systems, we have so far characterized the states in terms of two-component spinors. In this section, we consider some other methods of specifying a state. The spinor (γ_1, γ_2 real)

$$\chi = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} |c_1| e^{i\gamma_1} \\ |c_2| e^{i\gamma_2} \end{pmatrix} \tag{16.85}$$

characterizes a particular state. However, the same physical state can be described by different spinors, since χ depends on four real parameters, but measurements can

give us only two parameters: the relative probabilities $|c_1|^2 : |c_2|^2$ and the relative phase, $\gamma_1 - \gamma_2$, of c_1 and c_2. If χ is normalized to unity

$$|c_1|^2 + |c_2|^2 = 1 \tag{16.86}$$

the only remaining redundancy is the common phase factor of the components of χ, and this is acknowledged by postulating that χ and $e^{i\alpha}\chi$ (α: arbitrary, real) represent the same state.

An elegant and useful way of representing the state without the phase arbitrariness is to characterize it by the density operator defined in (14.19) as $\rho = |\Psi\rangle\langle\Psi|$, and discussed in detail in Section 15.5. In the context of two-state quantum mechanics, the *density matrix* of a pure state χ is

$$\rho = \chi\chi^\dagger = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}(c_1^* \quad c_2^*) = \begin{pmatrix} |c_1|^2 & c_1c_2^* \\ c_2c_1^* & |c_2|^2 \end{pmatrix} \tag{16.87}$$

subject to the normalization condition (16.86) which requires that

$$\text{trace } \rho = 1 \tag{16.88}$$

According to the probability doctrine of quantum mechanics, its knowledge exhausts all that we can find out about the state.

The expectation value $\langle A \rangle$ of any operator A is expressed in terms of ρ as:

$$\begin{aligned} \langle A \rangle = \chi^\dagger A \chi &= (c_1^* \quad c_2^*)\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \\ &= |c_1|^2 A_{11} + c_1c_2^* A_{21} + c_2c_1^* A_{12} + |c_2|^2 A_{22} \\ &= \text{trace}\begin{pmatrix} |c_1|^2 & c_1c_2^* \\ c_2c_1^* & |c_2|^2 \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \text{trace}(\rho A) \end{aligned} \tag{16.89}$$

in accord with Eq. (15.105).

Exercise 16.17. If A is a Hermitian matrix with eigenspinors u and v, corresponding to the distinct eigenvalues A_1' and A_2', show that the probability of finding A_1' in a measurement of A on the state χ is given by

$$|(u^\dagger \chi)|^2 = \text{trace}(\rho P_{A_1'}) = \text{trace}\left(\rho \frac{A_2' - A}{A_2' - A_1'}\right) = \text{trace}(\rho u u^\dagger) \tag{16.90}$$

where $P_{A_1'} = uu^\dagger$ represents the projection operator for the eigenvalue A_1'.

Like any 2×2 matrix, ρ can be expanded in terms of the Pauli matrices σ_x, σ_y, σ_z, and 1. Since ρ is Hermitian and its trace equals unity, it can according to (16.57) most generally be represented as

$$\boxed{\rho = \frac{1}{2}(1 + \mathbf{P} \cdot \boldsymbol{\sigma})} \tag{16.91}$$

where P_x, P_y, P_z are three real numbers given by

$$\begin{aligned} P_x &= 2\,\text{Re}(c_1^* c_2) \\ P_y &= 2\,\text{Im}(c_1^* c_2) \\ P_z &= |c_1|^2 - |c_2|^2 \end{aligned} \tag{16.92}$$

It is immediately verified that ρ has eigenvalues 0 and 1. The eigenspinor that corresponds to the latter eigenvalue is χ itself, i.e.,

$$\rho\chi = \chi \tag{16.93}$$

The other eigenspinor must be orthogonal to χ. The matrix ρ applied to it gives zero. Hence, when ρ is applied to an arbitrary state φ, we have

$$\rho\varphi = \chi(\chi^\dagger\varphi) \tag{16.94}$$

since χ is assumed to be normalized to unity. We thus see that ρ projects φ in the "direction" of χ. It follows that the density matrix is idempotent:

$$\rho^2 = \rho \tag{16.95}$$

Exercise 16.18. Show directly from (16.87) that the density matrix for a pure spin state is idempotent and has eigenvalues 0 and 1.

If (16.91) is required to be idempotent and the identity (16.59) is employed, we obtain

$$\mathbf{P}\cdot\mathbf{P} = P_x^2 + P_y^2 + P_z^2 = \mathbf{P}^2 = 1 \tag{16.96}$$

Hence, the state is characterized by two independent real parameters, as it should be.

The expectation value of σ_x in the state χ is

$$\langle\sigma_x\rangle = \text{trace}(\rho\sigma_x) = \frac{1}{2}\,\text{trace}\;\sigma_x + \frac{1}{2}\,\mathbf{P}\cdot\text{trace}(\boldsymbol{\sigma}\sigma_x) = \frac{1}{2}\,P_x\cdot\text{trace}(\sigma_x^2) = P_x$$

where use is made of the fact that trace $\boldsymbol{\sigma} = 0$ and $\sigma_x^2 = 1$. We get from this and analogous equations for σ_y and σ_z the simple formula

$$\boxed{\mathbf{P} = \langle\sigma\rangle = \text{trace}(\rho\boldsymbol{\sigma}) = \text{trace}(\boldsymbol{\sigma}\rho)} \tag{16.97}$$

proving that $\mathbf{P}$ transforms like a vector under rotations.

Combining (16.91) with (16.93), we find that the spinor χ is an eigenspinor of the matrix $\mathbf{P}\cdot\boldsymbol{\sigma}$:

$$\mathbf{P}\cdot\boldsymbol{\sigma}\chi = \chi \tag{16.98}$$

Hence, the unit vector $\mathbf{P}$ may legitimately be said to point in the direction of the particle's spin. The vector $\mathbf{P}$ is also known as the *polarization vector* of the state. It may be characterized by the two spherical coordinates, the polar and azimuthal angles, which specify a point on the unit sphere.

Exercise 16.19. Given a spinor

$$\chi = \begin{pmatrix} e^{i\alpha}\cos\delta \\ e^{i\beta}\sin\delta \end{pmatrix} \tag{16.99}$$

calculate the polarization vector $\mathbf{P}$ and construct the matrix U_R which rotates this state into $\begin{pmatrix}1\\0\end{pmatrix}$. Prove that the probability $p_{\hat{\mathbf{n}}}$ of finding this particle to be in a state represented by the polarization vector $\hat{\mathbf{n}}$ is

$$p_{\hat{\mathbf{n}}} = \frac{1}{2}\,\text{trace}[\rho(1 + \hat{\mathbf{n}}\cdot\boldsymbol{\sigma})] = \frac{1}{2}\,(1 + \mathbf{P}\cdot\hat{\mathbf{n}}) \tag{16.100}$$

and show that this result agrees with expectations for $\hat{\mathbf{n}} = \mathbf{P}$, $\hat{\mathbf{n}} = -\mathbf{P}$, and $\hat{\mathbf{n}} \perp \mathbf{P}$.

Although the language we have used in describing the properties of **P** refers to spin and rotations in ordinary space, the concepts have more general applicability, and the formalism allows us to define a "polarization vector" corresponding to the state of any two-level system. The polarization **P** is then a "vector" in an abstract three-dimensional Euclidean space, and the operator

$$\exp\left(-\frac{i}{2}\,\phi\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}\right)$$

induces "rotations" in this space. Examples are the isospin space in nuclear physics and the abstract polarization vector which can be defined to represent two isolated atomic states in interaction with an electromagnetic field as might be generated by laser light.

The formalism is particularly useful to describe the polarization states of electromagnetic radiation. Any two "orthogonal" polarization states may be chosen as the basis states for the representation, but the two (right and left) circular polarization states are usually preferred. The general elliptic polarization state of a light wave or photon is a superposition of the basis states in a two-dimensional complex vector space, which in optics is known as the *Jones vector* space. The elements of the corresponding density matrix are essentially the *Stokes parameters* of the polarization state. The vector **P** provides yet another representation of the polarization state of light. To combat the almost inevitable confusion caused by the double meaning of the term *polarization*, in the context of the polarization of light **P** is best referred to as the *Poincaré* vector, in recognition of the mathematician who introduced the two-sphere as a convenient tool for representing the elliptic polarization states.

The time evolution of the density matrix ρ can be obtained from the equation of motion for χ,

$$i\hbar\frac{d\chi}{dt} = H\chi \tag{16.101}$$

where H is assumed to be a Hermitian 2×2 matrix. Using the definition (16.87) of the density matrix, we obtain

$$\frac{\partial\rho}{\partial t} = \frac{d\chi}{dt}\chi^\dagger + \chi\frac{d\chi^\dagger}{dt} = \frac{1}{i\hbar}H\chi\chi^\dagger - \frac{1}{i\hbar}\chi\chi^\dagger H$$

or

$$\boxed{i\hbar\frac{\partial\rho}{\partial t} = H\rho - \rho H} \tag{16.102}$$

All of these equations are merely concrete matrix realizations of the general formalism of Sections 14.1 and 15.5.

Exercise 16.20. Derive the properties of the density matrix that represents a stationary state.

The equation of motion for any expectation value $\langle A\rangle$ is familiar:

$$i\hbar\frac{d\langle A\rangle}{dt} = \langle AH - HA\rangle + i\hbar\left\langle\frac{\partial A}{\partial t}\right\rangle \tag{16.103}$$

It is instructive to derive the equation of motion for the vector $\mathbf{P} = \langle \boldsymbol{\sigma} \rangle$. To obtain a simple formula, it is convenient to represent the Hamiltonian operator H as

$$H = \frac{1}{2}(Q_0 1 + \mathbf{Q} \cdot \boldsymbol{\sigma}) \tag{16.104}$$

where Q_0 and the three components of the vector $\mathbf{Q}$ are real numbers, which may be functions of time. By (16.103), (16.104), and the spin commutation relations summarized in the equation $\boldsymbol{\sigma} \times \boldsymbol{\sigma} = 2i\boldsymbol{\sigma}$, we derive

$$\frac{d\mathbf{P}}{dt} = \frac{d\langle \boldsymbol{\sigma} \rangle}{dt} = \frac{1}{i\hbar} \langle \boldsymbol{\sigma} H - H\boldsymbol{\sigma} \rangle = \frac{1}{2i\hbar} \langle \boldsymbol{\sigma}\mathbf{Q} \cdot \boldsymbol{\sigma} - \mathbf{Q} \cdot \boldsymbol{\sigma}\boldsymbol{\sigma} \rangle$$
$$= \frac{1}{2i\hbar} \langle \mathbf{Q} \times (\boldsymbol{\sigma} \times \boldsymbol{\sigma}) \rangle = \frac{1}{\hbar} \mathbf{Q} \times \langle \boldsymbol{\sigma} \rangle$$

or

$$\boxed{\hbar \frac{d\mathbf{P}}{dt} = \mathbf{Q} \times \mathbf{P}} \tag{16.105}$$

Since

$$\frac{d\mathbf{P}^2}{dt} = \frac{d(\mathbf{P} \cdot \mathbf{P})}{dt} = 2\mathbf{P} \cdot \frac{d\mathbf{P}}{dt} = \frac{2}{\hbar} \mathbf{P} \cdot (\mathbf{Q} \times \mathbf{P}) = 0$$

the vector $\mathbf{P}$ maintains a constant length. This is merely another way of saying that, when the Hamiltonian is Hermitian, the normalization of χ is conserved during the motion.

If $\mathbf{Q}$ is a *constant* vector, (16.105) implies that $\mathbf{P}$ precesses about $\mathbf{Q}$ with a constant angular velocity

$$\boldsymbol{\omega}_Q = \frac{\mathbf{Q}}{\hbar} \tag{16.106}$$

If

$$\mathbf{P}(0) = \mathbf{P}_0 \quad \text{and} \quad \mathbf{Q}/Q = \hat{\mathbf{Q}}$$

the solution of (16.105) is

$$\begin{aligned} \mathbf{P}(t) &= \hat{\mathbf{Q}}(\mathbf{P}_0 \cdot \hat{\mathbf{Q}}) + [\mathbf{P}_0 - \hat{\mathbf{Q}}(\mathbf{P}_0 \cdot \hat{\mathbf{Q}})] \cos \omega_Q t + \hat{\mathbf{Q}} \times \mathbf{P}_0 \sin \omega_Q t \\ &= \hat{\mathbf{Q}}(\mathbf{P}_0 \cdot \hat{\mathbf{Q}}) + \hat{\mathbf{Q}} \times (\mathbf{P}_0 \times \hat{\mathbf{Q}}) \cos \omega_Q t + \hat{\mathbf{Q}} \times \mathbf{P}_0 \sin \omega_Q t \qquad (16.107) \\ &= \mathbf{P}_0 \cos \omega_Q t + 2\hat{\mathbf{Q}}(\mathbf{P}_0 \cdot \hat{\mathbf{Q}}) \sin^2 \frac{\omega_Q t}{2} + \hat{\mathbf{Q}} \times \mathbf{P}_0 \sin \omega_Q t \end{aligned}$$

Exercise 16.21. Show that if $\mathbf{Q}$ is constant, $\mathbf{Q} \cdot \mathbf{P}$ and $(d\mathbf{P}/dt)^2$ are constants of the motion. Verify that (16.107) is the solution of (16.105). [See also Eq. (16.63).]

If $\mathbf{Q}$ is a constant vector and the initial polarization $\mathbf{P}_0$ is parallel to $\mathbf{Q}$, it is seen from (16.107) and Figure 16.3 that $\mathbf{P}$ is constant and equal to $\hat{\mathbf{Q}}$ or $-\hat{\mathbf{Q}}$. These two vectors represent the two stationary states of the system. Their energies are given by the eigenvalues of H, but only the energy difference, ΔE, is of physical interest. Since $\hat{\mathbf{Q}} \cdot \boldsymbol{\sigma}$ has the eigenvalues $+1$ and -1, the eigenvalues of H are $(Q_0 \pm Q)/2$ and

$$\Delta E = Q = \hbar\omega_Q \tag{16.108}$$

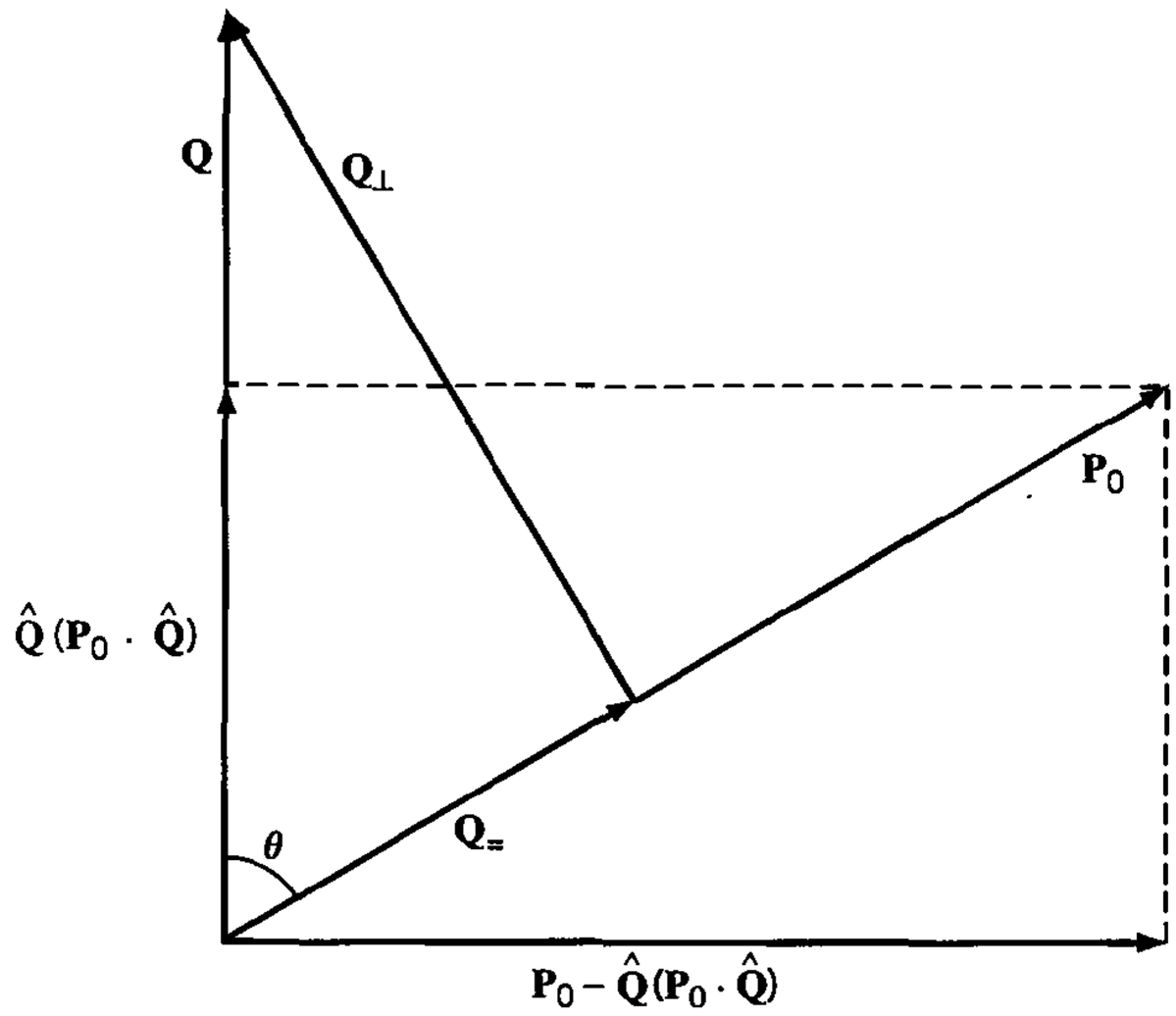


Figure 16.3. Precession of the spin polarization vector about **Q**. The initial polarization vector $\mathbf{P}_0$ and **Q** define the plane of the figure, and the precession angle θ is the angle between $\mathbf{P}_0$ and **Q**. The Rabi oscillations have the maximum amplitude $\sin^2 \theta$.

The probability of $\mathbf{P}(t)$ pointing in the direction $-\mathbf{P}_0$ at time t is

$$\frac{1}{2}\left[1 - \mathbf{P}_0 \cdot \mathbf{P}(t)\right] = (\mathbf{P}_0 \times \hat{\mathbf{Q}})^2 \sin^2 \frac{\omega_Q}{2} t = \frac{(\mathbf{P}_0 \times \mathbf{Q})^2}{Q^2} \sin^2 \frac{\omega_Q}{2} t \quad (16.109)$$

If, as indicated in Figure 16.3, we decompose the constant vector **Q** into two components parallel or longitudinal ($\mathbf{Q}_=$) and perpendicular or transverse ($\mathbf{Q}_\perp$) to the initial polarization $\mathbf{P}_0$,

$$\mathbf{Q} = \mathbf{Q}_= + \mathbf{Q}_\perp = \mathbf{P}_0(\mathbf{Q} \cdot \mathbf{P}_0) + \mathbf{P}_0 \times (\mathbf{Q} \times \mathbf{P}_0) \quad (16.110)$$

(16.109) can be written as

$$\boxed{\frac{1}{2}\left[1 - \mathbf{P}_0 \cdot \mathbf{P}(t)\right] = \frac{Q_\perp^2}{Q_=^2 + Q_\perp^2} \sin^2 \frac{\sqrt{Q_=^2 + Q_\perp^2}}{2\hbar} t = \sin^2 \theta \sin^2 \frac{Q}{2\hbar} t} \quad (16.111)$$

where θ is the angle between the polarization vector and **Q** (Figure 16.3). Formula (16.111) can be interpreted as describing periodic oscillations induced by the transverse field ($\mathbf{Q}_\perp$) between two energy levels split by the longitudinal field ($\mathbf{Q}_=$) by an amount $\Delta E_0 = Q_=$. Generically known as *Rabi oscillations*, these transitions between "spin up" and "spin down" eigenstates states of the unperturbed Hamiltonian, $H_0 = (1 + \mathbf{Q}_= \cdot \boldsymbol{\sigma})/2$, are caused by the constant perturbation $H - H_0 = \mathbf{Q}_\perp \cdot \boldsymbol{\sigma}/2$. In the special case of "resonance," when $\Delta E_0 = Q_= = 0$, the maximum amplitude of the Rabi oscillations is unity, and the initial state is totally depleted whenever $\omega_Q t = (2n + 1)\pi$. We emphasize that this analysis is exact and does not rely on any perturbation approximation.[5]

[5]For a full discussion of Rabi oscillations, with examples, see Cohen-Tannoudji, Diu, and Laloë (1977), Chapter IV.

The energy level splitting (16.108) is caused by the entire constant field **Q**. Transitions between the two stationary states can be induced if the spin system is exposed to a time-dependent oscillating field that has the same or a similar frequency as the spin precession. For example, if a spin 1/2 particle, whose degrees of freedom other than the spin can be neglected, is placed in a magnetic field **B**, the Hamiltonian can be written as

$$H = -\boldsymbol{\mu} \cdot \mathbf{B} = -\gamma \frac{\hbar}{2} \boldsymbol{\sigma} \cdot \mathbf{B} \tag{16.112}$$

The quantity γ is the *gyromagnetic ratio*, and the vector **Q** is given by

$$\mathbf{Q} = -\hbar\gamma\mathbf{B}$$

A constant field $\mathbf{B}_0$ causes a precession of **P** with angular velocity $\omega_0 = -\gamma B_0$.[6] If in addition an oscillating magnetic field with the same (or nearly the same) frequency is applied, the system will absorb or deliver energy, and the precession motion of **P** will be changed. These general principles are at the basis of all the magnetic resonance techniques that are so widely used in basic and applied science.

A special case of an *oscillating* field, for which a solution of the equation of motion can easily be obtained, is that in which the vector **Q** rotates uniformly about a fixed axis. Suppose that $\boldsymbol{\omega}$ is its angular velocity. It is advantageous to change over to a frame of reference which is rotating with the same angular velocity. Viewed from the rotating frame of reference, **Q** is a constant vector. If we denote the time rate of change of **P** with respect to the fixed system by $d\mathbf{P}/dt$, and with respect to the rotating system by $\partial\mathbf{P}/\partial t$, we have

$$\frac{d\mathbf{P}}{dt} = \frac{\partial\mathbf{P}}{dt} + \boldsymbol{\omega} \times \mathbf{P}$$

as is well known from the kinematics of rigid bodies; hence,

$$\hbar \frac{\partial\mathbf{P}}{dt} = (\mathbf{Q} - \hbar\boldsymbol{\omega}) \times \mathbf{P} \tag{16.113}$$

Since in the rotating frame of reference $\mathbf{Q} - \hbar\boldsymbol{\omega}$ is a constant vector, the problem has effectively been reduced to the previous one. Equation (16.113) can therefore be solved by transcribing the solution of (16.105) appropriately.

Exercise 16.22. If **Q** rotates uniformly about a fixed axis, the equation of motion (16.101) may conveniently be transformed to a frame of reference that rotates similarly. Derive the new Hamiltonian and show that it corresponds effectively to precession about the constant vector $\mathbf{Q} - \hbar\boldsymbol{\omega}$, providing an independent derivation of (16.113).

Exercise 16.23. If a constant magnetic field $\mathbf{B}_0$, pointing along the z axis, and a field $\mathbf{B}_1$, rotating with angular velocity ω in the xy plane, act in concert on a spin system (gyromagnetic ratio γ), calculate the polarization vector **P** as a function of time. Assume **P** to point in the z direction at $t = 0$. Calculate the Rabi oscillations in the rotating frame, and plot the average probability that the particle has "spin

[6]This is the quantum analogue of the classical Larmor precession described in Goldstein (1980), Section 5-9.

down'' as a function of ω/ω_0 for a value of $B_1/B_0 = 0.1$. Show that a resonance occurs when $\boldsymbol{\omega} = -\gamma \mathbf{B}_0$. (This arrangement is a model for all magnetic resonance experiments.)

Although so far in this section the density matrix ρ for spin states was assumed to represent a pure spinor state χ, almost every statement and equation involving ρ can be immediately applied to a mixed state, illustrating the general density operator theory of Section 15.5. The only exceptions are propositions that assume that ρ is idempotent or that the polarization vector satisfies $|\mathbf{P}| = 1$, since the conditions (16.95) or (16.96) are necessary and sufficient for the state to be pure and representable by a spinor χ.

A pure or mixed state is represented by a Hermitian density matrix whose eigenvalues are positive and sum to unity, as required by (16.88). For any density matrix the inequality (15.120) holds:

$$0 \leq \text{trace } \rho^2 \leq (\text{trace } \rho)^2 = 1 \tag{16.114}$$

In terms of the polarization vector, we have

$$\rho^2 = \rho + \frac{1}{4}(\mathbf{P} \cdot \mathbf{P} - 1)$$

If this identity is used in the inequality (16.114), we conclude that generally $|\mathbf{P}| \leq 1$, and that for a proper mixed state, i.e., one that is not a pure state, $|\mathbf{P}| < 1$. An unpolarized spin system has $\rho = (1/2)1$ and $\mathbf{P} = 0$. In spin-dependent scattering processes, which are the subject of the next section, proper mixed states representing incident particle beams are the rule rather than the exception.

7. ***Polarization and Scattering.*** The theory of scattering was developed in Chapter 13, neglecting the spin entirely. However, the forces that cause a beam of particles to be scattered may be spin-dependent, and it is then necessary to supplement the theory accordingly. The incident particles with spin one-half are represented by a wave function of the form

$$e^{ikz}\chi_{\text{inc}} = e^{ikz}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}, \qquad \chi^{\dagger}_{\text{inc}}\chi_{\text{inc}} = 1$$

Following the procedure of Chapter 13, we must look for asymptotic solutions of the Schrödinger equation which have the form

$$e^{ikz}\chi_{\text{inc}} + f(\theta, \varphi)\frac{e^{ikr}}{r} \tag{16.115}$$

but the scattering amplitude $f(\theta, \varphi)$ is now a two-component spinor. Spin-dependent scattering of a particle occurs, for instance, if the Hamiltonian has the form

$$H = \frac{\mathbf{p}^2}{2m} + V(r) + W(r)\mathbf{L} \cdot \boldsymbol{\sigma} \tag{16.116}$$

representing a spin-orbit interaction term in addition to a central force.

The superposition principle—and more specifically, the linearity of the Schrödinger equation—allows us to construct the solution (16.115) from the two particular solutions that correspond to $\chi_{\text{inc}} = \alpha$ and $\chi_{\text{inc}} = \beta$. These two special cases describe incident beams that are polarized along the direction of the initial momentum and

opposite to it. The polarization is said to be *longitudinal*. We are thus led to look for two solutions of the asymptotic form

$$\psi_1 \cong e^{ikz}\alpha + (S_{11}\alpha + S_{21}\beta)\frac{e^{ikr}}{r} \tag{16.117}$$

$$\psi_2 \cong e^{ikz}\beta + (S_{12}\alpha + S_{22}\beta)\frac{e^{ikr}}{r} \tag{16.118}$$

The quantities in parentheses are the appropriate scattering amplitudes.

Exercise 16.24. Show that the incident waves $e^{ikz}\alpha$ and $e^{ikz}\beta$ are eigenstates of J_z. What are the eigenvalues?

Multiplying (16.117) by c_1, and (16.118) by c_2, and adding the two equations, we obtain by superposition the more general solution

$$e^{ikz}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \frac{e^{ikr}}{r}\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \left(e^{ikz} + \frac{e^{ikr}}{r}S\right)\chi_{\text{inc}} \tag{16.119}$$

Here S stands for 2×2 *scattering matrix*

$$\boxed{S = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}} \tag{16.120}$$

S depends on the angles θ and φ, and on the momentum k. The scattering problem is solved if S can be determined as a function of these variables.

The form of S can be largely predicted by invariance arguments, although its dependence on the scattering angle θ can be worked out only by a detailed calculation, such as a phase shift analysis. Here we will only deduce the general form of the scattering matrix. The basic idea is to utilize the obvious constants of the motion that the symmetries of the problem generate. If A commutes with the Hamiltonian, then if ψ is an eigenfunction of H, $A\psi$ is also an eigenfunction of H, and both belong to the same energy. The state $A\psi$ may represent the same scattering state as ψ, or a different one of the same energy, depending on the asymptotic form of ψ.

Let us assume that, owing to spherical symmetry of the scattering potential, H is invariant under rotations and, according to Section 16.4, commutes with the components of $\mathbf{J}$. Expression (16.116) shows an example of a spin-dependent Hamiltonian with rotational symmetry. The incident waves in (16.117) and (16.118) are eigenstates of J_z with eigenvalues $+\hbar/2$ and $-\hbar/2$, respectively (Exercise 16.24). Since the operator J_z leaves the radial dependence of the scattered wave unchanged, the solutions (16.117) and (16.118) must both be eigenfunctions of J_z By requiring that

$$J_z\psi_1 = \hbar\left(\frac{1}{i}\frac{\partial}{\partial\varphi} + \frac{1}{2}\sigma_z\right)\psi_1 = \frac{\hbar}{2}\psi_1$$

and

$$J_z\psi_2 = \hbar\left(\frac{1}{i}\frac{\partial}{\partial\varphi} + \frac{1}{2}\sigma_z\right)\psi_2 = -\frac{\hbar}{2}\psi_2$$

it is easily seen that S_{11} and S_{22} can be functions of θ only and that the off-diagonal elements of the scattering matrix have the form

$$S_{12} = e^{-i\varphi} \times \text{function of } \theta, \qquad S_{21} = e^{i\varphi} \times \text{function of } \theta \tag{16.121}$$

Furthermore, the Hamiltonian H is assumed to be invariant under a reflection with respect to any coordinate plane. This is true for the spin-orbit interaction in (16.116), because both $\mathbf{L}$ and $\mathbf{S}$ or $\boldsymbol{\sigma}$ are axial vector operators, and their inner product is a scalar operator. The operator for reflection in the yz plane is $P_x\sigma_x$, where P_x simply changes x into $-x$, and σ_x has the effect of changing a spin state in such a way that $\langle\sigma_y\rangle$ and $\langle\sigma_z\rangle$ change sign, while $\langle\sigma_x\rangle$ remains unchanged, as behooves an axial (or pseudo-) vector. (For a more general explanation of reflections and parity for systems with spin, see Section 17.9.)

Since

$$\sigma_x\alpha = \beta, \qquad \sigma_x\beta = \alpha$$

the reflection in the yz plane changes the incident wave $e^{ikz}\alpha$ into $e^{ikz}\beta$ and leaves e^{ikr}/r invariant. Hence, (16.117) must go over into (16.118). In terms of spherical polar coordinates, P_x has the effect of changing φ into $\pi - \varphi$. It follows from this and (16.121) that

$$S_{11} = S_{22} = g(\theta), \qquad S_{21}(-\varphi, \theta) = -S_{12}(\varphi, \theta) = -e^{-i\varphi}h(\theta)$$

Consequently, we may write

$$S = \begin{pmatrix} g(\theta) & h(\theta)e^{-i\varphi} \\ -h(\theta)e^{i\varphi} & g(\theta) \end{pmatrix} = g(\theta)1 + ih(\theta)(\sigma_y \cos\varphi - \sigma_x \sin\varphi) \tag{16.122}$$

The unit vector $\hat{\mathbf{n}}(-\sin\varphi, \cos\varphi, 0)$ is normal to the plane of scattering and points in the direction of $\mathbf{k}_{\text{inc}} \times \mathbf{k}_{\text{scatt}}$. We conclude that the scattering matrix has the form

$$\boxed{S = g(\theta)1 + ih(\theta)\hat{\mathbf{n}} \cdot \boldsymbol{\sigma}} \tag{16.123}$$

The functions $g(\theta)$ and $h(\theta)$ are generalizations for the spin 1/2 case of the scattering amplitude $f(\theta)$ in Chapter 13. For rotationally invariant potentials, they can be parametrized by a generalization of the phase shift analysis of Section 13.5, but if they are to be computed from the assumed interaction, a set of coupled radial Schrödinger equations must ultimately be solved. The terminology "spin-flip" amplitude for $h(\theta)$ and "non-spin-flip" for $g(\theta)$ is self-explanatory.

Exercise 16.25. Show that the same scattering matrix is obtained by requiring reflection symmetry with respect to the xz plane.

Knowing S, we can calculate the intensity of the beam for a given direction. If (16.119) is the asymptotic form of the wave function, then by a straightforward generalization of the results of Chapter 13, the differential scattering cross section is found to be

$$\frac{d\sigma}{d\Omega} = (S\chi_{\text{inc}})^\dagger S\chi_{\text{inc}} = \chi_{\text{inc}}^\dagger S^\dagger S\chi_{\text{inc}} \tag{16.124}$$

which is merely the analogue of $|f(\theta)|^2$ for a particle with spin. If the density matrix ρ_{inc} describes the state of spin polarization of the incident beam, whether the state

be pure or—as is frequently the case—mixed, this expression may be generalized to

$$\boxed{\frac{d\sigma}{d\Omega} = \text{trace}\,(\rho_{\text{inc}} S^\dagger S)} \tag{16.125}$$

Since $S\chi_{\text{inc}}$ is the state of the particles scattered from an incident spin state χ_{inc} into the specified direction, the density matrix corresponding to the scattered part of the wave function is

$$\rho_{\text{scatt}} = \frac{S\rho_{\text{inc}}S^\dagger}{\text{trace}(S\rho_{\text{inc}}S^\dagger)} = \frac{S\rho_{\text{inc}}S^\dagger}{d\sigma/d\Omega} \tag{16.126}$$

Using the form (16.123) for the scattering matrix and

$$\rho_{\text{inc}} = \frac{1}{2}(1 + \mathbf{P}_0 \cdot \boldsymbol{\sigma}) \tag{16.127}$$

for the incident density matrix, we obtain the differential cross section in terms of the polarization $\mathbf{P}_0$ of the incident beam:

$$\boxed{\frac{d\sigma}{d\Omega} = |g|^2 + |h|^2 + i(g^*h - gh^*)\mathbf{P}_0 \cdot \hat{\mathbf{n}}} \tag{16.128}$$

The polarization of the scattered beam is

$$\mathbf{P} = \langle \boldsymbol{\sigma} \rangle = \text{trace}\,\rho_{\text{scatt}}\boldsymbol{\sigma} = \frac{\text{trace}(S\rho_{\text{inc}}S^\dagger\boldsymbol{\sigma})}{d\sigma/d\Omega} \tag{16.129}$$

If we use (16.123) to evaluate the trace, we obtain

$$\mathbf{P} = \frac{(|g|^2 - |h|^2)\mathbf{P}_0 + i(g^*h - gh^*)\hat{\mathbf{n}} + 2|h|^2\mathbf{P}_0 \cdot \hat{\mathbf{n}}\,\hat{\mathbf{n}} + (g^*h + gh^*)\mathbf{P}_0 \times \hat{\mathbf{n}}}{|g|^2 + |h|^2 + i(g^*h - gh^*)\,\mathbf{P}_0 \cdot \hat{\mathbf{n}}} \tag{16.130}$$

If the initial beam has transverse polarization and the scattering plane is chosen perpendicular to $\mathbf{P}_0$, or $\mathbf{P}_0 = P_0\hat{\mathbf{n}}$, it follows from (16.130) that

$$\mathbf{P} = \frac{(|g|^2 + |h|^2)P_0 + i(g^*h - gh^*)}{|g|^2 + |h|^2 + i(g^*h - gh^*)P_0}\,\hat{\mathbf{n}} \tag{16.131}$$

If the incident beam is unpolarized, $\mathbf{P}_0 = 0$, the scattered beam is polarized normal to the scattering plane:

$$\mathbf{P} = P\hat{\mathbf{n}} = i\,\frac{g^*h - gh^*}{|g|^2 + |h|^2}\,\hat{\mathbf{n}} \tag{16.132}$$

Exercise 16.26. Show that if the incident spin state is a pure transverse polarization state, the scattering amplitudes for the initial polarizations $\mathbf{P}_0 = \pm\hat{\mathbf{n}}$ are $g \pm ih$ and the scattering leaves the polarization unchanged, $\mathbf{P} = \mathbf{P}_0$.

Exercise 16.27. Show that the magnitude of the polarization given by (16.132) satisfies the condition $1 \geq |P| \geq 0$. *Hint:* Consider $|g - ih|$.

If the y axis is chosen to be along the direction of the transverse component of the polarization, $\mathbf{P}_0 - \mathbf{P}_0 \cdot \hat{\mathbf{k}}\,\hat{\mathbf{k}}$, we may write $\mathbf{P}_0 \cdot \hat{\mathbf{n}} = |\mathbf{P}_0 - \mathbf{P} \cdot \hat{\mathbf{k}}\,\hat{\mathbf{k}}| \cos\varphi$. With

these conventions, formula (16.128) for the differential cross section shows that the scattered intensity depends on the azimuthal angle as $I = a(\theta) + b(\theta) \cos \varphi$, in agreement with the empirical statement (16.7) in Section 16.1. In this way, we find substantiated our original supposition that the right-left asymmetry in the scattering of polarized beams of particles is a consequence of the particle spin.

Exercise 16.28. Assuming that $\mathbf{P}_0$ is perpendicular to the scattering plane, evaluate the *asymmetry* parameter A, defined as a measure of the right-left asymmetry by

$$A = \frac{(d\sigma/d\Omega)_+ - (d\sigma/d\Omega)_-}{(d\sigma/d\Omega)_+ + (d\sigma/d\Omega)_-} \tag{16.133}$$

where the subscripts + and − refer to the sign of the product $\mathbf{P}_0 \cdot \hat{\mathbf{n}}$. Show that if $\mathbf{P}_0 = \pm\hat{\mathbf{n}}$, the asymmetry A equals the degree of polarization P defined in (16.132). In particle polarization experiments, this quantity is referred to as the *analyzing power*.

8. ***Measurements, Probabilities, and Information.*** The spin formalism is so easy to survey that it lends itself particularly well to a demonstration of how quantum mechanics is to be interpreted, and how it relates to experiment, observation, and measurement. By using the 2×2 density matrix formalism to represent an arbitrary mixed spin state, we will be able to keep the discussion as general as possible. We assume that the spin state of the system is entirely specified by the density matrix ρ. Illustrating the general concepts of Section 15.5, we ask what kinds of ensembles might represent a known ρ, and what observables might be measured to determine an unknown ρ.

It is again convenient to represent the density matrix by the real-valued polarization three-vector $\mathbf{P}$, such that

$$\rho = \frac{1}{2}(1 + \mathbf{P} \cdot \boldsymbol{\sigma}) \tag{16.134}$$

Its eigenstates are represented by the pure-state density matrices:

$$\rho_+ = \frac{1}{2}(1 + \hat{\mathbf{P}} \cdot \boldsymbol{\sigma}) \quad \text{and} \quad \rho_- = \frac{1}{2}(1 - \hat{\mathbf{P}} \cdot \boldsymbol{\sigma}) \tag{16.135}$$

which correspond to eigenvalues $\bar{p}_+ = (1 + P)/2$ and $\bar{p}_- = (1 - P)/2$. The von Neumann entropy for this density matrix is, according to (15.128),

$$\begin{aligned} S(\rho) &= -\bar{p}_+ \ln \bar{p}_- - \bar{p}_- \ln \bar{p}_- \\ &= \ln 2 - \frac{1}{2}[(1 + P) \ln(1 + P) + (1 - P) \ln(1 - P)] \end{aligned} \tag{16.136}$$

The given density matrix may be realized by any ensemble of N pure states with polarization vectors $\hat{\mathbf{P}}_i$ and probabilities p_i with $\sum_i p_i = 1$, such that

$$\mathbf{P} = \sum_{i=1}^{N} p_i \hat{\mathbf{P}}_i \tag{16.137}$$

The Shannon mixing entropy (15.126) of this ensemble is

$$H(\mathscr{E}) = -\sum_{i=1}^{N} p_i \ln p_i \tag{16.138}$$

Exercise 16.29. As an example consider the 2×2 density matrix defined by the polarization vector

$$\mathbf{P} = \frac{1}{3}(\hat{\mathbf{x}} + \hat{\mathbf{y}} + \hat{\mathbf{z}}) \tag{16.139}$$

and realized by an ensemble $\mathscr{E}$ of the $N = 3$ equiprobable pure states that correspond to the spin pointing in the directions of the Cartesian coordinate vectors. For this mixed state, compute and compare the Shannon mixing entropy, $H(\mathscr{E})$, and the von Neumann entropy, $S(\rho)$.

We now consider the measurement of the observable $\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}$, which corresponds to projection operators (POM)

$$A_+ = \frac{1}{2}(1 + \boldsymbol{\sigma} \cdot \hat{\mathbf{n}}) \quad \text{and} \quad A_- = \frac{1}{2}(1 - \boldsymbol{\sigma} \cdot \hat{\mathbf{n}}) \tag{16.140}$$

The probability that the system is found with spin in the direction $\hat{\mathbf{n}}$ is the expectation value of the projection operator for the eigenstate of $\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}$:

$$\begin{aligned} p_{\hat{\mathbf{n}}} &= \frac{1}{2}\,\text{trace}[\rho(1 + \hat{\mathbf{n}} \cdot \boldsymbol{\sigma})] \\ &= \frac{1}{2}[1 + (\rho_{11} - \rho_{22})\, n_z + \rho_{12}(n_x + i n_y) + \rho_{12}^*(n_x - i n_y)] \end{aligned} \tag{16.141}$$

subject to the normalization condition

$$\text{trace}\,\rho = \rho_{11} + \rho_{22} = 1 \tag{16.142}$$

In terms of the polarization vector (Exercise 16.19),

$$p_{\hat{\mathbf{n}}} = \frac{1}{4}\,\text{trace}[(1 + \mathbf{P} \cdot \boldsymbol{\sigma})(1 + \hat{\mathbf{n}} \cdot \boldsymbol{\sigma})] = \frac{1}{2}(1 + \mathbf{P} \cdot \hat{\mathbf{n}}) \tag{16.143}$$

If three linearly independent observables $\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}_1$, $\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}_2$, $\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}_3$ are measured, using ensembles with the same density operator (although, of course, not the same particle), the matrix elements of ρ can be determined. This is similar to the description of the polarization state of a beam of light, which generally requires the measurement of the three Stokes parameters for its determination.

For example, the only possible results of a measurement of σ_z (or of any other component of $\boldsymbol{\sigma}$) are $+1$ or -1. By measuring σ_z for a very large number of replicas of the system, all prepared in the same state, we can determine the density matrix element $\rho_{11} = (1 + P_z)/2 = p_z$, which represents the relative frequency of finding "spin up" in the z direction. Other choices of the direction $\hat{\mathbf{n}}$ provide more information about the magnitudes and phases of the density matrix elements.

The outcome entropy for a measurement of $\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}$ is, according to (15.131),

$$H(\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}) = -\frac{1}{2}(1 + \mathbf{P} \cdot \hat{\mathbf{n}}) \ln\left[\frac{1}{2}(1 + \mathbf{P} \cdot \hat{\mathbf{n}})\right] - \frac{1}{2}(1 - \mathbf{P} \cdot \hat{\mathbf{n}}) \ln\left[\frac{1}{2}(1 - \mathbf{P} \cdot \hat{\mathbf{n}})\right] \tag{16.144}$$

As expected from Eq. (15.131), this entropy reaches its minimum value, the von Neumann entropy (16.136), when the measured spin points in the direction of the polarization vector: $\hat{\mathbf{n}} = \hat{\mathbf{P}}$. Figure 16.4 shows how the Shannon entropy for the outcome of the measurement of $\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}$ depends on $\mathbf{P} \cdot \hat{\mathbf{n}}$.

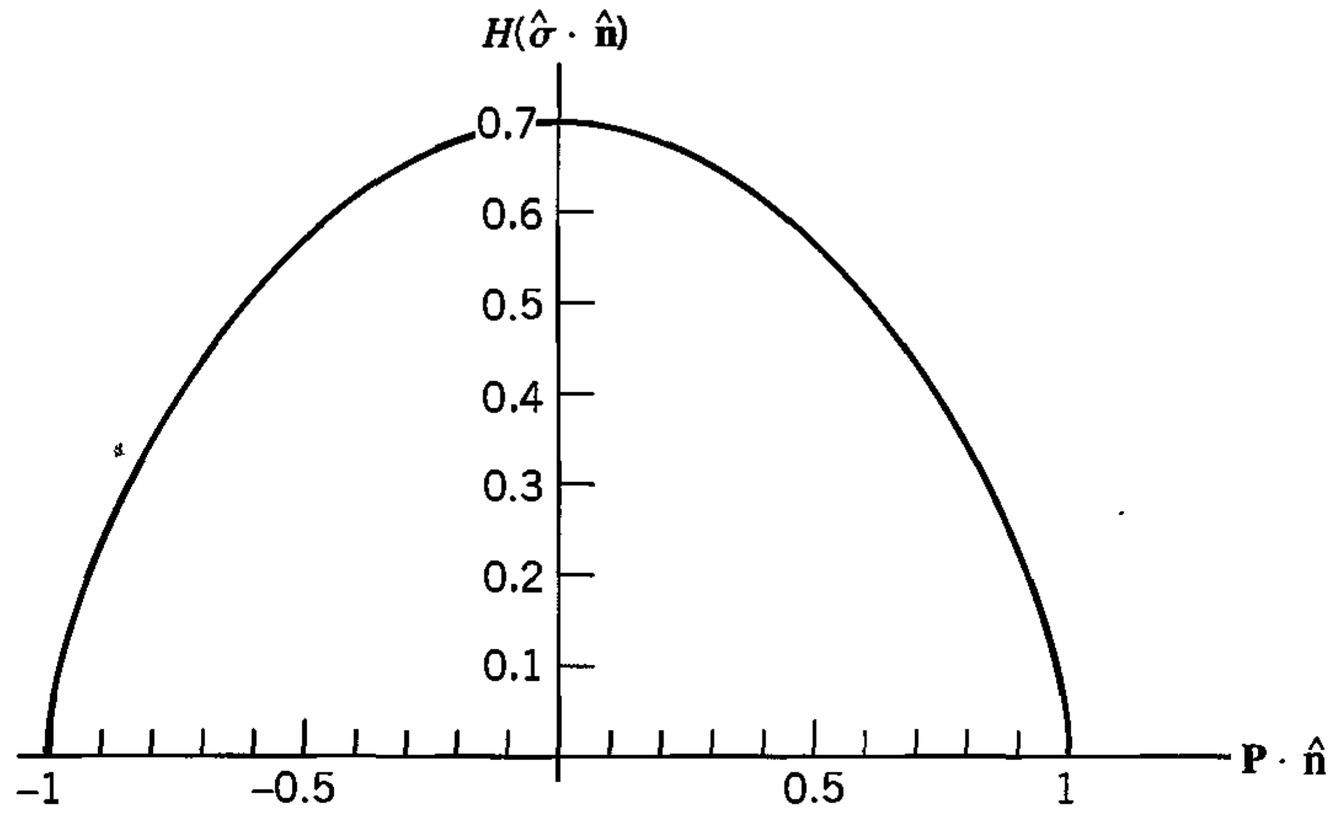


Figure 16.4. Outcome entropy $H(\boldsymbol{\sigma} \cdot \hat{\mathbf{n}})$ for a measurement of $\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}$ as a function of $\mathbf{P} \cdot \hat{\mathbf{n}}$.

Exercise 16.30. For the state specified by the polarization vector (16.139), calculate the Shannon entropy, $H(\boldsymbol{\sigma} \cdot \hat{\mathbf{n}})$, for the outcome of a measurement of $\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}$, with $\hat{\mathbf{n}}$ pointing along any one of the three coordinate axes. Compare the answer with the value of the von Neumann entropy of the state.

Exercise 16.31. If ρ represents the *pure* state,

$$\chi = c_1\alpha + c_2\beta = |c_1|e^{i\gamma_1}\alpha + |c_2|e^{i\gamma_2}\beta \tag{16.145}$$

and if $\hat{\mathbf{n}}$ is a unit vector in the yz plane making an angle θ with the z axis and $90° - \theta$ with the y axis, show that the probability for $\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}$ to yield the value $+1$ is

$$p_{\hat{\mathbf{n}}} = |c_1|^2 \cos^2\frac{\theta}{2} + |c_2|^2 \sin^2\frac{\theta}{2} - |c_1||c_2| \sin(\gamma_1 - \gamma_2) \sin\theta \tag{16.146}$$

Similarly, the probability for the value -1 is given by

$$p_{-\hat{\mathbf{n}}} = |c_1|^2 \sin^2\frac{\theta}{2} + |c_2|^2 \cos^2\frac{\theta}{2} + |c_1||c_2| \sin(\gamma_1 - \gamma_2) \sin\theta \tag{16.147}$$

Exercise 16.32. Write down the density matrix that represents the pure spin state

$$\chi = \sqrt{\frac{1}{3}}\,\alpha + \sqrt{\frac{2}{3}}\,\beta$$

and compare this with the density matrix for the mixed state about which we only know that the probability of "spin up" is one-third, and the probability of "spin down" is two-thirds. Calculate the von Neumann entropy for these two states.

Exercise 16.33. For a mixed state given by the density matrix

$$\rho = \frac{1}{75}\begin{pmatrix} 41 & 7 + 7i \\ 7 - 7i & 34 \end{pmatrix}$$

check the inequalities (15.120), and calculate the eigenvalues and eigenstates. Evaluate the von Neumann entropy, and compare this with the outcome entropy for a measurement of σ_z.

A molecular beam experiment of the Stern-Gerlach type has traditionally been regarded as the prototype of a measurement, fundamental to a proper understanding of quantum mechanics. When, as depicted in Figure 16.1, the z component of the spin is measured, there is a bodily separation of the particles that the experimenter subjects to the question, ''Is the spin up or down?'' The beam splits into two components made up, respectively, of those particles that respond with ''up'' or with ''down'' to this experimental question.

Before the particle interacts with the measuring apparatus, the preparation of its state is assumed to introduce no correlations between the spin and space degrees of freedom. Thus, initially the state has the simple product form

$$\rho_{\text{initial}} = \rho_{\text{spin}} \otimes \rho(\mathbf{r}, \mathbf{r}') \tag{16.148}$$

where ρ_{spin} denotes the spin state and $\rho(\mathbf{r}, \mathbf{r}')$ the purely spatial part of the density matrix. The probabilities for ''spin up'' and ''spin down'' in this state are

$$p_{\pm\hat{\mathbf{n}}} = \frac{1}{2}\,\text{trace}[\rho_{\text{spin}}(1 \pm \hat{\mathbf{n}} \cdot \boldsymbol{\sigma})]$$

The interaction with the Stern-Gerlach magnet causes the product state (16.148) to change into a more complicated correlated, or entangled, state. A careful analysis shows that in the region near the magnet where the two beams are well separated, the state of the particles can be represented as

$$\rho_{\text{final}} = p_{\hat{\mathbf{n}}}\,\frac{1}{2}\,(1 + \hat{\mathbf{n}} \cdot \boldsymbol{\sigma}) \otimes \rho_{\text{up}}(\mathbf{r}, \mathbf{r}') + p_{-\hat{\mathbf{n}}}\,\frac{1}{2}\,(1 - \hat{\mathbf{n}} \cdot \boldsymbol{\sigma}) \otimes \rho_{\text{down}}(\mathbf{r}, \mathbf{r}') \tag{16.149}$$

In this expression, $\rho_{\text{up}}(\mathbf{r}, \mathbf{r}')$ and $\rho_{\text{down}}(\mathbf{r}, \mathbf{r}')$ are spatial density matrices that describe the two separated particle beams. Usually, these spatial density matrices can be approximated in terms of wave packets moving along classical trajectories. The design of the apparatus ensures that they do not overlap and differ from zero only in the region traversed by the upper or lower beam, respectively. The upper component $\rho_{\text{up}}(\mathbf{r}, \mathbf{r}')$ is said to be *correlated* with the spin state in which $\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}$ is $+1$, and the down component $\rho_{\text{down}}(\mathbf{r}, \mathbf{r}')$ is correlated with the spin state in which $\boldsymbol{\sigma} \cdot \hat{\mathbf{n}}$ is -1. In the measurement, a particle reveals a spin ''up'' or ''down'' with probabilities equal to $p_{+\hat{\mathbf{n}}}$ and $p_{-\hat{\mathbf{n}}}$. If by some ingenious manipulation the two separated beams are recombined, additional terms must be included in (16.149) to account for the phase relations in the spin density matrix, which are lost if only the separated beams are considered.

In this connection, it is interesting to give some thought to a multiple Stern-Gerlach experiment in which two or more spin measurements are carried out in series. Let us assume again that σ_z is measured in the first experiment. If in the second experiment σ_z is remeasured, we will find that every particle in the upper beam has spin up, and every particle in the lower beam has spin down. Neither beam is split any further, confirming merely that conceptually the ideal Stern-Gerlach experiment is an exceptionally simple kind of measurement. Although it can change the state profoundly, from (16.148) to (16.149), this particular measurement does not alter the statistical distribution of the measured quantity (σ_z), nor does the spin state change between measurements. If in the second measurement the inhomogeneous magnetic field has a different direction, and thus a different component of the spin, say σ_y, is measured, we will find that each beam is split into two components of equal intensity, corresponding to the values $+1$ and -1 for σ_y (Figure 16.5).

This example shows the unavoidable effect which a measurement has on the

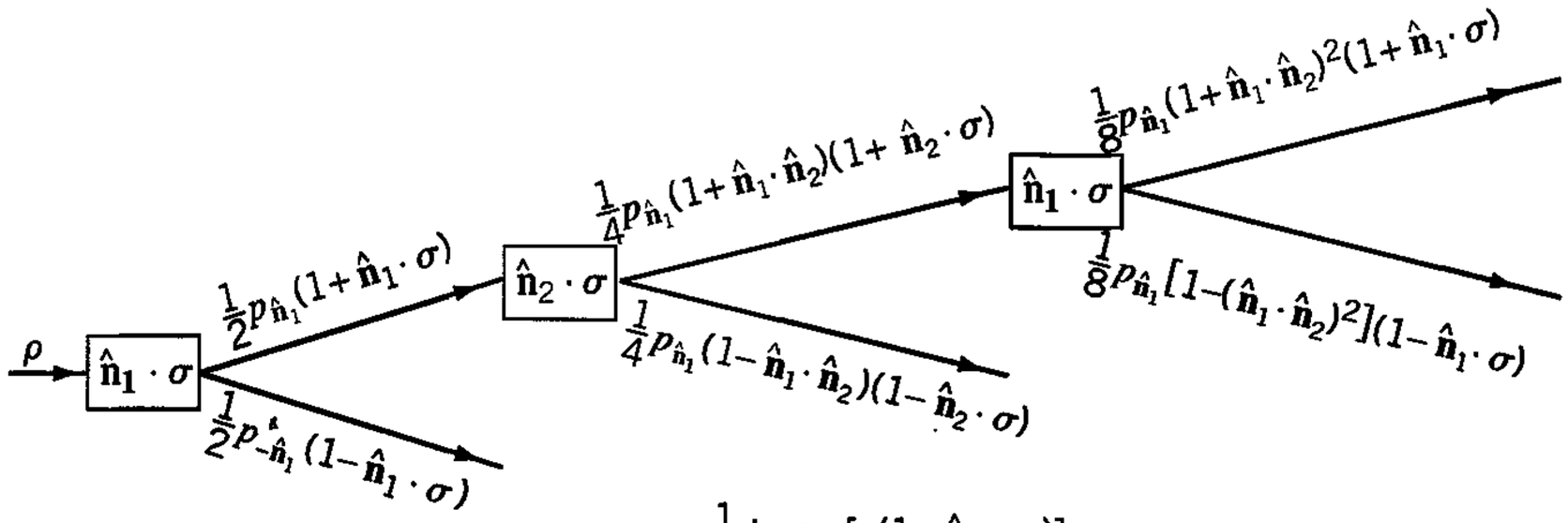


Figure 16.5. Successive Stern-Gerlach measurements of the spin projections $\hat{\mathbf{n}}_1 \cdot \boldsymbol{\sigma}$, $\hat{\mathbf{n}}_2 \cdot \boldsymbol{\sigma}$, $\hat{\mathbf{n}}_1 \cdot \boldsymbol{\sigma}$, producing pure "spin up" and "spin down" states. Each box symbolizes a beam splitting. The spin state of the incident beam is represented by the density matrix ρ. For each beam segment the spin component of the density matrix is specified. If $\hat{\mathbf{n}}_1 \neq \pm\hat{\mathbf{n}}_2$, the second beam splitter regenerates the "spin down" polarization state for direction $\hat{\mathbf{n}}_1$ from particles that entered it entirely with "spin up" along direction $\hat{\mathbf{n}}_1$.

system upon which the measurement is carried out. If ρ (short for ρ_{spin}) is the spin state before the measurement, and σ_y, rather than σ_z, is measured in a first experiment, then according to (16.134) the probability of finding $+1$ is

$$p_y = \frac{1}{2}\,\text{trace}[\rho(1 + \sigma_y)] = \frac{1}{2} + \frac{i}{2}(\rho_{12} - \rho_{12}^*) \tag{16.150}$$

whereas, if we precede this σ_y measurement by a measurement of σ_z, the probability of finding σ_y to be $+1$ is simply $\rho_{11}/2 + \rho_{22}/2 = (\rho_{11} + \rho_{22})/2 = 1/2$, in accordance with the common rule of compounding conditional probabilities. The probability p_y in (16.150) differs from $\rho_{11}/2 + \rho_{22}/2$ by an (off-diagonal) interference term, which the intervening σ_z measurement must wipe out if the probability interpretation of quantum mechanics is to be consistent. If in a third successive Stern-Gerlach measurement σ_z is measured again (Figure 16.5), we find anew a splitting of the beam, showing that the intervening measurement of σ_y has undone what the first σ_z measurement had accomplished. In the language of particle physics, we may say that the σ_y measurement has *regenerated* an amplitude for σ_z with value -1 in the branch in which the first measurement of σ_z had produced a pure $+1$ spin state.

In an ideal arrangement of this kind, two observables A and B are termed *compatible* if for any state of the system the results of a measurement of A are the same, whether or not a measurement of B precedes that of A. In other words, A and B are compatible if measuring B does not destroy the result of the determination of A. Clearly, this can happen only if the eigenstates of A are simultaneously also eigenstates of B. According to the arguments presented in Section 10.4, the necessary and sufficient condition for this is that the matrices representing A and B commute:

$$AB - BA = 0$$

Two observables are compatible if and only if the Hermitian matrices representing them commute.

For example, σ_z and σ_y are incompatible, for they do not commute; a state cannot simultaneously have a definite value of σ_z and σ_y. If we wish to measure σ_z and σ_y for a state ρ, two separate copies of the system must be used. The two components of the spin cannot be measured simultaneously on the same sample.

A measurement of the simple kind described by the initial state (16.148) and the final correlated state (16.149) is an example of an *ideal measurement* (sometimes called a *measurement of the first kind*) because the spatial separation of the two spin components allows the unambiguous identification of the two spin states and the (''impulsive'') measuring interaction leaves the two spatially separated spin states intact. If we consider the spin properties of the particles in isolation, the Stern-Gerlach device may be regarded as a spin filter that allocates fractions $p_{+\hat{\mathbf{n}}}$ and $p_{-\hat{\mathbf{n}}}$ of the particles definitely to the pure spin states represented by the density matrices $(1 + \hat{\mathbf{n}} \cdot \boldsymbol{\sigma})/2$ and $(1 - \hat{\mathbf{n}} \cdot \boldsymbol{\sigma})/2$, respectively. A correlated or entangled state like (16.149), in which the various eigenstate projections of the dynamical variable that is being measured are prevented from interfering after the measurement, is sometimes loosely interpreted by saying that the act of measurement ''puts'' the system into an eigenstate.

The acquisition of information provided by the measurement and the subsequent replacement of the original correlated state by selection of one or the other of its components with definite probabilities is conventionally referred to as the *reduction of the state*. In the spirit of the statistical interpretation of quantum mechanics, the reduction of the state—also known more dramatically as the *collapse of the wave packet*—is not meant to describe a physical process that affects the (probability) amplitudes by actual measurement manipulations. Only with this proviso is it defensible to say that after the reduction has taken place in an ideal measurement, the system *has* a definite value of the observable, namely, the eigenvalue determined by the measurement. A repetition of the measurement of the same quantity in the new state will now yield with certainty this very eigenvalue.

While the idealized Stern-Gerlach experiment illustrates many salient issues in quantum mechanics, the great variety of actual experiments defies any effort to classify all measurements systematically. Most measurements are more difficult to analyze, but for an understanding of the physical significance of quantum states it is sufficient to consider the simplest kind.

In the persistent debate about the foundations of quantum mechanics and the quantum theory of measurement, we take the position that the assignment of probabilities to the outcomes of various possible tests, acquired through experimental evidence, inspired guesswork, or other inferential procedures, is an indispensable part of the specification of a quantum system. In particular, in this view there is no reason to draw a line and make a qualitative distinction between a probability-free ''objective'' physical reality and the ''subjective'' realm of the observer who uses probabilities for interpreting the data. Rather, we regard the acquisition of information, and its evaluation in terms of probabilities, as an integral part of a full description of the physical system and its evolution.

Problems

1. The spin-zero neutral kaon is a system with two basis states, the eigenstates of σ_z, representing a particle K^0 and its antiparticle $\bar{K}^0$: The operator $\sigma_x = CP$ represents the combined parity (P) and charge conjugation (C), or particle-antiparticle, transformation and takes $\alpha = |K^0\rangle$ into $\beta = |\bar{K}^0\rangle$. The dynamics is governed by the Hamiltonian matrix

$$H = M - i\frac{\Gamma}{2}$$

where M and Γ are Hermitian 2×2 matrices, representing the mass-energy and decay properties of the system, respectively.[7] The matrix Γ is positive definite. A fundamental symmetry (under the combined CP and time reversal transformations) requires that $\sigma_x M^* = M\sigma_x$ and $\sigma_x \Gamma^* = \Gamma\sigma_x$.

(a) Show that in the expansion of H in terms of the Pauli matrices, the matrix σ_z is absent. Derive the eigenvalues and eigenstates of H in terms of the matrix elements of M and Γ. Are the eigenstates orthogonal?

(b) Assuming, as is the case to good approximation, that the Hamiltonian also satisfies the CP invariance conditions $\sigma_x M = M\sigma_x$ and $\sigma_x \Gamma = \Gamma\sigma_x$ show that H is normal, and construct its eigenstates, $|K_1^0\rangle$ and $|K_2^0\rangle$. If the measured lifetimes for these two decaying states are $\tau_1 = \hbar/\Gamma_1 = 0.9 \times 10^{-10}$ sec and $\tau_2 = \hbar/\Gamma_2 = 0.5 \times 10^{-7}$ sec, respectively, and if their mass difference is $m_2 - m_1 = 3.5 \times 10^{-6}$ eV/c^2, determine the numerical values of the matrix elements of M and Γ as far as possible.

(c) If the kaon is produced in the state K^0 at $t = 0$, calculate the probability of finding it still to be a K^0 at a later time t. What is the probability that it will be found in the $\bar{K}^0$ state? Plot these probabilities, exhibiting particle-antiparticle oscillations, as a function of time.

[7] See Perkins (1982) for experimental information on neutral kaons.

CHAPTER 17

Rotations and Other Symmetry Operations

Although symmetry arguments have already been used in almost every chapter, here we begin a systematic examination of the fundamental symmetries in quantum mechanics. The concepts are best understood by thinking about a concrete example. Rotations exhibit the interesting properties of many symmetry operations, and yet their theory is simple enough to keep the general features from being obscured by too much technical detail. If the theory of rotations is to be transferrable to other symmetries, it must be seen in the more abstract context of symmetry groups and their matrix representations. Much of the chapter is devoted to the practical problems of adding angular momenta and the extraction of symmetry-related properties of matrix elements of physical observables. In the last two sections, we deal with discrete symmetries (space reflection and time reversal) and their physical implications, and we return briefly to local gauge symmetries, which are distinctly different from global geometric symmetries.

1. ***The Euclidean Principle of Relativity and State Vector Transformations.*** The fundamental assumption underlying all applications of quantum mechanics is that ordinary space is subject to the laws of Euclidean geometry and that it is physically homogeneous and isotropic. By this we mean that we can move our entire physical apparatus from one place to another and we can change its orientation in space without affecting the outcome of any experiment. We say that there is no preferred position or orientation in space. The assumption that space is homogeneous and isotropic will be called the *Euclidean principle of relativity* because it denies that spatial location and orientation have any absolute significance.

Gravity seems at first sight to introduce inevitably a preferred direction, the vertical, into any experiment performed on or near the surface of the earth, but in quantum physics we are concerned primarily with atomic, nuclear, and particle processes in which gravitational effects play a negligible role. The apparent anisotropy of space can then usually be ignored, and the isotropy of space for such quantum processes can be tested directly by rotating the system at any desired angle. If gravitation cannot be neglected, as in some extremely sensitive neutron interferometry measurements,[1] there is again no conflict with the Euclidean principle of relativity, because we can imagine the earth to be part of the mechanical system and take its gravitational field into account when a rotation is performed. No violation of the Euclidean principle of relativity has ever been found in any laboratory experiment.

On a grander, astronomical and cosmological scale there are legitimate serious questions about the validity of the principle. Understanding the physics of the very early universe may require a fully developed theory that unites gravity with quantum mechanics. The scale at which quantum gravity is expected to be influential, called

[1]Werner (1994).

the *Planck scale*, is characterized, on purely dimensional grounds, by the Planck mass,

$$M_P c^2 = \left(\frac{\hbar c^5}{G}\right)^{1/2} \approx 10^{28} \text{ eV} = 10^{16} \text{ TeV}$$

The corresponding Planck length is of the order of 10^{-35} m, and the Planck time is 10^{-43} sec. These estimates make it clear why we will not be concerned with gravity.

Here we focus on the remarkable consequences that the Euclidean principle of relativity and its extension to the time dimension have for the structure of quantum mechanics. We will find that this principle severely restricts the possible forms that the quantum description of a given system can take.

A transformation that leaves the mutual relations of the physically relevant aspects of a system unaltered is said to be a *symmetry operation*. The Euclidean principle of relativity amounts to the assumption that geometric translations and rotations are symmetry operations. We first concentrate on rotations about an axis or a point, but in Section 17.9 we will extend the discussion to reflections. Nonrelativistic Galilean symmetry, involving transformations that include motions in time, was discussed in Section 4.7. The symmetry operations associated with the *Einstein principle of relativity* are based on Lorentz or Poincaré transformations and will be taken up in Chapters 23 and 24.

When a quantum system with a state vector Ψ is rotated in space to a new orientation, the state vector changes to Ψ'. The Euclidean principle of relativity requires that under rotation all probabilities be invariant, i.e., all inner products of two rotated states remain invariant in absolute value. We thus have a mapping of the vector space onto itself, $\Psi \mapsto \Psi'$, such that $|(\Psi', \Phi')|^2 = |(\Psi, \Phi)|^2$ for every pair of state vectors. Such a mapping is called an *isometry*. The mapping must be reversible, because we could equally well have started from the new orientation and rotated the system back to its old orientation. In the language of Section 16.3, we are considering *active* rotations.

Generally, we do not require invariance of inner products, which is the hallmark of unitary transformations, but only that the absolute values be invariant. Yet because of a remarkable theorem, we will ultimately be able to confine our attention essentially to unitary and antiunitary transformations. The reasoning given here applies to any symmetry operation and not just to rotations.

Theorem. If a mapping $\Psi \mapsto \Psi'$ of the vector space onto itself is given such that

$$|(\Psi', \Phi')|^2 = |(\Psi, \Phi)|^2 \tag{17.1}$$

then a second isometric mapping $\Psi' \mapsto \Psi''$, which is merely a phase change of every vector,

$$\Psi'' = e^{i\alpha(\Psi')}\Psi' \tag{17.2}$$

can be found such that

$$\Psi = \Psi_a + \Psi_b$$

is mapped into

$$\Psi'' = \Psi''_a + \Psi''_b$$

For the proof of this theorem the reader is referred to the literature.[2] The theorem shows that through rephasing of all vectors we can achieve a mapping that has one of the two fundamental properties of a linear operator: The transform of the sum of two vectors is equal to the sum of the transforms of the two vectors [see (9.43)]. It follows from this result and from (17.1) that

$$|(\Psi''_a, \Psi''_a + \Psi''_b)|^2 = |(\Psi_a, \Psi_a + \Psi_b)|^2$$

Hence, by applying (17.1) again,

$$(\Psi''_a, \Psi''_b) + (\Psi''_a, \Psi''_b)^* = (\Psi_a, \Psi_b) + (\Psi_a, \Psi_b)^*$$

or

$$\mathrm{Re}(\Psi''_a, \Psi''_b) = \mathrm{Re}(\Psi_a, \Psi_b)$$

Since the absolute value of the inner product (Ψ_a, Ψ_b) is invariant, we must have

$$\mathrm{Im}(\Psi''_a, \Psi''_b) = \pm \mathrm{Im}(\Psi_a, \Psi_b)$$

The + sign implies that

$$(\Psi''_a, \Psi''_b) = (\Psi_a, \Psi_b) \tag{17.3}$$

and

$$(\lambda\Psi)'' = \lambda\Psi'' \tag{17.4}$$

whereas the − sign implies that

$$(\Psi''_a, \Psi''_b) = (\Psi_a, \Psi_b)^* \tag{17.5}$$

and

$$(\lambda\Psi)'' = \lambda^*\Psi'' \tag{17.6}$$

Equation (17.4) expresses the second fundamental property of a *linear* operator [see (9.44)], and from condition (17.3) we infer that in the first case the transformation is *unitary*. Equation (17.6), on the other hand, characterizes an *antilinear* operator [see Eq. (9.46)].

It is easy to see the profound implications of this theorem. State vectors that differ by phase factors represent the same state, and a rephasing transformation has no physical significance. It follows that in studying the symmetry operations of a physical system we may confine ourselves to two simple transformations—those that are linear and those that are antilinear. Any more general symmetry transformation can be supplemented by a phase change and made to fall into one of these two fundamental categories, which are mutually exclusive. Note that the rephasing operation is generally not unitary because different state vectors are generally multiplied by different phase factors.

If the symmetry operation is a rotation, the antilinear case is excluded as a possibility because rotations can be generated continuously from the identity operation, which is inconsistent with complex conjugation of a multiplier. Antilinear transformations are important in describing the behavior of a system under time reversal, a topic to which we will return in Section 17.9.

[2]See Wigner (1959), Appendix to Chapter 20, p. 233; see also Bargmann (1964).

2. ***The Rotation Operator, Angular Momentum, and Conservation Laws.*** The result of the last section is that, if the Euclidean principle of relativity holds, rotations in quantum mechanics are represented by unitary transformations, validating the assumption made in Section 16.3. Although the discussion in Section 16.4 was phrased in terms of spinors describing the state of a spin one-half system, the formalism of rotation operators (or matrices) was in no way dependent on the special nature of the system. The unitary operator that in three-dimensional space rotates a state $|\Psi\rangle$ into $|\Psi'\rangle = U_R|\Psi\rangle$ has the form

$$U_R = \exp\left(-\frac{i}{\hbar}\,\hat{\mathbf{n}}\cdot\mathbf{J}\phi\right) \tag{17.7}$$

and the Hermitian generators of rotation, $\mathbf{J}$, must satisfy the commutation relations (16.44):

$$\boxed{[J_i, J_j] = i\hbar\varepsilon_{ijk}J_k \quad \text{or} \quad \varepsilon_{ijk}J_iJ_j = i\hbar J_k} \tag{17.8}$$

Since the trace of J_i vanishes, the operators (17.7) are unimodular. We know from Chapters 11 and 16 that orbital angular momentum operators $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ and spin angular momentum operators $\mathbf{S}$ satisfy the commutation relations (17.8). They are *realizations* of the generic angular momentum operator $\mathbf{J}$. Planck's constant $\hbar$ was introduced into the definition of the rotation operator in anticipation of the identification of the operators $\mathbf{J}$ as angular momentum for the system on whose states these operators act.

In Section 11.2, we determined the eigenvalues and eigenvectors for all the Hermitian operators $\mathbf{J}$ that satisfy the commutation relations (17.8), as well as the matrices that represent the generalized angular momentum. We now make use of the results obtained there. The eigenvalues of any component of $\mathbf{J}$, such as J_z, are $m\hbar$, and the eigenvalues of $\mathbf{J}^2$ are $j(j+1)\hbar^2$. The quantum number j takes on the values j = nonnegative integer or half-integer, and m is correspondingly integer or half-integer subject to the restriction $-j \le m \le j$. Suppressing any other relevant information that characterizes the state, we denote the eigenvectors by $|jm\rangle$. Since all nonnegative integers are expressible as $2j+1$, the angular momentum algebra can be realized in a vector subspace of any number of dimensions.

In constructing the rotation operator explicitly, we must take into account a further condition that arises because the same rotation R is represented by all operators of the form

$$U_R = \exp\left[-\frac{i}{\hbar}\,\hat{\mathbf{n}}\cdot\mathbf{J}(\phi + 2\pi k)\right] = \exp\left(-\frac{i}{\hbar}\,\hat{\mathbf{n}}\cdot\mathbf{J}\phi\right)\exp\left(-\frac{i}{\hbar}\,\hat{\mathbf{n}}\cdot\mathbf{J}2\pi k\right)$$

where k is an arbitrary integer. The factor $\exp(-2\pi k i\hat{\mathbf{n}}\cdot\mathbf{J}/\hbar)$ is a unimodular operator whose effect on an eigenstate $|jm\rangle$ of $\hat{\mathbf{n}}\cdot\mathbf{J}$ is simply to multiply it by $\exp(-im2\pi k) = (-1)^{2km}$. This is $+1$ for integer km and -1 for half-integer km. If a physical state were represented by a superposition of angular momentum eigenvectors with both integral and half-integral values of j, then since the components with integer j (and m) would remain unchanged while the components with half-integral j (and m) would change sign, application of the rotation operator $\exp(-2\pi k i\hat{\mathbf{n}}\cdot\mathbf{J}/\hbar)$ with k = odd would produce an entirely different state vector. Yet, for systems of point particles such a rotation is geometrically and physically equivalent to no rotation at all and behaves like the identity. In other words, the

mathematical framework allows for state vectors that have no counterpart in physical reality. In ordinary quantum mechanics these states are declared inadmissible by the imposition of a principle. This *superselection rule* has dramatic physical consequences: for instance, particles of half-integral spin cannot be created or destroyed singly, or more generally, in odd numbers (because of the way angular momenta add; see Section 17.5 as well as Section 24.5).

The general theory of angular momentum presents us with all the possible ways in which state vectors may transform under rotation in three dimensions. It does not, of course tell us which of these possibilities are realized in nature. We have already encountered two actual and important examples: the orbital angular momentum, $\mathbf{J} = \mathbf{L}$, and the spin angular momentum, $\mathbf{J} = \mathbf{S} = \hbar\boldsymbol{\sigma}/2$, of electrons, protons, neutrons, quarks, and so on. Both of these vector operators satisfy the commutation relations for angular momentum. They correspond to the values $j = l = 0, 1, 2, \ldots$ and $j = s = 1/2$, respectively. Generalizing these notions, we now identify as *angular momentum* any observable that is represented by a generator $\mathbf{J}$ (in units of $\hbar$) of infinitesimal rotations. In order to apply the theory, we must know something about the nature of the particular physical system under consideration. We must know the observables that describe it and how they behave under rotation. Thus, in the case of orbital angular momentum (Chapter 11), we were dealing with the transformation of a function ψ of the position coordinates x, y, z, or r, φ, θ, and we were led to the study of spherical harmonics. In the case of the spin (Chapter 16), we deduced the behavior of two-component spinors under rotation from the physical connection between the intrinsic angular momentum and magnetic moment, and from the vectorial character of these quantities. Other, more complex examples of angular momentum will appear shortly.

It should be stressed that three-dimensional Euclidean space, with its own topology, underlies the theory with which we are concerned here. Quantum systems of lower or higher dimensionality may require qualitatively different treatments. For example, a system that is confined to two dimensions may have structural characteristics that allow a physical distinction between $2\pi k$ rotations—the integer *winding number* k being an appropriate index. Even in three dimensions, we can conceive of idealized physical systems other than point particles (e.g., perfect solid or rigid bodies) for which it is meaningful to distinguish between odd and even multiples of 2π rotations.[3] We will return to this point in Section 17.4.

A symmetry principle like the Euclidean principle of relativity not only circumscribes the geometric structure of quantum mechanics, but also has important dynamical consequences, notably certain conservation laws, by which the theory can be tested experimentally. Although the same ideas are applicable to almost any symmetry property, to be explicit we will frame the discussion in terms of *conservation of angular momentum* which is a result of *rotational invariance*.

We assume that the dynamical system under consideration is characterized by a time-independent Hamiltonian, and in the Schrödinger picture evolves in time from its initial state $|\Psi(0)\rangle$ according to the equation of motion,

$$i\hbar \frac{d}{dt} |\Psi(t)\rangle = H(\alpha)|\Psi(t)\rangle \tag{17.9}$$

The Hamiltonian depends on the dynamical variables that describe the system, and by adding the parameter α we have explicitly allowed for the possibility that the

[3]Merzbacher (1962).

system may be acted upon by external forces, constraints, or other agents that are not part of the system itself. The division of the world into "the system" and "the environment" in which the system finds itself is certainly arbitrary. But if the cut between the system and its surroundings is made suitably, the system may often be described to a highly accurate approximation by neglecting its back action on the "rest of the world." It is in this spirit that the parameter α symbolizes the external fields acting on what, by an arbitrary but appropriate choice, we have delineated as the dynamical system under consideration.

We have seen that a rotation induces a unitary transformation U_R of the state vector of the system. If we insist that the external fields and constraints also participate in the same rotation, a new Hamiltonian $H(\alpha_R)$ is generated. The Euclidean principle of relativity in conjunction with the principle of causality asserts that, if the dynamical system and the external fields acting on it are rotated together, the two arrangements obtained from each other by rotation must be equivalent *at all times*, and $U_R|\Psi(t)\rangle$ must be a solution of

$$i\hbar \frac{d}{dt} U_R|\Psi(t)\rangle = H(\alpha_R)U_R|\Psi(t)\rangle \tag{17.10}$$

If the symmetry transformation itself is time-independent, comparison of (17.10) with (17.9) yields the important connection

$$H(\alpha_R)U_R = U_R H(\alpha) \tag{17.11}$$

If the symmetry transformation is time-dependent, compatibility of Eqs. (17.9) and (17.10) requires

$$H(\alpha_R) = U_R H(\alpha)U_R^\dagger + i\hbar \frac{\partial U_R}{\partial t} U_R^\dagger \tag{17.12}$$

It frequently happens that the effect of the external parameters on the system is invariant under rotation. In mathematical terms, we then have the equality

$$H(\alpha_R) = H(\alpha) \tag{17.13}$$

Hence, if the symmetry operator is time-independent, U_R commutes with H. Since, according to (17.7), U_R is a function of the Hermitian operator $\mathbf{J}$, the latter becomes a constant of the motion, as defined in Chapter 14. Indeed, the present discussion parallels that of Section 14.4, where the connection between invariance properties and conservation laws was discussed in general terms. Conservation of angular momentum is thus seen to be a direct consequence of invariance under all rotations.

As an important special case, the condition (17.13) obviously applies to an *isolated* system, which does not depend on any external parameters. We thus see that the isotropy of space, as expressed by the Euclidean principle of relativity, requires that the total angular momentum $\mathbf{J}$ of an isolated system be a constant of the motion.

Frequently, certain parts and variables of a system can be subjected separately and independently to a rotation. For example, the spin of a particle can be rotated independently of its position coordinates. In the formalism, this independence appears as the mutual commutivity of the operators $\mathbf{S}$ and $\mathbf{L}$ which describe rotations of the spin and position coordinates, respectively. If the Hamiltonian is such that no correlations are introduced between these two kinds of variables as the system evolves, then they may be regarded as dynamical variables of two separate subsys-

tems. In this case, invariance under rotation implies that both **S** and **L** commute separately with the Hamiltonian and that each is a constant of the motion. The nonrelativistic Hamiltonian of a particle with spin moving in a central-force field couples **L** with **S**, and as we saw in Section 16.4, includes a *spin-orbit* interaction term proportional to $\mathbf{L} \cdot \mathbf{S}$:

$$H = \frac{\mathbf{p}^2}{2m} + V(r) + W(r)\mathbf{L} \cdot \mathbf{S}$$

If the $\mathbf{L} \cdot \mathbf{S}$ term, which correlates spin and orbital motion, can be neglected in a first approximation, the zero-order Hamiltonian commutes with both **L** and **S**, and both of these are thus approximate constants of the motion. However, only the total angular momentum, $\mathbf{J} = \mathbf{L} + \mathbf{S}$, is rigorously conserved by the full Hamiltonian. We will see in Chapter 24 that in the relativistic theory of the electron even the free particle Hamiltonian does not commute with **L** or **S**.

Exercise 17.1. Discuss the rotational symmetry properties of a two-particle system, with its Hamiltonian,

$$H = \frac{\mathbf{p}_1^2}{2m_1} + \frac{\mathbf{p}_2^2}{2m_2} + V(|\mathbf{r}_1 - \mathbf{r}_2|)$$

(See Section 15.4.) Recall the expression (15.95) for the total angular momentum in terms of relative coordinates and the coordinates of the center of mass, and show that if the reduced mass is used, the standard treatment of the central-force problem in Chapters 11 and 12 properly accounts for the exchange of angular momentum between the two particles.

Exercise 17.2. How much rotational symmetry does a system possess, which contains a spinless charged particle moving in a central field and a uniform static magnetic field? What observable is conserved?

3. ***Symmetry Groups and Group Representations.*** Because of the paramount importance of rotational symmetry, the preceding sections of this chapter were devoted to a study of rotations in quantum mechanics. However, rotations are but one type of many symmetry operations that play a role in physics. It is worthwhile to introduce the general notion of a *group* in this section, because symmetry operations are usually elements of certain groups, and group theory classifies and analyzes systematically a multitude of different symmetries that appear in nature.

A *group* is a set of distinct elements $a, b, c, \ldots$, subject to the following four postulates:

1. To each ordered pair of elements a, b, of the group belongs a *product* ba (usually not equal to ab), which is also an element of the group. We say that the law of group multiplication or the multiplication table of the group is given. The product of two symmetry operations, ba, is the symmetry operation that is equivalent to the successive application of a and b, performed in that order.

2. $(ab)c = a(bc)$, i.e., the *associative law* holds. Since symmetry operations are usually motions or substitutions, this postulate is automatically satisfied.

3. The group contains an *identity* element e, with the property

$$ea = ae = a \quad \text{for every } a$$

4. Each element has an *inverse*, denoted by a^{-1}, which is also an element of the group and has the property

$$aa^{-1} = a^{-1}a = e$$

All symmetry operations are reversible and thus have inverses.

For example, rotations form a group in which the product ba of two elements b and a is defined as the single rotation that is equivalent to the two successive rotations a and b. By a rotation we mean the mapping of a physical system, or of a Cartesian coordinate frame, into a new physical system or coordinate frame obtainable from the old system by actually rotating it. The term *rotation* is, however, not to be understood as the physical motion that takes the system from one orientation to another. The intervening orientations that a system assumes during the motion are ignored, and two rotations are identified as equal if they lead from the same initial configuration to the same final configuration regardless of the way in which the operation is carried out. In the rotation group, generally, $ab \neq ba$. For instance, two successive rotations by $\pi/2$—one about the x axis and the other about the y axis—do not lead to the same overall rotation when performed in reverse order. The operation "no rotation" is the identity element for the rotation group.

Exercise 17.3. Use Eq. (16.62) to calculate the direction of the axis of rotation and the rotation angle for the compound rotation obtained by two successive rotations by $\pi/2$ about the x and y axes, respectively. Show that the result is different depending on the order in which the two rotations are performed. Convince yourself of the correctness of this result by rotating this book successively by 90° about two perpendicular axes.

Exercise 17.4. Show that the three Pauli spin matrices, σ_x, σ_y, σ_z, supplemented by the identity I do not constitute a group under matrix multiplication, but that if these matrices are multiplied by ± 1 and $\pm i$ a set of 16 matrices is obtained which meets the conditions for a group. Construct the group multiplication table.

A symmetry operation a transforms the state $|\Psi\rangle$ of a system into a state $|\Psi_a\rangle$. It was shown in Section 17.1 that under quite general conditions this transformation may be assumed to be either unitary, $|\Psi_a\rangle = U_a|\Psi\rangle$, or antilinear.

We assume here that the symmetry operations of interest belong to a group called a *symmetry group* of the system, which induces unitary linear transformations on the state vectors such that, if a and b are two elements of the group,

$$\boxed{U_{ba} = U_b U_a} \tag{17.14}$$

When (17.14) is translated into a matrix equation by introducing a complete set of basis vectors in the vector space of $|\Psi\rangle$, each element a of the group becomes associated with a matrix $D(a)$ such that

$$\boxed{D(ab) = D(a)D(b)} \tag{17.15}$$

That is, the matrices have the same multiplication table as the group elements to which they correspond. The set of matrices $D(a)$ is said to constitute a (matrix) *representation* of the group. Thus far in this book the term *representation* has been used mainly to describe a basis in an abstract vector space. In this chapter, the same

term will be used for the more specific *group representation*. The context usually establishes the intended meaning, and misunderstandings are unlikely to occur.

A change of basis changes the matrices of a representation according to the relation

$$\overline{D}(a) = S^{-1}D(a)S \tag{17.16}$$

as discussed in Section 9.5. From a group theoretical point of view, two representations that can be transformed into each other by a similarity transformation S are not really different, because the matrices $\overline{D}(a)$ obey the same multiplication rule (17.15) as the matrices $D(a)$. The two representations are called *equivalent*, and a transformation (17.16) is known as an *equivalence transformation*. Two representations are *inequivalent* if there is *no* transformation matrix S that takes one into the other. Since the operators U_a were assumed to be unitary, the representation matrices are also unitary if the basis is orthonormal. In the following, all representations $D(a)$ and all transformations S will be assumed in unitary form.

By a judicious choice of the orthonormal basis, we can usually reduce a given group representation to block structure, such that all the matrices $D(a)$ of the representation simultaneously break up into direct sums of smaller matrices arrayed along the diagonal:

$$D(a) = \begin{pmatrix} D_1(a) & 0 & 0 & . \\ 0 & D_2(a) & 0 & . \\ 0 & 0 & D_3(a) & . \\ . & . & . & . \end{pmatrix} \tag{17.17}$$

We suppose that each matrix of the representation acquires the same kind of block structure. If n is the dimension of D, each block $D_1, D_2, \ldots$ is a matrix of dimension $n_1, n_2, \ldots$, with $n_1 + n_2 + \ldots = n$. It is then obvious that the matrices D_1 by themselves constitute an n_1-dimensional representation. Similarly, D_2 gives an n_2-dimensional representation, and D_i is an n_i-dimensional representation. The original representation has thus been *reduced* to a number of simpler representations. The state vector (Hilbert) space has been similarly decomposed into a set of subspaces such that the unitary operators U_a reduce to a direct sum

$$U_a = U_a^{(1)} \oplus U_a^{(2)} \oplus \ldots$$

with each set of operators $U_a^{(1)}$ satisfying the group property (17.14).

If no basis can be found to reduce all D matrices of the representation simultaneously to block structure, the representation is said to be *irreducible*. Otherwise it is called *reducible*. Apart from an unspecified equivalence transformation, the decomposition into irreducible representations is unique. (If two essentially different decompositions into irreducible subspaces were possible, the subspace formed by the intersection of two irreducible subspaces would itself have to be irreducible, contrary to the assumption.) There is therefore a definite sense in stating the irreducible representations (or *irreps* in technical jargon), which make up a given reducible representation. Some of the irreducible matrix representations may occur more than once.

It is clearly sufficient to study all *inequivalent* irreducible representations of a group; all reducible representations are then built up from these. Group theory provides the rules for constructing systematically all irreducible representations from the group multiplication table. Which of these are relevant in the analysis of a par-

ticular physical system depends on the structure of the state vector space of the system.

The usefulness of the theory of group representations for quantum mechanics and notably the idea of irreducibility will come into sharper focus if the Schrödinger equation $H|\Psi\rangle = E|\Psi\rangle$ is considered. A symmetry operation, applied to all eigenstates, must leave the Schrödinger equation invariant so that the energies and transition amplitudes of the system are unaltered. The criterion for the invariance of the Schrödinger equation under the operations of the group is that the Hamiltonian commute with U_a for every element a of the group:

$$\boxed{[H, U_a] = 0} \tag{17.18}$$

In Section 17.2 the same condition was obtained by applying symmetry requirements to the dynamical equations, and the connection between conservation laws and constants of the motion was established. By studying the symmetry group, which gives rise to these constants of the motion, we can learn much about the eigenvalue spectrum of the Hamiltonian and the corresponding eigenfunctions.

If E is an n-fold degenerate eigenvalue of the Hamiltonian,

$$H|k\rangle = E|k\rangle \qquad (k = 1, 2, \ldots, n) \tag{17.19}$$

the degenerate eigenvectors $|k\rangle$ span a subspace and, owing to (17.18),

$$HU_a|k\rangle = U_aH|k\rangle = EU_a|k\rangle$$

Thus, if $|k\rangle$ is an eigenvector of H corresponding to the eigenvalue E, then $U_a|k\rangle$ is also an eigenvector and belongs to the same eigenvalue. Hence, it must be equal to a linear combination of the degenerate eigenvectors,

$$U_a|k\rangle = \sum_{j=1}^{n} |j\rangle D_{jk}(a) \tag{17.20}$$

where the $D_{jk}(a)$ are complex coefficients that depend on the group element. Repeated application of symmetry operations gives

$$U_bU_a|k\rangle = \sum_{j=1}^{n} U_b|j\rangle D_{jk}(a) = \sum_{j=1}^{n}\sum_{\ell=1}^{n} |\ell\rangle D_{\ell j}(b)D_{jk}(a) \tag{17.21}$$

But we also have

$$U_{ba}|k\rangle = \sum_{\ell=1}^{n} |\ell\rangle D_{\ell k}(ba) \tag{17.22}$$

By the assumption of (17.14), the left-hand sides of (17.21) and (17.22) are identical. Hence, comparing the right-hand sides, it follows that

$$\boxed{D_{\ell k}(ba) = \sum_{j=1}^{n} D_{\ell j}(b)D_{jk}(a)} \tag{17.23}$$

This is the central equation of the theory. It shows that the coefficients $D_{jk}(a)$ define a unitary representation of the symmetry group. If a vector lies entirely in the n-dimensional subspace spanned by the n degenerate eigenvectors of H, the operations of the group transform this vector into another vector lying entirely in the same subspace, i.e., the symmetry operations leave the subspace *invariant*.

Since any representation D of the symmetry group can be characterized by the irreducible representations that it contains, it is possible to classify the stationary states of a system by the irreducible representations to which the eigenvectors of H belong. A partial determination of these eigenvectors can thereby be achieved. The labels of the irreducible representations to which an energy eigenvalue belongs are the quantum numbers of the stationary state.

These considerations exhibit the mutual relationship between group theory and quantum mechanics: The eigenstates of (17.19) generate representations of the symmetry group of the system described by H. Conversely, knowledge of the appropriate symmetry groups and their irreducible representations can aid considerably in solving the Schrödinger equation for a complex system. If all symmetries of a system are recognized, much can be inferred about the general character of the eigenvalue spectrum and the nature of the eigenstates. The use of group theoretical methods, applied openly or covertly, is indispensable in the study of the structure and the spectra of complex nuclei, atoms, molecules, and solids. The Schrödinger equation for such many-body systems is almost hopelessly complicated, but its complexity can be reduced and a great deal of information inferred from the various symmetry properties, such as translational and rotational symmetry, reflection symmetry, and symmetry under exchange of identical particles.

The observation of symmetric patterns and structures, as in crystals and molecules, suggest the use of *finite* groups, i.e., transformation groups with a finite number of elements. Often details about forces and interactions are unknown, or the theory is otherwise mathematically intractable, as in the case of strongly interacting elementary particles (quantum chromodynamics). However, the dynamical laws are understood to be subject to certain general symmetry principles, such as invariance under rotations, Lorentz transformations, charge conjugation, interchange of identical particles, ''rotation'' in isospin space, and, at least approximately, the operation of the group *SU(3)* in a three-dimensional vector space corresponding to intrinsic degrees of freedom. The irreducible representations of the groups which correspond to these symmetries provide us with the basic quantum numbers and selection rules for the system, allowing classification of states, without requiring us to solve the complete dynamical theory.

The utility of group representations in quantum mechanics is not restricted to systems whose Hamiltonian exhibits perfect invariance under certain symmetry transformations. Although the symmetry may only be approximate and the degeneracy of the energy eigenstates can be broken to a considerable degree, the states may still form so-called *multiplets*, which under the action of the group operations transform according to an irreducible representation. Thus, a set of these states can be characterized by the labels, or ''good quantum numbers,'' of the representation with which they are identified or to which they are said to ''belong.'' An understanding of the pertinent symmetry groups for a given system not only offers considerable physical insight, but as we will see in Sections 17.7 and 17.8, can also simplify the technical calculation of important matrix elements in theory. For example, the orbital angular momentum operator $\mathbf{L}^2$, whose eigenvalues $\ell(\ell + 1)\hbar^2$ will be seen in the next section to label the irreducible representations of the group of rotations of the position coordinates alone, commutes with the Hamiltonian of an electron in an atom that is exposed to an external uniform magnetic field (but not an electric field) and, also with the spin-orbit interaction. Therefore, the quantum numbers ℓ characterize multiplets in atomic spectra.

The *continuous* groups that are of particular interest in quantum mechanics are various groups of linear transformations, conveniently expressible in terms of matrices, which—like the rotations in Section 16.3—can be generated by integration from infinitesimal transformations. The elements of such *Lie groups* are specified by a minimal set of independent, continuously variable real parameters—three, in the case of ordinary rotations—and the corresponding generators. The algebra of the generators, interpreted as (Hermitian) matrices or operators with their characteristic commutation relations, is the mathematical tool for obtaining the irreducible representations of these groups. An important category of continuous groups are the so-called *semi-simple Lie groups*, which are of particular physical relevance and also have attractive mathematical properties (analogous to the richness of the theory of *analytic* functions). Examples of important semi-simple Lie groups are the n-dimensional groups $O(n)$ of the real orthogonal matrices; their subgroups $SO(n)$ composed of those matrices that have a determinant equal to $+1$ (with the letter S standing for *special*); and the special unitary groups $SU(n)$. In Chapter 16 we saw that the rotation group $R(3)$, which is isomorphic to $O(3)$, is intimately related to $SU(2)$. This connection will be developed further in Section 17.4.

Exercise 17.5. An n-dimensional proper real orthogonal matrix $SO(n)$, i.e., a matrix whose inverse equals its transpose and which has determinant equal to unity, can be expressed as $\exp(X)$, where X is a real-valued skew-symmetric matrix. Show that the group of *special orthogonal* matrices $SO(n)$ has $n(n-1)/2$ independent real parameters. (Compare Exercise 16.3.) Similarly, show that the group $SU(n)$ has $n^2 - 1$ real parameters.

4. ***The Representations of the Rotation Group.*** The representations of the rotation group $R(3)$, which is our prime example, are generated from the rotation operator (17.7),

$$U_R = \exp\left(-\frac{i}{\hbar}\,\hat{\mathbf{n}} \cdot \mathbf{J}\phi\right) \tag{17.7}$$

The rotations in real three-dimensional space are characterized by three independent parameters, and correspondingly there are three Hermitian generators, J_x, J_y, J_z, of infinitesimal rotations. They satisfy the standard commutation relations for the components of angular momentum, (17.8). The eigenvectors of one and only one of them, usually chosen to be J_z, can serve as basis vectors of the representation, thus diagonalizing J_z. In other Lie groups, the maximum number of generators that commute with each other and can be simultaneously diagonalized is usually greater than one. This number is called the *rank* of the group. The groups $O(4)$ and $SU(4)$ both have rank two.

The central theorem on group representations is *Schur's* (second) *Lemma*:

If the matrices $D(a)$ form an irreducible representation of a group and if a matrix M commutes with all $D(a)$,

$$[M, D(a)] = 0 \quad \text{for every } a \tag{17.24}$$

then M is a multiple of the identity matrix.

This result encourages us to look for a normal operator C which commutes with all the generators, and thus with every element, of the given symmetry group. The aim

is to find an operator C whose eigenvalues can be used to characterize and classify the irreducible representations of the group. If the operator has distinct eigenvalues $c_1, c_2, \ldots$, it can by a suitable choice of the basis vectors be represented in diagonal form as

$$C = \begin{pmatrix} c_1 I & 0 & 0 & 0 \\ 0 & c_2 I & 0 & 0 \\ 0 & 0 & . & 0 \\ 0 & 0 & 0 & . \end{pmatrix} \tag{17.25}$$

where the identity matrices have dimensions corresponding to the multiplicities (degeneracies) of the eigenvalues $c_1, c_2, \ldots$, respectively. In this basis, all matrices representing the group elements are reduced to block structure as in (17.17).

Exercise 17.6. Show that if $D(a)$ commutes with C, the matrix elements of $D(a)$ which connect basis vectors that belong to two distinct eigenvalues of C (e.g., $c_1 \neq c_2$) are zero.

If the reduction to block structure produces irreducible representations of the group, the eigenvalues $c_1, c_2, \ldots$, of the operator C are convenient numbers (*quantum numbers*) which may serve as labels classifying the irreducible representations. Since, depending on the nature of the vector (Hilbert) space of the physical system under consideration, any particular irreducible representation may appear repeatedly, additional labels α are usually needed to identify the basis vectors completely and uniquely.

In general, more than one operator C is needed so that the eigenvalues will provide a complete characterization of all irreducible representations of a symmetry group. For the important class of the *semi-simple Lie groups*, the rank of the group is equal to the number of mutually commuting independent *Casimir* operators sufficient to characterize all irreducible representations. For the rotation group $R(3)$ of rank one, the Casimir operator, which commutes with every component of $\mathbf{J}$, is chosen to be the familiar square magnitude of the angular momentum operator,

$$\mathbf{J}^2 = J_x^2 + J_y^2 + J_z^2 = \frac{1}{2}(J_+J_- + J_-J_+) + J_z^2 \tag{17.26}$$

with eigenvalues $j(j+1)\hbar^2$. The nonnegative integral or half-integral angular momentum quantum number j fully characterizes the irreducible representations of the rotation group in three dimensions. (In four dimensions, two Casimir operators and their quantum numbers are needed.) As is customary, we choose the common eigenvectors of $\mathbf{J}^2$ and J_z as our basis vectors and denote them by $|\alpha jm\rangle$. Since the quantum numbers α are entirely unaffected by rotation, they may be omitted in some of the subsequent formulas, but they will be reintroduced whenever they are needed.

From Section 11.2 we copy the fundamental equations:

$$\begin{aligned} J_z|j\,m\rangle &= m\hbar|j\,m\rangle \\ J_+|j\,m\rangle &= \sqrt{(j-m)(j+m+1)}\hbar|j\,m+1\rangle \\ J_-|j\,m\rangle &= \sqrt{(j+m)(j-m+1)}\hbar|j\,m-1\rangle \end{aligned} \tag{17.27}$$

The vector space of the system at hand thus decomposes into a number of disjoint $(2j+1)$-dimensional subspaces whose intersection is the null vector and which are

invariant under rotation. An arbitrary rotation is represented in one of the subspaces by the matrix

$$D^{(j)}_{m'm}(R) = \langle jm' | U_R | jm \rangle = \langle jm' | \exp\left(-\frac{i}{\hbar} \hat{\mathbf{n}} \cdot \mathbf{J}\phi \right) | jm \rangle \tag{17.28}$$

Owing to (9.63) and (9.64), the representation matrix can be written as

$$D^{(j)}(R) = \exp\left(-\frac{i}{\hbar} \hat{\mathbf{n}} \cdot \mathbf{J}\phi \right) \tag{17.29}$$

where now $\mathbf{J}$ stands for the *matrix* whose elements are $\langle jm' | \mathbf{J} | jm \rangle$.

The simplicity of (17.28) and (17.29) is deceptive, for the components of $\mathbf{J}$ other than J_z are represented by nondiagonal matrices, and the detailed dependence of the matrix elements of $D^{(j)}(R)$ on the quantum numbers and on the rotation parameters $\hat{\mathbf{n}}$ and ϕ is quite complicated. For small values of j, we can make use of the formula (10.31) to construct the rotation matrices in terms of the first $2j$ powers of the matrix $\hat{\mathbf{n}} \cdot \mathbf{J}$:

$$\boxed{D^{(j)}(R) = \sum_{k=-j}^{j} e^{-ik\phi} \prod_{\substack{k'=-j \\ k' \neq k}}^{j} \frac{k'\hbar - \hat{\mathbf{n}} \cdot \mathbf{J}}{(k' - k)\hbar}} \tag{17.30}$$

Exercise 17.7. Using (17.30) and the explicit form of the angular momentum matrices, work out the rotation matrices for $j = 0$, $1/2$, and 1.

Exceptional simplification occurs for the subgroup of two-dimensional rotations about the "axis of quantization," the z axis if J_z is chosen to be diagonal. For such special rotations

$$D^{(j)}_{m'm}(\phi) = e^{-im\phi} \delta_{m',m} \tag{17.31}$$

The representation matrices are also simple for infinitesimal rotations, i.e., when $\phi \ll 1$. In this case, by expanding (17.28) to first order in $\boldsymbol{\varepsilon} = \phi\, \hat{\mathbf{n}}$, and applying (17.27), we get

$$\boxed{\begin{aligned} D^{(j)}_{m'm}(\varepsilon\hat{\mathbf{n}}) = \delta_{m',m} &- \frac{i\varepsilon_x + \varepsilon_y}{2} \sqrt{(j - m)(j + m + 1)}\, \delta_{m',m+1} \\ &- \frac{i\varepsilon_x - \varepsilon_y}{2} \sqrt{(j + m)(j - m + 1)}\, \delta_{m',m-1} - i\varepsilon_z\, m\delta_{m',m} \end{aligned}} \tag{17.32}$$

A further diagonalization of all rotation matrices is clearly impossible, since diagonalizing, for instance, J_x would necessarily undo the diagonalization of J_z. Hence, for a given value of j, the irreducibility of the representations just obtained is verified. All continuous unitary irreducible representations of the rotation group are obtained by allowing j to assume the values $j = 0, 1/2, 1, 3/2, \ldots$.

Again we must comment on the half-integral values of j. From the point of view of infinitesimal rotations, these are on the same footing as the integral values, as is confirmed by (17.32). However, when finite rotations are considered, an important distinction arises, for upon rotation about a fixed angle by $\phi = 2\pi$, (17.31) shows that the matrix $D^{(j)}(2\pi) = -1$ is obtained. Yet, the system has been restored to its original configuration, and the resulting rotation is equivalent to the operation "no

rotation at all." Hence, for half-integral values of j, a *double-valued representation* of the group is produced: to any rotation R correspond two distinct matrices differing by a sign. By the strictest definition, these $D^{(j)}$ matrices do not constitute a representation of the rotation group. Instead, as we saw in Chapter 16, they represent the *unitary unimodular* group in two dimensions, $SU(2)$. This group is said to be the *universal covering group* of the rotation group. With this understanding, we can safely regard them as double-valued representations of the rotation group. They appear naturally in the reduction of the matrix representing the rotation operator U_R. Owing to the usual phase ambiguities of quantum mechanics, they have their place in the theory, provided that U_R contains *only* double-valued representations.

Exercise 17.8. Show that the representation matrices $D^{(1/2)}(R)$ are equivalent to the matrices U_R in Eq. (16.62).

The general expression for the representation matrices $D(R)$ is most conveniently obtained if the rotation is parametrized in terms of the three Euler angles, α, β, γ, rather than in terms of $\hat{\mathbf{n}}$ and ϕ. As indicated in Figure 17.1, α is the angle of a rotation about the z axis of the fixed frame-of-reference. This rotation carries the y axis into the y'' direction, and the second rotation by the angle β takes place around this *nodal line*. This rotation carries the z axis into the z' direction, defined by the azimuthal angle α and the polar angle β. The final rotation is by the angle γ about the new z' axis. Hence, the complete rotation operator is

$$U_R \doteq \exp\left(-\frac{i}{\hbar}\,\gamma\hat{\mathbf{z}}'\cdot\mathbf{J}\right)\exp\left(-\frac{i}{\hbar}\,\beta\hat{\mathbf{y}}''\cdot\mathbf{J}\right)\exp\left(-\frac{i}{\hbar}\,\alpha\hat{\mathbf{z}}\cdot\mathbf{J}\right) \qquad (17.33)$$

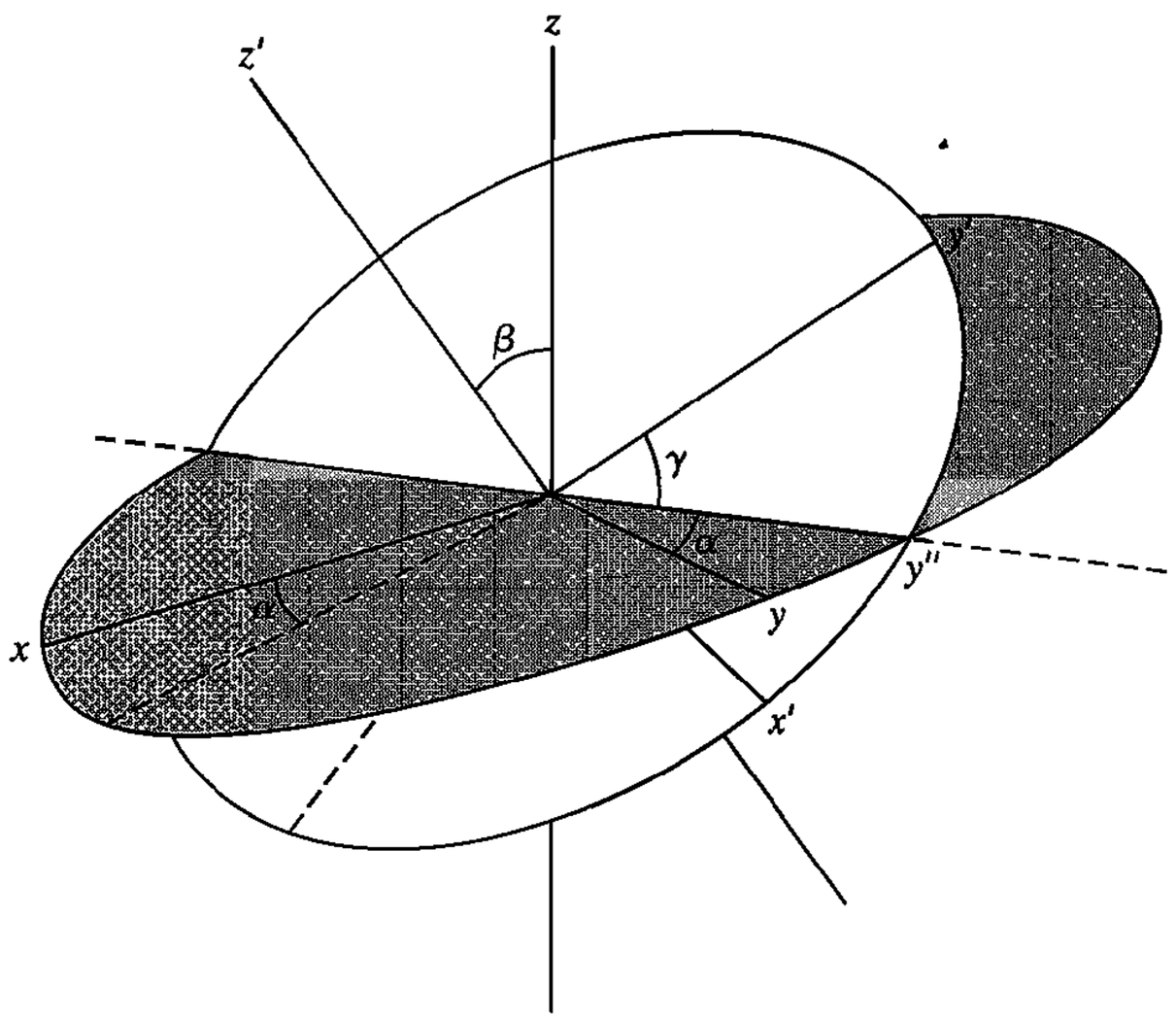


Figure 17.1. Euler angles α (rotation about the z axis), β (rotation about the y'' axis or nodal line), and γ (rotation about the z' axis), transforming xyz into $x'y'z'$. All intermediate axes are shown as dashed lines. The Euler angles α and β are the spherical polar coordinates of the z' axis with respect to the xyz coordinate system (see Figure 11.3).

It is preferable to transform the right-hand side into a product of rotations about the fixed axes $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, $\hat{\mathbf{z}}$. By inspection of Figure 17.1, it is easy to see that

$$e^{-(i/\hbar)\beta\hat{\mathbf{y}}''\cdot\mathbf{J}} = e^{-(i/\hbar)\alpha\hat{\mathbf{z}}\cdot\mathbf{J}}e^{-(i/\hbar)\beta\hat{\mathbf{y}}\cdot\mathbf{J}}e^{(i/\hbar)\alpha\hat{\mathbf{z}}\cdot\mathbf{J}} \tag{17.34}$$

and

$$e^{-(i/\hbar)\gamma\hat{\mathbf{z}}'\cdot\mathbf{J}} = e^{-(i/\hbar)\beta\hat{\mathbf{y}}''\cdot\mathbf{J}}e^{-(i/\hbar)\gamma\hat{\mathbf{z}}\cdot\mathbf{J}}e^{(i/\hbar)\beta\hat{\mathbf{y}}''\cdot\mathbf{J}} \tag{17.35}$$

Substitution of (17.34) and (17.35) into (17.33) yields the simple result

$$U_R = e^{-(i/\hbar)\alpha\hat{\mathbf{z}}\cdot\mathbf{J}}e^{-(i/\hbar)\beta\hat{\mathbf{y}}\cdot\mathbf{J}}e^{-(i/\hbar)\gamma\hat{\mathbf{z}}\cdot\mathbf{J}} = e^{-(i/\hbar)\alpha J_z}e^{-(i/\hbar)\beta J_y}e^{-(i/\hbar)\gamma J_z} \tag{17.36}$$

The matrix elements of the irreducible representation characterized by j are then

$$\begin{aligned} D^{(j)}_{m'm}(\alpha, \beta, \gamma) &= \langle jm'|e^{-(i/\hbar)\alpha J_z}\, e^{-(i/\hbar)\beta J_y}\, e^{-(i/\hbar)\gamma J_z}|jm\rangle \\ &= e^{-i\alpha m'}e^{-i\gamma m}\langle jm'|e^{-(i/\hbar)\beta J_y}|jm\rangle \\ &= e^{-i\alpha m'}e^{-i\gamma m}\langle jm'|\exp\left(-\beta\frac{J_+ - J_-}{2\hbar}\right)|jm\rangle \end{aligned} \tag{17.37}$$

We leave the details of working out explicit formulas for the representation matrices for general values of j to the myriad of specialized treatises on angular momentum[4] and the rotation group, but we note two particularly useful symmetry relations for the matrix elements:

$$D^{(j)}_{m'm}{}^*(\alpha, \beta, \gamma) = D^{(j)}_{mm'}(-\gamma, -\beta, -\alpha) \tag{17.38}$$

$$D^{(j)}_{m'm}{}^*(\alpha, \beta, \gamma) = (-1)^{m-m'}D^{(j)}_{-m',-m}(\alpha, \beta, \gamma) \tag{17.39}$$

Exercise 17.9. Using the unitary property of U_R, show that the inverse rotation R^{-1} is generated either by first rotating by $-\gamma$ around $\hat{\mathbf{z}}'$, then by $-\beta$ around $\hat{\mathbf{y}}'$, and finally by $-\alpha$ around $\hat{\mathbf{z}}$, or equivalently first rotating by $-\alpha$ around $\hat{\mathbf{z}}$, then by $-\beta$ around $\hat{\mathbf{y}}$, and finally by $-\gamma$ again around $\hat{\mathbf{z}}$. Prove the symmetry relation (17.38). [The relation (17.39) is related to time reversal symmetry and can be proved at the end of the chapter.]

Exercise 17.10. Work out the rotation matrices in terms of Euler angles for $j = 0$, $1/2$, and 1, and compare the results with the results of Exercise 17.7.

If j is integral, $j = \ell$, we can establish the connection between the rotation matrices and the eigenfunctions of orbital angular momentum. Expressed in terms of a (single particle) coordinate basis, the eigenvectors of $J_z(L_z)$ and $\mathbf{J}^2(\mathbf{L}^2)$ can be chosen to be the spherical harmonics:

$$\langle\mathbf{r}|\ell m\rangle = Y^m_\ell(\theta, \varphi) \tag{17.40}$$

A rotation transforms an eigenvector $|\ell m\rangle$ of L_z into an eigenvector $|\ell m\rangle'$ of $L_{z'}$, with the same values of ℓ and m, z' being the new axis obtained by the rotation from the z axis. Hence,

$$\boxed{|\ell m\rangle' = U_R|\ell m\rangle = \sum_{m'}|\ell m'\rangle D^{(\ell)}_{m'm}(R)} \tag{17.41}$$

[4]Rose (1957), Brink and Satchler (1968), Biedenharn and Louck (1981), Thompson (1994). See also Feagin (1993), Chapters 16–19, for a computer-oriented approach.

and, if this is multiplied on the left by $|\mathbf{r}\rangle$,

$$\boxed{Y_\ell^m(\theta', \varphi') = \sum_{m'=-\ell}^{\ell} Y_\ell^{m'}(\theta, \varphi) D_{m'm}^{(\ell)}(R)} \tag{17.42}$$

Here θ, φ, and θ', φ' are coordinates of the *same* physical point.

Equation (17.42) can be inverted by using the unitary property of the representation:

$$Y_\ell^m(\theta, \varphi) = \sum_{m'=-\ell}^{\ell} D_{mm'}^{(\ell)}{}^*(R) Y_\ell^{m'}(\theta', \varphi') \tag{17.43}$$

Consider in particular a point on the new z' axis, $\theta' = 0$. Using (11.98), we find the connection between the D matrices and spherical harmonics:

$$\boxed{D_{m0}^{(\ell)}{}^*(R) = \sqrt{\frac{4\pi}{2\ell + 1}} Y_\ell^m(\beta, \alpha)} \tag{17.44}$$

where β, α are the spherical polar coordinates of the new z' axis in the old coordinate system (see Figure 17.1). If this is substituted into (17.42) for $m = 0$, we get

$$Y_\ell^0(\theta', \varphi') = \sqrt{\frac{4\pi}{2\ell + 1}} \sum_{m=-\ell}^{\ell} Y_\ell^m(\theta, \varphi) Y_\ell^m{}^*(\beta, \alpha) \tag{17.45}$$

By (11.99), we finally retrieve the addition theorem for spherical harmonics, (11.100):

$$P_\ell(\cos\theta') = \frac{4\pi}{2\ell + 1} \sum_{m=-\ell}^{\ell} Y_\ell^m{}^*(\beta, \alpha) Y_\ell^m(\theta, \varphi) \tag{17.46}$$

5. ***The Addition of Angular Momenta.*** If two distinct physical systems or two distinct sets of dynamical variables of one system, which are described in two different vector spaces, are merged, the states of the composite system are vectors in the *direct product space* of the two previously separate vector spaces. The mathematical procedure was outlined in Section 15.4. A common rotation of the composite system is represented by the direct product of the rotation operators for each subsystem and may, as an application of Eq. (15.82), be written as

$$\exp\left(-\frac{i}{\hbar}\,\hat{\mathbf{n}}\cdot\mathbf{J}_1\phi\right) \otimes \exp\left(-\frac{i}{\hbar}\,\hat{\mathbf{n}}\cdot\mathbf{J}_2\phi\right) = \exp\left[-\frac{i}{\hbar}\,\hat{\mathbf{n}}\cdot(\mathbf{J}_1 \otimes 1 + 1 \otimes \mathbf{J}_2)\phi\right] \tag{17.47}$$

The operator

$$\mathbf{J} = \mathbf{J}_1 \otimes 1 + 1 \otimes \mathbf{J}_2 \tag{17.48}$$

often written more simply, if less accurately, as

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2 \tag{17.49}$$

is the *total angular momentum* of the composite system. We sometimes say that $\mathbf{J}_1$ and $\mathbf{J}_2$ are *coupled* to the resultant $\mathbf{J}$.

Exercise 17.11. Prove Eq. (17.47) and show that $\mathbf{J}_1 \otimes 1$ commutes with $1 \otimes \mathbf{J}_2$.

An important example of adding commuting angular momentum operators was already encountered in Section 16.4 and again touched upon in Section 17.3, where **L** and **S** were combined into **J**. The present section is devoted to the formal and general solution of adding any two commuting angular momenta.

Since each component of $\mathbf{J}_1$ commutes with each component of $\mathbf{J}_2$, which separately satisfy the usual commutation relations for angular momentum, the total angular momentum **J** also satisfies the angular momentum commutation relations:

$$[J_i, J_j] = i\hbar\varepsilon_{ijk}J_k \tag{17.50}$$

The problem of the addition of two angular momenta consists of obtaining the eigenvalues of J_z and $\mathbf{J}^2$ and their eigenvectors in terms of the direct product of the eigenvectors of J_{1z}, $\mathbf{J}_1^2$ and of J_{2z}, $\mathbf{J}_2^2$. The normalized simultaneous eigenvectors of the four operators $\mathbf{J}_1^2$, $\mathbf{J}_2^2$, J_{1z}, J_{2z}, can be symbolized by the direct product kets,

$$\boxed{|j_1 j_2 m_1 m_2\rangle = |j_1 m_1\rangle \otimes |j_2 m_2\rangle = |j_1 m_1\rangle|j_2 m_2\rangle} \tag{17.51}$$

These constitute a basis in the product space. From this basis we propose to construct the eigenvectors of J_z, $\mathbf{J}^2$ which form a new basis.

The operators $\mathbf{J}_1^2$ and $\mathbf{J}_2^2$ commute with every component of **J**. Hence,

$$[J_z, \mathbf{J}_1^2] = [J_z, \mathbf{J}_2^2] = [\mathbf{J}^2, \mathbf{J}_1^2] = [\mathbf{J}^2, \mathbf{J}_2^2] = 0$$

and the eigenvectors of J_z and $\mathbf{J}^2$ can be required to be simultaneously eigenvectors of $\mathbf{J}_1^2$ and $\mathbf{J}_2^2$ as well. (But $\mathbf{J}^2$ does *not* commute with J_{1z} or J_{2z}!) In the subspace of the simultaneous eigenvectors of $\mathbf{J}_1^2$ and $\mathbf{J}_2^2$ with eigenvalues j_1 and j_2, respectively, we can thus write the transformation equation

$$\boxed{|j_1 j_2 j m\rangle = \sum_{m_1 m_2} |j_1 j_2 m_1 m_2\rangle\langle j_1 j_2 m_1 m_2|j_1 j_2 j m\rangle} \tag{17.52}$$

connecting the two sets of normalized eigenvectors. The summation needs to be carried out only over the two eigenvalues m_1 and m_2, for j_1 and j_2 can be assumed to have fixed values. Thus, the problem of adding angular momenta is the problem of determining the transformation coefficients

$$\langle j_1 j_2 m_1 m_2|j_1 j_2 j m\rangle$$

These elements of the transformation matrix are called *Clebsch-Gordan* or *Wigner coefficients* or *vector-addition* or *vector-coupling coefficients*. When no confusion is likely, we simplify the notation and write

$$\boxed{\langle j_1 j_2 m_1 m_2|j_1 j_2 j m\rangle \equiv \langle m_1 m_2|j m\rangle}$$

for it is understood that j_1 and j_2 are the maximum values of m_1 and m_2.[5] Commas will be inserted between quantum numbers in a Clebsch-Gordan coefficient only if they are needed for clarity. As an abbreviation, we will refer to the Clebsch-Gordan coefficients as C-G coefficients.

[5]For a complete discussion of Wigner coefficients and the many notations that are in use for them and their variants, see Biedenharn and Louck (1981).

If the operator $J_z = J_{z_1} + J_{z_2}$ is applied to (17.52) and if the eigenvalue conditions

$$J_z|j_1 j_2 jm\rangle = m\hbar\,|j_1 j_2 jm\rangle$$
$$(J_{z_1} + J_{z_2})|j_1 j_2 m_1 m_2\rangle = (m_1 + m_2)\hbar\,|j_1 j_2 m_1 m_2\rangle$$

are used, it can immediately be concluded that the m quantum numbers are subject to the *selection rule*:

$$\langle j_1 j_2 m_1 m_2 | j_1 j_2 jm\rangle = 0 \quad \text{if} \quad m \neq m_1 + m_2 \tag{17.53}$$

Applying J_- and J_+ to (17.52), we obtain the following recursion relations for the C-G coefficients:

$$\sqrt{(j \pm m)(j \mp m + 1)}\langle m_1 m_2 | j, m \mp 1\rangle = \sqrt{j_1 \mp m_1)(j_1 \pm m_1 + 1)}\langle m_1 \pm 1, m_2 | jm\rangle + \sqrt{(j_2 \mp m_2)(j_2 \pm m_2 + 1)}\langle m_1, m_2 \pm 1 | jm\rangle \tag{17.54}$$

Exercise 17.12. Derive the recursion relation (17.54), as indicated.

To appreciate the usefulness of these equations, we set $m_1 = j_1$ and $m = j$ in (17.54), using the upper signs. Owing to the selection rule (17.53), for nontrivial results m_2 can have only the value $m_2 = j - j_1 - 1$, and we find that

$$\sqrt{2j}\,\langle j_1, j - j_1 - 1 | j, j - 1\rangle = \sqrt{(j_2 - j + j_1 + 1)(j_2 + j - j_1)}\,\langle j_1, j - j_1 | j\, j\rangle \tag{17.55}$$

Hence, if $\langle j_1, j - j_1 | j\, j\rangle$ is known, $\langle j_1, j - j_1 - 1 | j, j - 1\rangle$ can be determined. From these two C-G coefficients we can then calculate $\langle j_1 - 1, j - j_1 | j, j - 1\rangle$ by again applying (17.54), but this time using the lower signs and setting $m_1 = j_1$, $m_2 = j - j_1$, $m = j - 1$.

Continuing in this manner, we can use the recursion relations (17.54) to give for fixed values of j_1, j_2, and j all the C-G coefficients in terms of just one of them, namely,

$$\langle j_1 j_2 j_1, j - j_1 | j_1 j_2 j\, j\rangle \tag{17.56}$$

The absolute value of this coefficient is determined by normalization (see below). From the selection rule (17.53) we deduce that the C-G coefficient (17.56) is different from zero only if

$$-j_2 \leq j - j_1 \leq j_2$$

this being the range of values of m_2. Hence, j is restricted to the range

$$j_1 - j_2 \leq j \leq j_1 + j_2$$

But since j_1 and j_2 are on a symmetric footing, we could equally well have expressed all C-G coefficients in terms of $\langle j_1 j_2, j - j_2, j_2 | j_1 j_2 j\, j\rangle$ and then have concluded that j must be resticted by the condition

$$j_2 - j_1 \leq j \leq j_1 + j_2$$

The last two inequalities together show that the three angular momentum quantum numbers must satisfy the so-called *triangular condition*

$$\boxed{|j_1 - j_2| \leq j \leq j_1 + j_2} \tag{17.57}$$

i.e., the three nonnegative integer or half-integers j_1, j_2, j must be such that they could constitute the three sides of a triangle. Since $m = m_1 + m_2$ ranges between $-j$ and $+j$, it follows that j can assume only the values

$$\boxed{j = j_1 + j_2, j_1 + j_2 - 1, \ldots, |j_1 - j_2| + 1, |j_1 - j_2|} \tag{17.58}$$

Hence, either all three quantum numbers j_1, j_2, j are integers, or two of them are half-integral and one is an integer.

Exercise 17.13. Show that a symmetry exists among the three quantum numbers j_1, j_2, j, and that in addition to (17.57) they satisfy the equivalent relations

$$|j - j_2| \leq j_1 \leq j + j_2 \quad \text{and} \quad |j - j_1| \leq j_2 \leq j + j_1$$

We observe that for fixed values of j_1 and j_2 (17.52) gives a complete new basis in the $(2j_1 + 1)(2j_2 + 1)$-dimensional vector space spanned by the kets $|j_1 j_2 m_1 m_2\rangle$. Indeed, the new kets $|j_1 j_2 j\ m\rangle$, with j given by the allowed values (17.58), are again $(2j_1 + 1)(2j_2 + 1)$ in number, and being eigenkets of Hermitian operators they are also orthogonal. Since the old and new bases are both normalized to unity, the C-G coefficients must constitute a unitary matrix. Furthermore, it is clear from the recursion relations for the C-G coefficients that all of them are real numbers if one of them, say $\langle j_1 j_2, j_1, j - j_1 | j_1 j_2 j\ j\rangle$, is chosen real. If this is done, the C-G coefficients satisfy the condition

$$\sum_{m_1 m_2} \langle j_1 j_2 m_1 m_2 | j_1 j_2 j\ m\rangle \langle j_1 j_2 m_1 m_2 | j_1 j_2 j' m'\rangle = \delta_{mm'} \delta_{jj'} \tag{17.59}$$

or inversely

$$\sum_{jm} \langle j_1 j_2 m_1 m_2 | j_1 j_2 j m\rangle \langle j_1 j_2 m_1' m_2' | j_1 j_2 j m\rangle = \delta_{m_1 m_1'} \delta_{m_2 m_2'} \tag{17.60}$$

The double sums in these two equations can be substantially simplified by the use of the selection rule (17.53). Condition (17.59) in conjunction with the recursion relations determines all C-G coefficients except for a sign. It is conventional to choose the sign by demanding that the C-G coefficient $\langle j_1 j_2, j_1, j - j_1 | j_1 j_2 j\ j\rangle$ be *real and positive*. Extensive numerical tables and computer codes are available for C-G coefficients.[6]

Without proof, we note the most useful symmetry relations for C-G coefficients:

$$\langle j_1 j_2 m_1 m_2 | j_1 j_2 j\ m\rangle = (-1)^{j - j_1 - j_2} \langle j_2 j_1 m_2 m_1 | j_2 j_1 j\ m\rangle = \langle j_2 j_1, -m_2, -m_1 | j_2 j_1 j, -m\rangle \tag{17.61}$$

Exercise 17.14. Calculate the C-G coefficients needed to couple the two angular momenta $j_1 = 3/2$ and $j_2 = 1$ to the possible j values, and express the coupled states $|j_1 j_2 j\ m\rangle$ in terms of the uncoupled states $|j_1 j_2 m_1 m_2\rangle$.

Exercise 17.15. Verify the values of two special C-G coefficients:

$$\langle j1j0 | j1j\ j\rangle = \sqrt{\frac{j}{j+1}} \quad \text{and} \quad \langle j2j0 | j2\ j\ j\rangle = \sqrt{\frac{j(2j-1)}{(j+1)(2j+3)}} \tag{17.62}$$

[6]See Thompson (1994) for references.

Determine the value of the trivial C-G coefficient $\langle j0m0|j0jm\rangle$. How does it depend on the value of m?

When we deal with spin one-half particles, the C-G coefficients for $j_2 = 1/2$ are frequently needed. From the recursion relations (17.54), the normalization condition (17.59), and the convention that $\langle j_1, 1/2, j_1, j - j_1|j_1, 1/2, j, j\rangle$ shall be real and positive, we obtain the values

$$\langle j_1, 1/2, m - 1/2, 1/2|j_1\ 1/2, j_1 \pm 1/2, m\rangle = \pm\sqrt{\frac{j_1 \pm m + 1/2}{2j_1 + 1}}$$
$$\langle j_1, 1/2, m + 1/2, -1/2|j_1, 1/2, j_1 \pm 1/2, m\rangle = \sqrt{\frac{j_1 \mp m + 1/2}{2j_1 + 1}} \tag{17.63}$$

since the allowed j values are $j = j_1 \pm 1/2$.

Exercise 17.16. Work out the results (17.63) from the recursion relations for C-G coefficients and the normalization and standard phase conditions.

If the coordinate representation is used for the eigenstates of orbital angular momentum $\mathbf{L}$ and if, as usual, the eigenstates of S_z,

$$\alpha = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \beta = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

are chosen as a basis to represent the spin, the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$ and its eigenvectors may be represented in the direct product spin $\otimes$ coordinate basis. We will denote the common eigenstates of J_z and $\mathbf{J}^2$ by $\mathcal{Y}_\ell^{jm}$. Using the C-G coefficients (17.63) with $j_1 = \ell$, we have

$$\boxed{\mathcal{Y}_\ell^{jm} = \mathcal{Y}_\ell^{\ell \pm 1/2, m} = \frac{1}{\sqrt{2\ell + 1}} \begin{pmatrix} \pm\sqrt{\ell \pm m + \frac{1}{2}} & Y_\ell^{m-1/2} \\ \sqrt{\ell \mp m + \frac{1}{2}} & Y_\ell^{m+1/2} \end{pmatrix}} \tag{17.64}$$

for the eigenstates with $j = \ell \pm 1/2$. These coupled eigenstates can be expected to play an important role in the quantum mechanics of single-electron and single-nucleon systems.

Exercise 17.17. Apply

$$\mathbf{J}^2 = \mathbf{L}^2 + \mathbf{S}^2 + 2\mathbf{L} \cdot \mathbf{S} = \mathbf{L}^2 + \mathbf{S}^2 + L_+S_- + L_-S_+ + 2L_zS_z \tag{17.65}$$

to Eq. (17.64) and verify that $\mathcal{Y}_\ell^{jm}$ is an eigenstate of $\mathbf{J}^2$.

The formulas (17.63) are also useful when the commuting spins of two particles are added, or coupled, to give the *total spin*

$$\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2 \tag{17.66}$$

If the two particles have spin one-half, the direct product spin space of the complete system is four-dimensional, and a particular basis is spanned by the uncoupled eigenvectors of $\mathbf{S}_{1_z}$ and $\mathbf{S}_{2_z}$:

$$\alpha_1 \otimes \alpha_2 = \alpha_1\alpha_2, \quad \alpha_1 \otimes \beta_2 = \alpha_1\beta_2, \quad \beta_1 \otimes \alpha_2 = \beta_1\alpha_2, \quad \beta_1 \otimes \beta_2 = \beta_1\beta_2 \tag{17.67}$$

By letting $j_2 = 1/2$ in the expressions (17.63), we obtain the appropriate C-G coefficients, allowing us to write the coupled eigenstates of S_z and $\mathbf{S}^2$ in the form:

$$\boxed{|\tfrac{1}{2}\,\tfrac{1}{2}\,0,\,0\rangle = \frac{1}{\sqrt{2}}(\alpha_1 \otimes \beta_2 - \beta_1 \otimes \alpha_2) = \frac{1}{\sqrt{2}}(\alpha_1\beta_2 - \beta_1\alpha_2)} \tag{17.68}$$

and

$$\boxed{\begin{aligned} |\tfrac{1}{2}\,\tfrac{1}{2}\,1,\,+1\rangle &= \alpha_1 \otimes \alpha_2 = \alpha_1\alpha_2 \\ |\tfrac{1}{2}\,\tfrac{1}{2}\,1,\,0\rangle &= \frac{1}{\sqrt{2}}(\alpha_1 \otimes \beta_2 + \beta_1 \otimes \alpha_2) = \frac{1}{\sqrt{2}}(\alpha_1\beta_2 + \beta_1\alpha_2) \\ |\tfrac{1}{2}\,\tfrac{1}{2}\,1,\,-1\rangle &= \beta_1 \otimes \beta_2 = \beta_1\beta_2 \end{aligned}} \tag{17.69}$$

If we write the eigenvalues of $\mathbf{S}^2$ in the usual form $\hbar^2 S(S+1)$, the total spin quantum numbers $S = 0$ and $S = 1$ characterize two irreducible representations of $R(3)$ and the eigenstates of $\mathbf{S}^2$. The state (17.68) corresponds to $S = 0$ and is called a *singlet* state. The three states (17.69) correspond to $S = 1$ and are said to be the members of a *triplet*; they are, successively, eigenstates of S_z with eigenvalues $+\hbar$, 0, and $-\hbar$.

Exercise 17.18. Starting with the uncoupled basis (17.67), work out the 4×4 matrices S_z and $\mathbf{S}^2$, and show by explicit diagonalization that the singlet and triplet states are the eigenvectors with the appropriate eigenvalues.

6. ***The Clebsch-Gordan Series.*** The direct products of the matrices of two representations of a group again constitute a representation of the same group. The latter is usually reducible even if the original two representations are irreducible. The product representation can then be reduced, or decomposed, giving us generally new irreducible representations.

For the rotation group in three-space, Eqs. (17.47) and (17.48) show that the problem of reducing the direct product representation $D^{(j_1)}(R) \otimes D^{(j_2)}(R)$ is intimately related to the problem of adding two commuting angular momenta, $\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2$. Under rotations, the state vector in the $(2j_1 + 1) \times (2j_2 + 1)$-dimensional direct product space transforms according to $D^{(j_1)} \otimes D^{(j_2)}$, and the reduction consists of determining the invariant subspaces contained in this space. Since the irreducible representations are characterized by the eigenvalues of $\mathbf{J}^2$, the integral or half-integral quantum number j, which can assume values between $|j_1 - j_2|$ and $j_1 + j_2$ in integral steps, labels the several representations into which the direct product representation decomposes. It follows from the last section that in the reduction each of these irreducible representations of the rotation group appears once and only once. Formally, this fact is expressed by the equivalence

$$D^{(j_1)} \otimes D^{(j_2)} \to \begin{pmatrix} D^{(j_1+j_2)} & 0 & 0 & 0 \\ 0 & D^{(j_1+j_2-1)} & 0 & 0 \\ 0 & 0 & . & 0 \\ 0 & 0 & 0 & D^{|j_1-j_2|} \end{pmatrix} \tag{17.70}$$

In particle physics, it is customary to identify the irreducible representations by their dimensionality. For the rotation group, which has (except for equivalence transformations produced by changing bases) precisely one irreducible representation of

each dimension, the boldface number $\mathbf{2j + 1}$ thus characterizes an irreducible representation unambiguously. The relation (17.70) can then be concisely, if symbolically, written as

$$(\mathbf{2j_1 + 1}) \otimes (\mathbf{2j_2 + 1}) = [\mathbf{2(j_1 + j_2) + 1}] \oplus [\mathbf{2(j_1 + j_2 - 1) + 1}] \oplus \ldots \oplus [\mathbf{2|j_1 - j_2| + 1}] \tag{17.71}$$

Exercise 17.19. Show by explicit counting that the matrices on both sides of (17.70) have the same dimension and that (17.71) is correct. Verify this for a few examples by letting j_1 and j_2 assume some simple values, e.g., 0, 1/2, 1, 3/2, or 2.

The Clebsch-Gordan coefficients furnish the unitary transformation from the basis $|j_1 j_2 m_1 m_2\rangle$, in which the matrices and $D^{(j_1)}$ and $D^{(j_2)}$ represent rotations, to the basis $|j_1 j_2 j m\rangle$, in which $D^{(j)}$ represents rotations. Hence, noting also that the C-G coefficients are real, we may write the equivalence (17.70) immediately and explicitly as

$$\boxed{D^{(j_1)}_{m_1' m_1}(R) D^{(j_2)}_{m_2' m_2}(R) = \sum_{j=|j_1-j_2|}^{j_1+j_2} \sum_{mm'} \langle j_1 j_2 m_1 m_2 | j_1 j_2 j m\rangle \langle j_1 j_2 m_1' m_2' | j_1 j_2 j m'\rangle D^{(j)}_{m' m}(R)} \tag{17.72}$$

This expansion is called the *Clebsch-Gordan series*. As a useful application of this equation, we call on the identity (17.44) to obtain for $j_1 = \ell_1$, $j_2 = \ell_2$, $j = \ell$ and appropriate choices for the m's:

$$Y^{m_1}_{\ell_1}(\theta, \varphi) Y^{m_2}_{\ell_2}(\theta, \varphi) = \sum_{\ell} \sqrt{\frac{(2\ell_1 + 1)(2\ell_2 + 1)}{4\pi(2\ell + 1)}} \times \langle \ell_1 \ell_2 00 | \ell_1 \ell_2 \ell 0\rangle \langle \ell_1 \ell_2 m_1 m_2 | \ell_1 \ell_2 \ell, m_1 + m_2\rangle Y^{m_1+m_2}_{\ell}(\theta, \varphi) \tag{17.73}$$

From this we find the value of the frequently used integral,

$$\boxed{\int [Y^{m_3}_{\ell_3}(\theta, \varphi)]^* Y^{m_1}_{\ell_1}(\theta, \varphi) Y^{m_2}_{\ell_2}(\theta, \varphi) d\Omega = \sqrt{\frac{(2\ell_1 + 1)(2\ell_2 + 1)}{4\pi(2\ell_3 + 1)}} \times \langle \ell_1 \ell_2 00 | \ell_1 \ell_2 \ell_3 0\rangle \langle \ell_1 \ell_2 m_1 m_2 | \ell_1 \ell_2 \ell_3 m_3\rangle} \tag{17.74}$$

Exercise 17.20. By use of the unitary condition for the D matrices and the orthogonality of the C-G coefficients, derive from (17.72) the linear homogeneous relation for the C-G coefficients:

$$\sum_{m_2'} \langle j_1 j_2 m_1' m_2' | j_1 j_2 j_3 m_3\rangle D^{(j_2)}_{m_2' m_2}(R) = \sum_{m_1 m} D^{(j_3)}_{m_3 m}(R) \langle j_1 j_2 m_1 m_2 | j_1 j_2 j_3 m\rangle D^{(j_1)}_{m_1' m_1}{}^*(R) \tag{17.75}$$

Show that for infinitesimal rotations this equation is identical with the recursion relations (17.54) for the C-G relations.

7. ***Tensor Operators and the Wigner-Eckart Theorem.*** So far this chapter has been concerned mainly with the behavior under rotation of state vectors and wave functions. This section focuses on the rotational transformation properties of the operators of quantum mechanics and generalizes the concepts introduced in Sections

11.1 and 16.3. The operators corresponding to various physical quantities will be characterized by their behavior under rotation as scalars, vectors, and tensors. From a knowledge of this behavior alone, much information will be inferred about the structure of the matrix elements. Such information is useful in many applications of quantum mechanics.

Let us suppose that a rotation R takes a state vector Ψ by a unitary transformation U_R into the state vector $\Psi' = U_R\Psi$. A *vector operator* $\mathbf{A}$ for the system is defined as an operator whose expectation value $\langle\mathbf{A}\rangle$ is a vector that rotates together with the physical system. It is convenient to consider the rotationally invariant scalar product $\langle\mathbf{A}\rangle \cdot \hat{\mathbf{e}}$ where $\hat{\mathbf{e}}$ denotes an arbitrary vector that also rotates with the system into the vector $\hat{\mathbf{e}}'$. Then we require that

$$\langle\Psi|\mathbf{A}|\Psi\rangle \cdot \hat{\mathbf{e}} = \langle\Psi'|\mathbf{A}|\Psi'\rangle \cdot \hat{\mathbf{e}}' \tag{17.76}$$

Since Ψ is an arbitrary state, we infer the condition

$$U_R\mathbf{A}U_R^\dagger \cdot \hat{\mathbf{e}} = \mathbf{A} \cdot \hat{\mathbf{e}}' \tag{17.77}$$

Under a right-handed rotation parametrized by the axis $\hat{\mathbf{n}}$ and the angle of rotation ϕ, we can, according to (16.36), express the relation between $\hat{\mathbf{e}}$ and $\hat{\mathbf{e}}'$ as

$$\hat{\mathbf{e}}' = [1 + (1 - \cos\phi)\,\hat{\mathbf{n}} \times (\hat{\mathbf{n}} \times) + \sin\phi\,\hat{\mathbf{n}} \times]\,\hat{\mathbf{e}} \tag{17.78}$$

or in matrix form, using a Cartesian basis $(\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}})$,

$$e_i' = \sum_{j=1}^{3} R_{ij}e_j \tag{17.79}$$

where

$$R = \begin{pmatrix} (1 - n_x^2)\cos\phi + n_x^2 & n_xn_y(1 - \cos\phi) - n_z\sin\phi & n_xn_z(1 - \cos\phi) + n_y\sin\phi \\ n_yn_x(1 - \cos\phi) + n_z\sin\phi & (1 - n_y^2)\cos\phi + n_y^2 & n_yn_z(1 - \cos\phi) - n_x\sin\phi \\ n_zn_x(1 - \cos\phi) - n_y\sin\phi & n_zn_y(1 - \cos\phi) + n_x\sin\phi & (1 - n_z^2)\cos\phi + n_z^2 \end{pmatrix} \tag{17.80}$$

If (17.79) is substituted in (17.77), the condition for the components of a vector operator is seen to be

$$U_RA_jU_R^\dagger = \sum_i A_iR_{ij} \tag{17.81}$$

Exercise 17.21. To first order, work out the rotation matrix R for an infinitesimal rotation characterized by the vector $\boldsymbol{\varepsilon} = \varepsilon\hat{\mathbf{n}}$, and check that this agrees with the result from formula (16.38).

Exercise 17.22. By carrying out two successive rotations, verify that (17.81) is consistent with the properties of the rotation group and its representation by the R matrices.

Again it is helpful to confine the discussion to infinitesimal rotations, since finite rotations can be generated by integration (exponentiation). If (17.78) is applied to an infinitesimal rotation by an angle $\phi = \varepsilon$, and substituted in (17.77), we find that

$$\left(1 - \frac{i}{\hbar}\,\varepsilon\hat{\mathbf{n}} \cdot \mathbf{J}\right)\mathbf{A}\left(1 + \frac{i}{\hbar}\,\varepsilon\hat{\mathbf{n}} \cdot \mathbf{J}\right) = \mathbf{A} - \varepsilon\hat{\mathbf{n}} \times \mathbf{A}$$

or

$$\boxed{[\mathbf{A}, \hat{\mathbf{n}} \cdot \mathbf{J}] = i\hbar\hat{\mathbf{n}} \times \mathbf{A}} \tag{17.82}$$

in agreement with the commutation relations (11.19) and (16.42). **A** vector operator **A** must satisfy (17.82) for arbitrary $\hat{\mathbf{n}}$. Of course, **J** itself is a vector operator and satisfies (17.82) for $\mathbf{A} = \mathbf{J}$.

Whether or not a vector **A** constitutes a vector *operator* for a system depends on the definition of the physical system and the structure of its angular momentum operator **J**. As an example, consider the case where the coordinate operators x, y, z provide a complete description of the dynamical system, such as a particle without spin. In this case, we can identify **J** with $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. The quantities **r**, **p**, **L** are all vector operators for this system, as can be verified by using the fundamental commutation relations between **r** and **p** and checking the commutation relations (17.82). On the other hand, an *external* electric field **E** acting on the system does not in general make up a *vector operator* for this system—even though **E** is a vector. Neither condition (17.81) nor condition (17.82) is satisfied by such a field, as long as **E** is external to the system and not subject to rotation together with it. But **E** would become a proper vector operator if the system were enlarged so as to include the electromagnetic field in the dynamical description. This would result in a much more complicated operator **J**, and the commutation relations (17.82) would then be satisfied by **E**. The theory will be generalized along these lines in Chapter 23.

If the system is a particle with spin, **S** becomes a vector operator provided that **J** is now taken to be $\mathbf{L} + \mathbf{S}$; that is, the spin wave function must be rotated together with the space wave function.

A vector is a tensor of rank one. The left-hand side of (17.81), $A_j' = U_R A_j U_R^\dagger$, can according to (9.128) be interpreted as the operator into which A_j is transformed by the rotation. If A_j maps a state Ψ into Φ, $A_j' = U_R A_j U_R^\dagger$ maps $U_R\Psi$ into $U_R\Phi$. The Cartesian components of $\mathbf{A}'$ are linear combinations of the components of **A**, and the expansion coefficients R_{ij} are the matrix elements of the fundamental three-dimensional representation of the rotation group in a basis spanned by the Cartesian unit vectors.

We generalize these ideas and define a *tensor operator* as a set of operators which, in analogy to (17.81), under rotation induce a representation of the rotation group. Generally, such a representation is reducible. (In particular, the usual Cartesian tensors induce reducible representations if the rank of the tensor exceeds 1.) The reduction decomposes the space of the tensors into rotationally invariant subspaces, suggesting that we define as a building block of the theory the so-called *irreducible spherical tensor operator* of nonnegative integer rank k. This is a set of $2k + 1$ operators T_k^q, with $q = -k, -k + 1, \ldots k - 1, k$, which satisfy the transformation law

$$\boxed{U_R T_k^q U_R^\dagger = \sum_{q'=-k}^{k} T_k^{q'} D_{q'q}^{(k)}(R)} \tag{17.83}$$

As a special case (for $k = 1$), we may apply (17.83) to the (irreducible) vector operator **r** and compare the formula with (17.81). If we apply the rotational transformation of spherical harmonics, as given by Eq. (17.42), to the case $k = 1$ and employ the explicit form of the spherical harmonics Y_1^m, from (11.82), we find

$$\left(-\frac{x' + iy'}{\sqrt{2}}, z', \frac{x' - iy'}{\sqrt{2}}\right) = \left(-\frac{x + iy}{\sqrt{2}}, z, \frac{x - iy}{\sqrt{2}}\right) D^{(1)}(R) \tag{17.84}$$

where the representation matrix $D^{(1)}(R)$ is the matrix (17.29) for $j = 1$. Hence, the Cartesian components of any vector operator **A** define an irreducible spherical tensor operator T_1 of rank 1 by the relations

$$T_1^{+1} = -\frac{A_x + iA_y}{\sqrt{2}}, \qquad T_1^0 = A_z, \qquad T_1^{-1} = \frac{A_x - iA_y}{\sqrt{2}} \tag{17.85}$$

Exercise 17.23. Construct the 3×3 matrix for the transformation (17.85) and show that it is unitary. How does this matrix transform the rotation matrix R into $D^{(1)}(R)$?

Armed with the definition of an irreducible tensor operator, we can now formulate and prove the *Wigner-Eckart theorem*, which answers the following question: If T_k^q ($q = k, j - 1, \ldots -k + 1, -k$) is an irreducible tensor operator, how much information about its matrix elements in the angular momentum basis can be inferred?

To find the answer, let us take the matrix element of the operator equation (17.83) between the states $\langle \alpha' j' m' |$ and $| \alpha j m \rangle$. In addition to the angular momentum quantum numbers, the labels α and α' signify the totality of all quantum numbers needed to specify the basis states of the system completely. We obtain

$$\langle \alpha' j' m' | U_R T_k^q U_R^\dagger | \alpha j m \rangle = \sum_{q'=-k}^{k} \langle \alpha' j' m' | T_k^{q'} | \alpha j m \rangle D_{q'q}^{(k)}(R)$$

Using the definition (17.28) of the matrix elements of U_R, we convert this equation into the set of simultaneous equations

$$\sum_{\mu\mu'} D_{m'\mu'}^{(j')}(R) \langle \alpha' j' \mu' | T_k^q | \alpha j \mu \rangle D_{m\mu}^{(j)*}(R) = \sum_{q'} \langle \alpha' j' m' | T_k^{q'} | \alpha j m \rangle D_{q'q}^{(k)}(R) \tag{17.86}$$

A glance at this equation and (17.75) shows that the two have exactly the same structure. If the D are known, the linear homogeneous relations (17.75) determine the C-G coefficients for given values of j_1, j_2, j_3, except for a common factor. Hence, (17.86) must similarly determine the matrix elements of the tensor operator. If we identify $j = j_1$, $k = j_2$, and $j' = j_3$, we can conclude that the two solutions of these linear homogeneous recursion relations must be proportional:

$$\boxed{\langle \alpha' j' m' | T_k^q | \alpha j m \rangle = \langle jkmq | jkj'm' \rangle \langle \alpha' j' \parallel T_k \parallel \alpha j \rangle} \tag{17.87}$$

This important formula embodies the *Wigner-Eckart theorem.*

The constant of proportionality $\langle \alpha' j' \parallel T_k \parallel \alpha j \rangle$ is called the *reduced matrix element* of the irreducible spherical tensor operator T_k. It depends on the nature of the tensor operator, on the angular momentum quantum numbers j and j', and on the "other" quantum numbers α and α', but *not* on the quantum numbers m, m', and q which specify the orientation of the system.

The Wigner-Eckart theorem provides us with a fundamental insight, because it separates the geometric and symmetry-related properties of the matrix elements from the other physical properties that are contained in the reduced matrix element. Since the C-G coefficients are readily available, the theorem also has great practical value.

From the fundamental properties (17.53) and (17.57) of the C-G coefficients, we infer the angular momentum *selection rules* for irreducible spherical tensor operators. The matrix element $\langle \alpha' j' m' | T_k^q | \alpha j m \rangle$ vanishes unless

$$\boxed{q = m' - m} \tag{17.88}$$

and the numbers, j, j', and k satisfy the triangular condition:

$$\boxed{|j - j'| \leq k \leq j + j'} \tag{17.89}$$

In particular, it follows that a scalar operator ($k = 0$) has nonvanishing matrix elements only if $m = m'$ and $j = j'$. The selection rules for a vector operator ($k = 1$) are

$$\Delta m \equiv m' - m = 0, \pm 1 \quad \text{and} \quad \Delta j \equiv j' - j = 0, \pm 1$$

with $j = j' = 0$ excluded.

Exercise 17.24. Derive the angular momentum selection rules for an irreducible second-rank (*quadrupole*) tensor operator.

Although (17.83) defines a tensor operator, to test a given set of $2k + 1$ operators T_k^q for its rotational behavior and irreducibility, it is much easier and sufficient to check this condition for infinitesimal rotations. For these, (17.83) becomes in first order,

$$\boxed{[\hat{\mathbf{n}} \cdot \mathbf{J}, T_k^q] = \sum_{q'=-k}^{k} T_k^{q'} \langle kq' | \hat{\mathbf{n}} \cdot \mathbf{J} | kq \rangle} \tag{17.90}$$

where $\hat{\mathbf{n}}$ points along the arbitrarily chosen axis of rotation. Using the familiar expressions (17.27) for the matrix elements of the angular momentum components, we derive from (17.90) the standard commutation relations

$$\boxed{\begin{aligned} [J_z, T_k^q] &= q\hbar T_k^q \\ [J_+, T_k^q] &= \sqrt{(k - q)(k + q + 1)}\hbar T_k^{q+1} \\ [J_-, T_k^q] &= \sqrt{(k + q)(k - q + 1)}\hbar T_k^{q-1} \end{aligned}} \tag{17.91}$$

Exercise 17.25. Show that by successive application of infinitesimal rotations (exponentiation) the argument leading from (17.83) to (17.90) can be reversed, suggesting (though not proving) the sufficiency of the commutation relations (17.90) or (17.91) as a test for an irreducible spherical tensor operator.

Exercise 17.26. Show that the trace of any irreducible spherical tensor operator vanishes, except those of rank 0 (scalar operators).

Exercise 17.27. Show that the tensor operators $S_k^q = (-1)^q T_k^{-q\dagger}$ and T_k^q transform in the same way under rotation. Prove that

$$\langle \alpha' j \parallel S_k \parallel \alpha j \rangle = \langle \alpha j \parallel T_k \parallel \alpha' j \rangle^*$$

and deduce the identity

$$\langle jkmq | jkjm' \rangle = (-1)^q \langle jkm', -q | jkjm \rangle \tag{17.92}$$

Exercise 17.28. If S_k^q and T_k^q are two irreducible spherical tensor operators of rank k, prove that the *contracted* operator

$$\sum_{q=-k}^{+k} (-1)^q S_k^q T_k^{-q} \tag{17.93}$$

is a tensor operator of rank zero, i.e., a scalar operator. For $k = 1$, show that this scalar operator is just the inner product $\mathbf{S} \cdot \mathbf{T}$.

Exercise 17.29. Prove that every irreducible spherical tensor operator of rank k can be expressed as a linear combination of the basic constituent tensor operators

$$T_k^q(\alpha' j', \alpha j) = \sum_{mm'} |\alpha' j' m'\rangle\langle jkmq|jkj'm'\rangle\langle \alpha jm| \tag{17.94}$$

with the appropriate reduced matrix elements serving as the coefficients in the linear combination.

Although the Wigner-Eckart theorem has been discussed in this section entirely in the context of rotations in ordinary three-dimensional space, the concepts developed here are of great generality and can be extended to other symmetries and their group representations.

8. ***Applications of the Wigner-Eckart Theorem.*** The static electric moments of a charge distribution, such as an atom or a nucleus, are examples of tensor operators. They are best defined in terms of the perturbation energy of the particles in an external electric field $\mathbf{E} = -\nabla\phi$. In first order, a particle of charge q contributes

$$E = q \int \psi^*(\mathbf{r})\phi(\mathbf{r})\psi(\mathbf{r})\, d^3r \tag{17.95}$$

to the interaction energy. If the sources of the electric field are at large distances, ϕ satisfies Laplace's equation and may be written in the form

$$\phi(\mathbf{r}) = \sum_{k=0}^{\infty} \sum_{q=-k}^{+k} A_k^q r^k Y_k^q(\theta, \varphi) \tag{17.96}$$

the origin being chosen at the center of mass of the atom or nucleus. The energy can therefore be expressed as the *electric multipole expansion*,

$$E = \int \rho(\mathbf{r})\phi(\mathbf{r})\, d^3r = \sum_{kq} A_k^q \int \rho(\mathbf{r}) r^k Y_k^q\, d^3r \tag{17.97}$$

At each order of k, the field is characterized by $2k + 1$ constants A_k^q, and there is correspondingly a spherical tensor of rank k, $r^k Y_k^q(\hat{\mathbf{r}})$, the *electric 2^k-pole moment*. Since spherical harmonics transform irreducibly under rotations in ordinary space, the components of these tensors obviously pass the test (17.91) for irreducible spherical tensor operators, provided that the rotations are generated by the operator $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. Their matrix elements may therefore be evaluated by use of the Wigner-Eckart theorem.

Exercise 17.30. Prove that the static 2^k-pole moment of a charge distribution has zero expectation value in any state with angular momentum $j < k/2$. Verify this property explicitly for the quadrupole moment by use of the formulas (17.62) and the Wigner-Eckart theorem.

Exercise 17.31. For a particle in a central-force potential, with energy basis functions separable in radial and spherical polar coordinates, use (17.74) to evaluate the reduced matrix elements of the electric multiple moment operators.

Next, as an application of the Wigner-Eckart theorem to a vector operator **A** and to the special vector operator **J**, we use the relation

$$\langle \alpha' j' m' | A^q | \alpha j m \rangle = \frac{\langle \alpha' j' \parallel \mathbf{A} \parallel \alpha j \rangle}{\langle \alpha' j' \parallel \mathbf{J} \parallel \alpha j \rangle} \langle \alpha' j' m' | J^q | \alpha j m \rangle \tag{17.98}$$

For $j' = j$ this may be further developed by noting that by virtue of the simple properties of **J** we must have for the scalar operator $\mathbf{J} \cdot \mathbf{A}$:

$$\langle \alpha' j m | \mathbf{J} \cdot \mathbf{A} | \alpha j m \rangle = C \langle \alpha' j \parallel \mathbf{A} \parallel \alpha j \rangle$$

According to the results derived in Section 17.7, the constant C must be independent of the nature of the vector operator **A** and of the quantum numbers m, α, and α'. Hence, it can be evaluated by choosing $\alpha' = \alpha$ and $\mathbf{A} = \mathbf{J}$:

$$\langle \alpha j m | \mathbf{J} \cdot \mathbf{J} | \alpha j m \rangle = j(j+1)\hbar^2 = C \langle \alpha j \parallel \mathbf{J} \parallel \alpha j \rangle$$

Substituting the reduced matrix elements of **A** and **J** from the last two equations into (17.98), we obtain the important result

$$\boxed{\langle \alpha' j m' | A^q | \alpha j m \rangle = \frac{\langle \alpha' j m | \mathbf{J} \cdot \mathbf{A} | \alpha j m \rangle}{\hbar^2 j(j+1)} \langle j m' | J^q | j m \rangle} \tag{17.99}$$

by which the j-diagonal matrix elements of a vector operator **A** are expressed in terms of the matrix element of the scalar operator $\mathbf{J} \cdot \mathbf{A}$ and other known quantities. Since the right-hand side of (17.99) can be loosely interpreted as the projection of **A** on **J**, this formula contains the theoretical justification for the *vector model* of angular momentum. It can be used as the starting point of the derivation of the matrix elements of magnetic moment operators that are important in spectroscopy.

Exercise 17.32. Prove from the Wigner-Eckart theorem and formula (17.62) that the reduced matrix element of **J** is

$$\langle \alpha j' \parallel \mathbf{J} \parallel \alpha j \rangle = \sqrt{j(j+1)}\hbar \delta_{jj'} \tag{17.100}$$

Contributions to the magnetic moment of an atom (or nucleus) arise from the orbital motion of the charged particles and the intrinsic spins in the system. Generally, the magnetic moment operator of an atom may be assumed to have the form

$$\boldsymbol{\mu} = -\frac{e}{2mc}(g_L \mathbf{L} + g_S \mathbf{S}) \tag{17.101}$$

where m is the electron mass.[7] Since both **L** and **S** are vector operators with respect to the rotations generated by the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$, $\boldsymbol{\mu}$ is also a vector operator.

According to the Wigner-Eckart theorem, all matrix elements of $\boldsymbol{\mu}$ in the angular momentum basis are proportional to each other. We therefore speak of *the magnetic moment* of the atom and, in so doing, have reference to the expectation value

$$\mu = \langle \alpha j j | \mu_z | \alpha j j \rangle = \langle j 1 j 0 | j 1 j j \rangle \langle \alpha j \parallel \boldsymbol{\mu} \parallel \alpha j \rangle = \sqrt{\frac{j}{j+1}} \langle \alpha j \parallel \boldsymbol{\mu} \parallel \alpha j \rangle \tag{17.102}$$

[7]Haken and Wolf (1993), Mizushima (1970). For an older, but very thorough, discussion, see Condon and Shortley (1935).

In the classical limit ($j \to \infty$) the reduced matrix element becomes identical with the magnetic moment μ.

The application of (17.99) to the magnetic moment vector operator (17.101) gives immediately

$$\mu = \langle \alpha jj | \mu_z | \alpha jj \rangle = \frac{-e}{2m\hbar c(j+1)} \langle \alpha jj | g_L \mathbf{L} \cdot \mathbf{J} + g_S \mathbf{S} \cdot \mathbf{J} | \alpha jj \rangle \quad (17.103)$$

But

$$2\mathbf{L} \cdot \mathbf{J} = \mathbf{J}^2 + \mathbf{L}^2 - \mathbf{S}^2 \quad \text{and} \quad 2\mathbf{S} \cdot \mathbf{J} = \mathbf{J}^2 + \mathbf{S}^2 - \mathbf{L}^2$$

Hence,

$$\mu = \frac{-e}{4m\hbar c(j+1)} \langle \alpha jj | (g_L + g_S)\mathbf{J}^2 + (g_L - g_S)(\mathbf{L}^2 - \mathbf{S}^2) | \alpha jj \rangle \quad (17.104)$$

If L-S coupling describes the state of the atom, for which the magnetic moment is to be evaluated, $|\alpha jj\rangle$ is an approximate eigenstate of the total orbital and the total spin angular momenta, which are symbolized by α. Using the usual notation J, L, and S for the three angular momentum quantum numbers that characterize the atomic state, and the values $g_L = 1$ and $g_S \approx 2$ for the electrons, we obtain for the magnetic moment of the atom,

$$\mu = -\frac{e\hbar J}{2mc}\left[1 + \frac{J(J+1) - L(L+1) + S(S+1)}{2J(J+1)}\right] = -\frac{e\hbar}{2mc} g_J J \quad (17.105)$$

The expression in the bracket (g_J) represents the general form for L-S coupling of the *Landé g-factor*.

9. *Reflection Symmetry, Parity, and Time Reversal.* The Euclidean principle of relativity may be supplemented by the further assumption that space has no intrinsic *chirality*, or handedness: Processes take place in the same way in a physical system and its mirror image, obtained from one another by *reflection* with respect to a plane. We may call this assumption the *extended Euclidean principle of relativity*. By definition, a reflection with respect to the yz plane changes x into $-x$, and leaves y and z unchanged; similarly, p_x is transformed into $-p_x$ and p_y and p_z remain unchanged.

In any test of the extended Euclidean principle of relativity, an important difference between rotational symmetry and reflection symmetry must be remembered. In rotating a system from one orientation to another, we can proceed gradually and take it through a continuous sequence of rigid displacements, all of which are equivalent. This is not so in the case of reflection, since it is impossible to transform a system into its mirror image without distorting it into some intermediate configurations that are physically very different from the original system. It is therefore not at all obvious how quantities, seemingly unrelated to coordinate displacements, such as the electric charge of a particle, should be treated in a reflection if symmetry is to be preserved. Only experience can tell us whether it is possible to create a suitable mirror image of a complex physical system in accordance with the assumption of reflection symmetry.

In developing the mathematical formalism, it is convenient to consider *inversions* through a fixed origin ($\mathbf{r} \to -\mathbf{r}$) instead of plane reflections. This is no limitation, since any reflection can be obtained by an inversion followed by a rotation. Conversely, an inversion is the same as three successive reflections with respect

to perpendicular planes. Corresponding to an inversion, there is a unitary operator U_P with the following transformation properties:

$$U_P\mathbf{r}U_P^\dagger = -\mathbf{r}, \qquad U_P\mathbf{p}U_P^\dagger = -\mathbf{p} \tag{17.106}$$

If the system is completely described by spatial coordinates, these requirements are met by the operator U_P, defined by

$$U_P|\mathbf{r}'\rangle = |-\mathbf{r}'\rangle \tag{17.107}$$

Since two successive inversions are the same as the identity operation, we may require that

$$U_P^2 = 1 \tag{17.108}$$

choosing an arbitrary phase factor to be 1. The eigenvalues of U_P are ± 1 and are said to define the *parity* of the eigenstate. If the system is specified by $\mathbf{r}$, the eigenstates are the states with even and odd wave functions, since

$$U_P \int |\mathbf{r}'\rangle\, d^3r'\, \{\langle \mathbf{r}'|\Psi\rangle \pm \langle -\mathbf{r}'|\Psi\rangle\} = \pm \int |\mathbf{r}'\rangle\, d^3r'\, \{\langle \mathbf{r}'|\Psi\rangle \pm \langle -\mathbf{r}'|\Psi\rangle\} \tag{17.109}$$

In the context of one-dimensional systems, these properties of wave functions under reflection were discussed in Section 5.2.

The Euclidean principle of relativity connects physical laws with geometry. We expect that rotations and inversions commute. If this is assumed, we have

$$[U_P, \mathbf{J}] = 0 \quad \text{or} \quad U_P^\dagger \mathbf{J} U_P = \mathbf{J} \tag{17.110}$$

confirming that $\mathbf{J}$ is an *axial* vector operator. It follows from (17.110) that the eigenstates $|jm\rangle$ have definite parity and that all substates of a given j value have the same parity, since they are obtained by successive application of J_+. These conclusions are illustrated by the discussion of the parity of spherical harmonics in Section 11.4. Their parity is given by $(-1)^\ell$.

A mirror reflection in a plane whose normal is the unit vector $\hat{\mathbf{n}}$ can be thought of as an inversion followed (or preceded) by a rotation by an angle π around the axis $\hat{\mathbf{n}}$. Thus, the unitary operator describing such a mirror reflection is the product

$$U_P e^{-(i/\hbar)\hat{\mathbf{n}}\cdot\mathbf{J}\pi} = e^{-(i/\hbar)\hat{\mathbf{n}}\cdot\mathbf{J}\pi} U_P \tag{17.111}$$

If the system is a spin 1/2 particle, the inversion operator U_P is a multiple of the identity, since $\mathbf{J} = (\hbar/2)\boldsymbol{\sigma}$, and there is no other matrix that commutes with every component of $\boldsymbol{\sigma}$. Hence, we may choose $U_P = 1$ for the inversion. The operator for a mirror reflection, (17.111), reduces to

$$e^{-i(\pi/2)\hat{\mathbf{n}}\cdot\boldsymbol{\sigma}} = -i\hat{\mathbf{n}}\cdot\boldsymbol{\sigma} \tag{17.112}$$

Exercise 17.33. Show that $\exp\left(\dfrac{i\pi H}{\hbar\omega}\right)$ is (except for a phase factor) the reflection operator for the coordinate x, if H is the Hamiltonian of a linear harmonic oscillator along the x axis with frequency ω.

Exercise 17.34. A rotation in three-dimensional space can be replaced by the product of two mirror reflections whose intersection is the axis of rotation. Working with the representation (17.112) of a reflection, construct the product of two such

reflections and show that it represents a rotation. Relate the angle of rotation to the angle between the mirror planes.

If the Hamiltonian of a system is invariant under inversion, the *parity operator* U_P is a constant of the motion, and all eigenvectors of H may be assumed to have definite parity. Parity is conserved for a particle in a central field even in the presence of a uniform magnetic field (Zeeman effect), but the presence of a uniform electric field causes states of opposite parity to be mixed (see Chapter 18).

There are parity selection rules for the matrix elements of operators which transform simply under inversion. An *even* operator is characterized by the property

$$U_P A U_P^\dagger = A \tag{17.113}$$

and has nonvanishing matrix elements between states of definite parity only if the two states have the same parity. Similarly,

$$U_P B U_P^\dagger = -B \tag{17.114}$$

characterizes an *odd* operator, which has nonvanishing matrix elements between states of definite parity only if the two states have opposite parity. For instance, if $-e\mathbf{r}$ is the electric dipole moment operator, the expectation value of this operator is zero in any state of definite parity. More generally, an atom or nucleus in a state of definite parity has no electric 2^k-pole moment corresponding to odd values of k. If, as is known from the properties of the weak interactions, conservation of parity is only an approximate symmetry, small violations of this selection rule are to be expected.

A system is said to exhibit symmetry under *time reversal* if, at least in principle, its time development may be reversed and all physical processes run backwards, with initial and final states interchanged. Symmetry between the two directions of motion in time implies that to every state Ψ there corresponds a time-reversed state $\Theta\Psi$ and that the transformation Θ preserves the values of all probabilities, thus leaving invariant the absolute value of any inner product between states.

From Section 17.1 we know that in determining Θ, only unitary and antiunitary transformations need to be considered. The physical significance of Θ as the time reversal operator requires that, while spatial relations must remain unchanged, all velocities must be reversed. Hence, we postulate the conditions

$$\Theta \mathbf{r} \Theta^{-1} = \mathbf{r}, \qquad \Theta \mathbf{p} \Theta^{-1} = -\mathbf{p}, \quad \Theta \mathbf{J} \Theta^{-1} = -\mathbf{J} \tag{17.115}$$

If the time development of the system is given by

$$\Psi(t) = e^{-(i/\hbar)Ht}\, \Psi(0)$$

time reversal symmetry demands that the time-reversed state $\Theta\Psi(0)$ evolves into

$$\Theta\Psi(-t) = e^{-(i/\hbar)Ht}\, \Theta\Psi(0)$$

From the last two equations we obtain the condition

$$\boxed{e^{-(i/\hbar)Ht}\, \Theta = \Theta e^{(i/\hbar)Ht}} \tag{17.116}$$

if the theory is to be invariant under time reversal.

If Θ were unitary, condition (17.116) would be equivalent to

$$\Theta H + H\Theta = 0$$

If such an operator Θ existed, every stationary state Ψ_E of the system with energy E would be accompanied by another stationary state $\Theta\Psi_E$ with energy $-E$. The change of sign of the energy is in conflict with our classical notions about the behavior of the energy if all velocities are reversed, and it is inconsistent with the existence of a lower bound to the energy. Hence, Θ cannot be unitary.

If Θ is assumed to be *antiunitary* so that

$$\Theta\lambda\Psi = \lambda^*\Theta\Psi \tag{17.117}$$

and

$$(\Theta\Psi_a, \Theta\Psi_b) = (\Psi_b, \Psi_a) \tag{17.118}$$

for any two states, invariance under time reversal as expressed by (17.116) requires that

$$\boxed{\Theta H - H\Theta = 0} \tag{17.119}$$

Although the operator Θ commutes with the Hamiltonian, it is not a constant of the motion because the equation of motion (14.18) holds only for linear operators.

Exercise 17.35. Deduce the commutation relation (17.119) from (17.116).

Obviously, a double reversal of time, corresponding to the application of Θ^2 to all states, has no physical consequences. If, as is the case in ordinary quantum mechanics, a one-to-one correspondence may be set up between physical states and (admissible) state vectors, with only a phase factor remaining arbitrary, Θ^2 must satisfy the condition

$$\Theta^2\Psi = C\Psi$$

with the same constant $|C| = 1$ for all Ψ. From this condition, owing to the antiunitary property of Θ the following chain of equalities flows:

$$C\Theta\Psi = \Theta^2\Theta\Psi = \Theta^3\Psi = \Theta\Theta^2\Psi = \Theta C\Psi \doteq C^*\Theta\Psi$$

Hence, C is real and either $C = 1$ or $C = -1$, depending on the nature of the system. We also see from the identity

$$(\Theta\Psi, \Psi) = (\Theta\Psi, \Theta^2\Psi) = C(\Theta\Psi, \Psi) \tag{17.120}$$

that in the case $C = -1$,

$$(\Theta\Psi, \Psi) = 0 \tag{17.121}$$

showing that for such systems time-reversed states are orthogonal.

As a corollary, if $C = -1$ and the Hamiltonian is invariant under time reversal, the energy eigenstates may be classified in degenerate time-reversed pairs. This property is known as *Kramers degeneracy*.

For a particle without spin, the operator Θ is defined by its action on the coordinate basis:

$$\Theta|\mathbf{r}'\rangle = |\mathbf{r}'\rangle \tag{17.122}$$

Since Θ is antilinear, it is represented by complex conjugation of the wave function:

$$\langle\mathbf{r}'|\Theta\Psi\rangle = \langle\mathbf{r}'|\Psi\rangle^* = \psi^*(\mathbf{r}') \tag{17.123}$$

in agreement with the conclusions reached in Section 3.5.

Exercise 17.36. Derive (17.123) from (17.122). Formulate time reversal in the momentum representation.

From the time reversal behavior of angular momentum **J**, which according to (17.115) anticommutes with Θ, it is easy to infer that the basis vectors $|\alpha jm\rangle$ transform under time reversal according to the relation

$$\boxed{\Theta|\alpha jm\rangle = e^{i\delta}(-1)^m|\alpha j, -m\rangle} \tag{17.124}$$

where the real phase constant δ may depend on j and α but not on m. By applying Θ again to Eq. (17.124) and using the antilinear property of Θ, we find that

$$\Theta^2|\alpha jm\rangle = (-1)^{2j}|\alpha jm\rangle \tag{17.125}$$

Hence, $\Theta^2 = +1$ if j is integral, and $\Theta^2 = -1$ if j is half-integral. Kramers degeneracy implies that in atoms with an odd number of electrons (or nuclei with an odd number of constituent nucleons, or baryons with an odd number of quarks) the energy levels are doubly degenerate. Since it does not destroy time reversal symmetry, this remains true even in the presence of a static electric field. A magnetic field violates time reversal symmetry and splits the degeneracy.

Exercise 17.37. Prove Eqs. (17.124) and (17.125). Show that Θ commutes with the rotation operator $U_R = e^{-(i/\hbar)\phi\hat{\mathbf{n}}\cdot\mathbf{J}}$, and use this information to derive the symmetry relation (17.39) for the representation matrices:

$$D^{(j)}_{m'm}{}^*(R) = (-1)^{m-m'}D^{(j)}_{-m',-m}(R) \tag{17.39}$$

Exercise 17.38. Show that for a particle with spin one-half in the usual σ_z basis α and β, time reversal may be represented by $\sigma_y K$, where K stands for complex conjugation.

Exercise 17.39. Show that for a particle of spin one-half, orbital angular momentum ℓ, and total angular momentum j, the eigenstates $\mathcal{Y}^{jm}_{\ell}$ defined in (17.64) transform under time reversal into $\pm(-1)^m\,\mathcal{Y}^{j,-m}_{\ell}$, the sign depending on whether $j = \ell - 1/2$ or $j = \ell + 1/2$.

Exercise 17.40. Show that the single-particle orbital angular momentum eigenfunctions with quantum numbers ℓ and m in the momentum representation are spherical harmonics. Show that the choice of phase implied by $\langle\mathbf{p}|\ell m\rangle = i^{\ell}Y^m_{\ell}(\hat{\mathbf{p}})$ leads to consistent and simple time reversal transformation properties for the angular momentum eigenfunctions in momentum space. Compare with the time reversal transformation properties of the angular momentum eigenfunctions in the coordinate representation.

Although time reversal invariance does not lead to any conservation law, selection rules may be inferred just as from rotation and reflection invariance, because the important tensor operators usually have simple transformation properties under time reversal. An irreducible spherical tensor operator T^q_k is said to be *even* or *odd* with respect to Θ if it satisfies the condition

$$\Theta T^q_k \Theta^{-1} = \pm(-1)^q T^{-q}_k \tag{17.126}$$

The + sign refers to tensors that are even under time reversal, and the − sign to tensors that are odd.

By taking the matrix element of the operator equation (17.126) between states of sharp angular momentum and using (17.124), as well as the antilinearity of Θ, we derive:

$$\langle \alpha' j' m' | T_k^q | \alpha j m \rangle = \pm e^{i(\delta - \delta')} \langle \alpha' j', -m' | T_k^{-q} | \alpha\, j, -m \rangle^*$$

If we confine our attention to matrix elements that are diagonal in α and j, it follows from (17.126), the Wigner-Eckart theorem, and (17.61) that

$$\langle \alpha j \parallel T_k \parallel \alpha j \rangle = \pm(-1)^k \langle \alpha j \parallel T_k \parallel \alpha j \rangle^* \tag{17.127}$$

If the reduced matrix element $\langle \alpha j \parallel T_k \parallel \alpha j \rangle$ is real, it is zero for tensor operators that have odd rank k and that are even under time reversal. (It is also zero for tensor operators of even rank, which are odd under time reversal.)

Exercise 17.41. Derive the results (17.126) and (17.127).

Electric multipole moments are even under time reversal. Hence, in sharp angular momentum states the expectation values of the electric dipole moment ($k = 1$) and all other odd-rank electric moments vanish as a consequence of time reversal symmetry. We have already noted that they also vanish for states of definite parity as a result of space reflection symmetry. We conclude that the observation of a static electric dipole moment in a stationary state of definite angular momentum can be expected only if *both* space reflection and time reversal symmetries are violated by the dynamical interactions. Since the discovery of parity violation in weak interactions, the search for experimental evidence for an electric dipole moment in the neutron has been motivated by the need to determine the effect of the expected small violation of time reversal symmetry, which has its origin in the interactions of particles produced in high-energy interactions. When quantum mechanics is extended to (relativistic) interactions, which cause the creation and destruction of particles, *charge* or *particle-antiparticle conjugation* (C) joins space reflection (P) and time reversal (T) as a third discrete symmetry. It is enforced by the strong and electromagnetic interactions, but violated in the realm of weak interactions. Some comments will be found in Chapter 24.

Exercise 17.42. Show that the symmetry operations Θ and U_P commute and derive the transformation properties of the fundamental dynamical variables (position, momentum, angular momentum) under the action of the combined inversion-time reversal operation of the antilinear operator ΘU_P.

10. ***Local Gauge Symmetry.*** The applications of symmetry in quantum mechanics treated in this chapter have so far been concerned with spacetime transformations of the states. We conclude this chapter with a brief discussion of *gauge symmetries*. These symmetries have come to dominate contemporary theories of particles and fields, but they can also be introduced in the framework of ordinary quantum mechanics.

Local gauge invariance made its first appearance in Section 4.6, where we saw that the dynamical equations of quantum mechanics for a charged particle in an

external electromagnetic field can be cast in gauge-invariant form if the wave function and the potential are transformed jointly according to the following scheme:

$$\psi' = U\psi \tag{17.128}$$

$$\frac{\hbar}{i}\nabla - \frac{q}{c}\mathbf{A}' = U\left(\frac{\hbar}{i}\nabla - \frac{q}{c}\mathbf{A}\right)U^\dagger, \qquad i\hbar\frac{\partial}{\partial t} - q\phi' = U\left(i\hbar\frac{\partial}{\partial t} - q\phi\right)U \tag{17.129}$$

These relations can also be written as

$$\mathbf{A}' = U\mathbf{A}U^\dagger + \frac{i\hbar c}{q}U(\nabla U^\dagger), \qquad \phi' = U\phi U^\dagger - \frac{i\hbar}{q}U\frac{\partial U^\dagger}{\partial t} \tag{17.130}$$

We found this local gauge transformation to be implemented by the unitary operator,

$$U = e^{iqf/\hbar c} \tag{17.131}$$

where $f(\mathbf{r}, t)$ is an arbitrary smooth function of the spacetime coordinates. The result of this gauge transformation is the mapping

$$\boxed{\mathbf{A} \mapsto \mathbf{A}' = \mathbf{A} + \nabla f, \qquad \phi \mapsto \phi' = \phi - \frac{1}{c}\frac{\partial f}{\partial t}, \qquad \psi \mapsto \psi' = e^{iqf/\hbar c}\psi} \tag{17.132}$$

It has become customary to say that the principle of local gauge symmetry in quantum mechanics, or invariance of the theory under the transformation (17.128), requires the introduction of a *gauge field* $(\mathbf{A}, \phi)$, which obeys the rule (17.130) for gauge transformations. If gauge invariance is accepted as a fundamental law of physics, the electromagnetic field can be regarded as a consequence of this symmetry. The local operators U represent the one-dimensional unitary group, $U(1)$. This symmetry group is commutative, or *Abelian*.

Without any specific reference to electromagnetic interactions, Abelian gauge fields also arose in Section 8.6 when we considered the adiabatic approximation for molecules. By separately treating the fast electron motion, with the nuclear coordinates serving as slowly varying parameters, we found that the Schrödinger equation (8.107) for the much heavier nuclei generally includes a gauge field.

In the domain of physics where classical correspondence arguments are unavailable for constructing the appropriate quantum mechanical laws, the imposition of local gauge symmetry serves as a powerful principle from which theories for interacting particles can be derived. The basic idea is similar to the scheme embodied in (17.128)–(17.132), generalized to particles that have multicomponent wave functions, corresponding to intrinsic degrees of freedom that are unrelated to their spacetime transformation properties. Typical examples of such intrinsic coordinates are isospin or color ''quantum numbers'' labeling the states.[8]

To illustrate the consequences of the adoption of the gauge symmetry principle, we choose the generators of the non-Abelian group $SU(2)$ as the observables associated with the intrinsic degrees of freedom. These observables are assumed to commute with all operators related to spacetime transformations, such as position, momentum, and angular momentum. For purposes of identification, we call the intrinsic degree of freedom *isospin*, and for simplicity we assume that the particle has no ordinary spin. The wave function is now a two-component column matrix in isospin

[8]For further descriptive comments, see Section 21.1.

space, analogous to the spin space of Chapter 16. We denote by τ_1, τ_2, τ_3 the three Hermitian 2×2 isospin matrices, which are the generators of $SU(2)$. They are analogous to the Pauli matrices σ_x, σ_y, σ_z defined in Section 16.4, and their commutation relations are

$$[\tau_i, \tau_j] = 2i\varepsilon_{ijk}\tau_k \tag{17.133}$$

The traces of the matrices τ_1, τ_2, τ_3 are zero. Generalizing the previous scheme, we *postulate* invariance of the wave equation for the isospinor $\psi(\mathbf{r}, t)$ under the local gauge symmetry operations induced by the group $SU(2)$. These operations are represented by

$$U = e^{(i/\hbar)\alpha} \tag{17.134}$$

where U is a unimodular unitary matrix, and

$$\alpha(\mathbf{r}, t) = \frac{1}{2}\sum_{i=1}^{3} \alpha_i(\mathbf{r}, t)\tau_i \tag{17.135}$$

is a traceless 2×2 matrix with an arbitrary smooth spacetime dependence. The three real-valued local parameters $\alpha_i(\mathbf{r}, t)$ specify the element of the group.

The relations (17.128)–(17.130) now generalize to

$$\psi'(\mathbf{r}, t) = e^{(i/\hbar)\alpha(\mathbf{r},t)}\,\psi(\mathbf{r}, t) \tag{17.136}$$

$$\frac{\hbar}{i}\nabla - \mathbf{G}' = U\left(\frac{\hbar}{i}\nabla - \mathbf{G}\right)U^\dagger, \qquad i\hbar\frac{\partial}{\partial t} - g' = U\left(i\hbar\frac{\partial}{\partial t} - g\right)U^\dagger \tag{17.137}$$

or, equivalently,

$$\mathbf{G}' = U\mathbf{G}U^\dagger + i\hbar U(\nabla U^\dagger), \qquad g' = UgU^\dagger - i\hbar U\frac{\partial U^\dagger}{\partial t} \tag{17.138}$$

Here, $\mathbf{G}$ and g are a vector and a scalar in ordinary space, respectively, but they are 2×2 matrices in isospin space. The gauge-covariant structures in Eqs. (17.137) are the building blocks from which the gauge-invariant dynamical laws are constructed.

For Lie groups like $SU(n)$, it is sufficient to implement the symmetry conditions by considering infinitesimal transformations (17.134), choosing $|\alpha_i|/\hbar \ll 1$. For these, (17.138) reduces to

$$\mathbf{G}' = \mathbf{G} + \nabla\alpha + \frac{i}{\hbar}[\alpha, \mathbf{G}], \qquad g' = g - \frac{\partial\alpha}{\partial t} + \frac{i}{\hbar}[\alpha, g] \tag{17.139}$$

These transformation equations for the gauge fields can be rewritten, if we decompose the matrices $\mathbf{G}$ and g, as

$$\mathbf{G} = \frac{1}{2}\sum_{i=1}^{3} \mathbf{G}_i\tau_i \quad \text{and} \quad g = \frac{1}{2}\sum_{i=1}^{3} g_i\tau_i \tag{17.140}$$

Instead of a single four-vector gauge field represented by the potentials $(\mathbf{A}, \phi)$, we now have *three* such fields, $(\mathbf{G}_i, g_i)$, with $i = 1, 2, 3$. Their infinitesimal gauge transformation properties are obtained by substituting (17.135) and (17.140) in (17.139), and using the commutation relations (17.133) to give

$$\boxed{\mathbf{G}'_\ell = \mathbf{G}_\ell + \nabla\alpha_\ell - \frac{1}{\hbar}\varepsilon_{\ell jk}\alpha_j\,\mathbf{G}_k, \qquad g'_\ell = g_\ell - \frac{\partial\alpha_\ell}{\partial t} - \frac{1}{\hbar}\varepsilon_{\ell jk}\alpha_j\,g_k} \tag{17.141}$$

These transformation equations show explicitly how the gauge fields must transform if they are to ensure local gauge invariance of the quantum mechanical wave equation. The last terms on the right-hand sides arise from the non-Abelian character of the symmetry group and indicate the greater complexity of the new gauge fields in comparison with the electromagnetic field.

Exercise 17.43. For an infinitesimal gauge transformation derive (17.139) from (17.138) and work out the transformation equations (17.141).

The gauge-invariant *Yang-Mills* field equations for the field tensors derived from the gauge (vector potential) fields are analogous to Maxwell's equations. However, because of the non-Abelian character of the theory, they are nonlinear, exhibit self-couplings, and pose formidable mathematical difficulties. Further discussion is found in textbooks on quantum field theory.[9] Local gauge symmetries associated with intrinsic dynamical variables that transform as the generators of various symmetry groups $SU(n)$ are the basis of our present-day understanding of strong and weak interactions in particle physics. Similar ideas have also been found useful in many-body theories and applications to condensed-matter physics. From the perspective of abstract differential geometry (topology), gauge fields are examples of *connections*, the derived field tensors are interpreted as *curvatures*, and the line integral of the gauge field over a closed loop (like Berry's phase in Section 8.6) is a *holonomy*.

Although in this chapter all symmetries were discussed in the context of nonrelativistic quantum mechanics, the dynamic gauge symmetries as well as the purely geometric symmetries can be readily adapted to the formalism of relativistic quantum mechanics. The equations of this section, in particular, can be transcribed in Lorentz-covariant form without modification. This is not wholly surprising, since the electromagnetic field and its quantum, the massless photon, are intrinsically relativistic.

Exercise 17.44. Defining for a non-Abelian gauge field theory the analogue $\mathbf{B}$ of the magnetic field as the 2×2 isospin matrix:

$$\mathbf{B} = \frac{1}{2}\sum_{i=1}^{3} \mathbf{B}_i \tau_i = \left(\nabla - \frac{i}{\hbar}\mathbf{G}\right) \times \mathbf{G} = \nabla \times \mathbf{G} - \frac{i}{\hbar}\mathbf{G} \times \mathbf{G} \qquad (17.142)$$

prove that under a finite gauge transformation,

$$\mathbf{B}' = U\mathbf{B}U^{\dagger} \qquad (17.143)$$

and that the definition (17.142) ensures that the ''magnetic'' energy,

$$\frac{1}{2}\sum_{i=1}^{3} \mathbf{B}_i \cdot \mathbf{B}_i = \text{trace}(\mathbf{B} \cdot \mathbf{B}) \qquad (17.144)$$

is gauge invariant. Similarly, construct the analogue of the electric field $\mathbf{E}$.

Exercise 17.45. If instead of $SU(2)$ the group $SU(3)$ induces the local gauge transformations of the system, how many gauge fields would there be? (Refer to Exercise 17.5.)

[9]Gross (1993). For an elementary introduction, see Halzen and Martin (1984).

Problems

1. In the notation of (17.64) the state of a spin one-half particle with sharp total angular momentum j, m, is:

$$a\mathcal{Y}^{jm}_{j-1/2} + b\mathcal{Y}^{jm}_{j+1/2}$$

Assume this state to be an eigenstate of the Hamiltonian with no degeneracy other than that demanded by rotation invariance.

(a) If H conserves parity, how are the coefficients a and b restricted?

(b) If H is invariant under time reversal, show that a/b must be imaginary.

(c) Verify explicitly that the expectation value of the electric dipole moment $-e\,\mathbf{r}$ vanishes if either parity is conserved or time reversal invariance holds (or both).

2. A particle with spin one-half (lambda hyperon) decays at rest ($\ell = 0$) into two particles with spin one-half (nucleon) and spin zero (pion). (See Figure 17.2.)

(a) Show that, in the representation in which the relative momentum of the decay products is diagonal, the final state wave functions corresponding to $m = \pm 1/2$ may, because of conservation of total angular momentum, be written in the form

$$\langle \mathbf{p} | \tfrac{1}{2}\,\tfrac{1}{2} \rangle = A_S\alpha + A_P(\cos\theta\,\alpha + e^{i\varphi}\sin\theta\,\beta)$$
$$\langle \mathbf{p} | \tfrac{1}{2},\,-\tfrac{1}{2} \rangle = A_S\beta - A_P(\cos\theta\,\beta - e^{-i\varphi}\sin\theta\,\alpha)$$

where θ is the angle between the momentum vector $\mathbf{p}$ of the spin zero particle and the z-axis of quantization. (Neglect any interactions between the decay products in the final state.)

(b) If the initial spin one-half particle is in a state with polarization $\mathbf{P}$ ($|\mathbf{P}| \leq 1$), show that the (unnormalized) angular distribution can be written as $W = 1 + \lambda\mathbf{P}\cdot\hat{\mathbf{p}}$. Evaluate λ in terms of A_s and A_p.

3. From the results of Problem 2, construct the density matrix ρ for the spin state of the final spin one-half particle (nucleon).

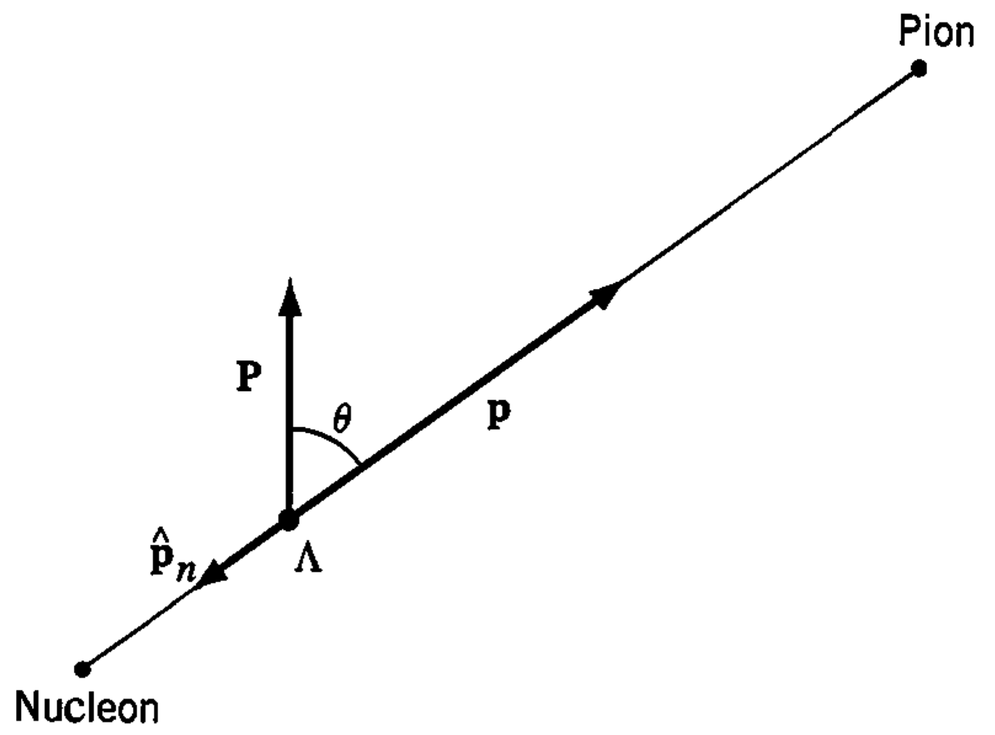


Figure 17.2. Kinematics of the decay of a spin-polarized hyperon at rest into a nucleon and a pion. The polarization vector is **P**. The magnitude of the relative momentum **p** equals that of the momenta of the two decay products in the rest frame of the hyperon. The unit vector $\hat{\mathbf{p}}_n$ points in the direction of emission of the nucleon.

(a) Prove that the polarization of the spin one-half decay products can be written in the form

$$\langle \boldsymbol{\sigma} \rangle = \text{trace}(\rho\boldsymbol{\sigma}) = \frac{1}{1 - \lambda \hat{\mathbf{p}}_n \cdot \mathbf{P}} [-(\lambda - \hat{\mathbf{p}}_n \cdot \mathbf{P})\hat{\mathbf{p}}_n + \mu \hat{\mathbf{p}}_n \times \mathbf{P} + \nu(\hat{\mathbf{p}}_n \times \mathbf{P}) \times \hat{\mathbf{p}}_n]$$

where $\hat{\mathbf{p}}_n$ is the unit vector in the direction of emission of the spin one-half particle, and $\mathbf{P}$ denotes the initial polarization of the decaying particles. Determine the real-valued quantities μ and ν in terms of A_S and A_P.

(b) Show that $\lambda^2 + \mu^2 + \nu^2 = 1$.

(c) Discuss the simplifications that occur in the expressions for the angular distribution and final state polarization if conservation of parity or invariance under time reversal is assumed for the decay-inducing interaction.

4. The Hamiltonian of the positronium atom in the $1S$ state in a magnetic field B along the z axis is to good approximation

$$H = A\mathbf{S}_1 \cdot \mathbf{S}_2 + \frac{eB}{mc}(S_{1_z} - S_{2_z})$$

if all higher energy states are neglected. The electron is labeled as particle 1 and the positron as particle 2. Using the *coupled representation* in which $\mathbf{S}^2 = (\mathbf{S}_1 + \mathbf{S}_2)^2$ and $S_z = S_{1_z} + S_{2_z}$ are diagonal, obtain the energy eigenvalues and eigenvectors and classify them according to the quantum numbers associated with constants of the motion.

Empirically, it is known that for $B = 0$ the frequency of the $1^3S \rightarrow 1^1S$ transition is 2.0338×10^5 MHz and that the mean lifetimes for annihilation are 10^{-10} s for the singlet state (two-photon decay) and 10^{-7} s for the triplet state (three-photon decay). Estimate the magnetic field strength B which will cause the lifetime of the longer lived $m = 0$ state to be reduced ("quenched") to 10^{-8} s.

5. An alternative to the usual representation for states of a particle with spin one-half, in which the simultaneous eigenstates of $\mathbf{r}$ and σ_z are used as a basis, is to employ a basis spanned by the simultaneous eigenstates of $\mathbf{r}$ and $\kappa = \boldsymbol{\sigma} \cdot \hat{\mathbf{r}}/2$. Show that the operators $\mathbf{S}^2$, $\mathbf{J}^2$, J_z, κ commute and that their eigenfunctions may be represented as

$$\langle \hat{\mathbf{r}}\kappa' \,|\, jm\kappa'' \rangle \propto D^{(j)*}_{m\kappa'}(\varphi, \theta)\delta_{\kappa'\kappa''}$$

where φ, θ denotes the rotation that turns the z axis into the direction of $\mathbf{r}$. Can this representation be generalized to particles of higher spin? Can an analogous basis be constructed in the momentum representation?[10]

6. The magnetic moment operator for a nucleon of mass m_n is $\boldsymbol{\mu} = e(g_\ell \mathbf{L} + g_S \mathbf{S})/2m_n c$, where $g_\ell = 1$ and $g_S = 5.587$ for a proton, $g_\ell = 0$ and $g_S = -3.826$ for a neutron. In a central field with an additional spin-orbit interaction, the nucleons move in shells characterized by the quantum numbers ℓ and $j = \ell \pm 1/2$. Calculate the magnetic moment of a single nucleon as a function of j for the two kinds of nucleons, distinguishing the two cases $j = \ell + 1/2$ and $j = \ell - 1/2$. Plot j times the effective gyromagnetic ratio versus j, connecting in each case the points by straight-line segments (*Schmidt lines*).

7. The angular momentum operator or generator of infinitesimal rotations may be represented in terms of Euler angles, which specify the orientation of a rectangula

[10]The *helicity operator* $\boldsymbol{\sigma} \cdot \mathbf{p}$ will be encountered in Chapter 24. See also Section 35.3 in Gottfri (1966).

coordinate system $Ox'y'z'$ anchored in a rigid body (Figure 17.1). In this representation, the components of angular momentum along the z, y'', and z' axes are

$$J_z = \frac{\hbar}{i}\frac{\partial}{\partial\alpha}, \qquad J_{y''} = \frac{\hbar}{i}\frac{\partial}{\partial\beta}, \qquad J_{z'} = \frac{\hbar}{i}\frac{\partial}{\partial\gamma}$$

(a) Prove that the components of angular momentum along the x and y axes can be expressed in terms of the Euler angles as

$$J_x = \frac{\hbar}{i}\left(-\sin\alpha\frac{\partial}{\partial\beta} - \cos\alpha\cot\beta\frac{\partial}{\partial\alpha} + \frac{\cos\alpha}{\sin\beta}\frac{\partial}{\partial\gamma}\right)$$

$$J_y = \frac{\hbar}{i}\left(\cos\alpha\frac{\partial}{\partial\beta} - \sin\alpha\cot\beta\frac{\partial}{\partial\alpha} + \frac{\sin\alpha}{\sin\beta}\frac{\partial}{\partial\gamma}\right)$$

(b) Check the commutation relations for the operators J_x, J_y, J_z.

(c) Work out the operator $\mathbf{J}^2$ in terms of the Euler angles.

(d) Using relation (17.37), show that the matrix elements $D^{(\ell)*}_{m'm}(\alpha, \beta, \gamma)$ are eigenfunctions of the differential operators J_z, $J_{z'}$ and $\mathbf{J}^2$ with eigenvalues $m'\hbar$, $m\hbar$, and $\ell(\ell + 1)\hbar^2$, respectively.

8. In the quantum mechanical treatment of a rigid body with principal moments of inertia, I_1, I_2, and I_3, it is convenient to use Euler angles as the coordinates to specify the orientation of the system. If the z' axis is the axis of a symmetric top, its Hamiltonian is

$$H = \frac{1}{2I_1}(\mathbf{J}^2 - \mathbf{J}_{z'}^2) + \frac{1}{2I_3}J_{z'}^2$$

Using the results of Problem 7, obtain the energy eigenvalues and the corresponding eigenfunctions.

9. A system that is invariant under rotation is perturbed by a quadrupole interaction

$$V = \sum_{q=-2}^{2} C_q T_2^q$$

where the C_q are constant coefficients and T_2^q are the components of an irreducible spherical tensor operator, defined by one of its components:

$$T_2^2 = (J_x + iJ_y)^2$$

(b) Deduce the conditions for the coefficients C_q if V is to be Hermitian.

(b) Consider the effect of the quadrupole perturbation on the manifold of a degenerate energy eigenstate of the unperturbed system with angular momentum quantum number j, neglecting all other unperturbed energy eigenstates. What is the effect of the perturbation on the manifold of an unperturbed $j = 1/2$ state?

(c) If $C_{\pm 2} = C_0$ and $C_{\pm 1} = 0$, calculate the perturbed energies for a $j = 1$ state, and plot the energy splittings as a function of the interaction strength C_0. Derive the corresponding unperturbed energy eigenstates.

CHAPTER 18

Bound-State Perturbation Theory

In Section 8.3, we introduced the elements of perturbation theory for the time-independent Schrödinger equation. We are now ready to study in greater generality and more detail a successive approximation scheme for the discrete bound energy eigenstates of the Hamiltonian H, known as *Rayleigh-Schrödinger perturbation theory*. Other iteration schemes have been devised in the search for more effective convergent perturbation methods, especially in many-body theory, but the Rayleigh-Schrödinger version serves our goal: to obtain a manageable algorithm for approximate energy eigenvalues and stationary states and to illustrate the theory with applications in atoms. We also return briefly to the variational method for stationary states and its relation to perturbation theory.

1. ***The Perturbation Method.*** The Rayleigh-Schrödinger perturbation method applies to the discrete energy levels of a physical system whose Hamiltonian operator H can be partitioned into two Hermitian operators:

$$H = H_0 + gV \tag{18.1}$$

Of these, H_0 will be regarded as the unperturbed part and gV as the perturbation. In (18.1), g is a real number, which will be used provisionally for bookkeeping purposes as an expansion parameter. We can let it become zero, in which case the Hamiltonian reduces to the unperturbed one, H_0, or we may let it grow to its full value, which may be chosen as $g = 1$. (In Section 8.3, g was assumed to have its full value throughout, and therefore g did not appear explicitly. In later sections of this chapter, we will also set $g = 1$.)

The eigenvalues and eigenvectors of H are functions of g. Simple perturbation theory applies when these eigenvalues and eigenvectors can be expanded in powers of g (at least in the sense of an asymptotic expansion) in the hope that for practical calculations only the first few terms of the expansion need be considered.

The eigenvalue problem we wish to solve is

$$H\Psi_n = E_n\Psi_n \tag{18.2}$$

We suppose that the related unperturbed eigenvalue problem,

$$H_0\Psi_n^{(0)} = E_n^{(0)}\Psi_n^{(0)} \tag{18.3}$$

has already been solved. Assuming, first, that no degeneracy has occurred here, let us inquire what happens to the energy eigenvalues and the corresponding eigenvectors as we allow g to grow continuously from zero to some finite value. In this process, the energy will change to

$$E_n = E_n^{(0)} + \Delta E_n \tag{18.4}$$

and the eigenvector will change to

$$\Psi_n = \Psi_n^{(0)} + \Delta\Psi_n \tag{18.5}$$

Perturbation theory assumes that the changes ΔE_n and $\Delta\Psi_n$ are small, where "small" must yet be precisely defined. We substitute expressions (18.1), (18.4), and (18.5) into (18.2) and project the equation in the direction of the unperturbed eigenvector $\Psi_n^{(0)}$. Using Eq. (18.3), we thus deduce for the *energy shift* of the nth level:

$$\boxed{\Delta E_n = \frac{(\Psi_n^{(0)}, gV\Psi_n)}{(\Psi_n^{(0)}, \Psi_n)}} \tag{18.6}$$

This is an exact relation, but it contains the unknown perturbed eigenvector Ψ_n in the inner products on the right-hand side.

The fundamental idea of perturbation theory is to assume that both the eigenvalues and eigenvectors of H can be expanded in powers of the perturbation parameter g, and to determine the coefficients in the perturbation expansions:

$$E_n = E_n^{(0)} + gE_n^{(1)} + g^2E_n^{(2)} + \ldots \tag{18.7}$$

and

$$\Psi_n = \Psi_n^{(0)} + g\Psi_n^{(1)} + g^2\Psi_n^{(2)} + \ldots \tag{18.8}$$

In writing (18.8), we have assumed that the unperturbed energy level, $E_n^{(0)}$, is *nondegenerate*. If there is degeneracy, the perturbation method that is based on the expansion (18.8) is still appropriate, provided that for some reason (such as symmetry properties) it is known beforehand to which specific state $\Psi_n \to \Psi_n^{(0)}$ in the subspace of the unperturbed eigenstates the perturbed state reduces as $g \to 0$. When such knowledge is unavailable, the discussion must be refined and *degenerate perturbation theory* used (Section 18.5).

The expansion (18.8) of the eigenvector Ψ_n in powers of g is not unique, for if Ψ_n is a solution of (18.2), then $(a_0 + a_1g + a_2g^2 + \ldots)\Psi_n$, which differs from Ψ_n by a constant factor (global gauge transformation), is an equally good solution, but its expansion in powers of g may be entirely different from (18.8). In spite of this apparent ambiguity in the eigenvectors, all physically observable quantities, such as the energies (18.7) and other expectation values of operators calculated from Ψ_n, have unique expansions in g.

If in (18.6) Ψ_n is approximated by its leading term, the known unperturbed state $\Psi_n^{(0)}$, we obtain the *first approximation*,

$$\boxed{\Delta E_n \approx gE_n^{(1)} = g\,\frac{(\Psi_n^{(0)}, V\Psi_n^{(0)})}{(\Psi_n^{(0)}, \Psi_n^{(0)})}} \tag{18.9}$$

which, if the normalization is accounted for, is the same result as formula (8.33), though in a different notation.

To go beyond the first approximation in a systematic fashion, we substitute the two expressions (18.7) and (18.8) into (18.2) and compare terms of equal power in g to obtain the successive approximation equations:

$$H_0\Psi_n^{(0)} = E_n^{(0)}\Psi_n^{(0)} \tag{18.10}$$

$$H_0\Psi_n^{(1)} + V\Psi_n^{(0)} = E_n^{(0)}\Psi_n^{(1)} + E_n^{(1)}\Psi_n^{(0)} \tag{18.11}$$

$$H_0\Psi_n^{(2)} + V\Psi_n^{(1)} = E_n^{(0)}\Psi_n^{(2)} + E_n^{(1)}\Psi_n^{(1)} + E_n^{(2)}\Psi_n^{(0)} \tag{18.12}$$

Of these equations, (18.10) is identical with (18.3) and gives nothing new. Equation (18.11) can be rewritten as

$$\boxed{(H_0 - E_n^{(0)})\Psi_n^{(1)} = (E_n^{(1)} - V)\Psi_n^{(0)}} \tag{18.13}$$

The right-hand side of this equation is known except for the value of $E_n^{(1)}$; the unknown vector $\Psi_n^{(1)}$ stands on the left. Thus (18.13) is an *inhomogeneous linear equation* for $\Psi_n^{(1)}$. If we try to solve Eq. (18.13) for $\Psi_n^{(1)}$ by dividing by the operator $H_0 - E_n^{(0)}$, we encounter an obstacle: $H_0 - E_n^{(0)}$ has zero eigenvalues and is therefore not invertible. It is worthwhile to review the properties of equations like (18.13), since all successive approximations and many other problems of quantum mechanics lead to the same type of equation—whether they are formulated as linear equations in a vector space, as matrix equations, or as differential or integral equations.

2. ***Inhomogenous Linear Equations.*** Assuming temporarily that the right-hand side is completely known, (18.13) is of the general type,

$$Au = v \tag{18.14}$$

where A is a given Hermitian operator with a complete set of eigenvectors, v is a given vector, and u is sought. For our purposes, it is quite immaterial whether A is represented by a square matrix, in which case u and v are column matrices, or by a differential operator, so that u and v are ordinary functions. We can treat this equation in a general abstract vector space, where u and v are vectors, thus encompassing many special cases simultaneously. But to most readers the theorems to be stated will be familiar from the theory of systems of simultaneous linear equations, where A is a finite dimensional square matrix.

The alternatives to be distinguished are these:

1. *Either* the homogeneous equation

$$Au' = 0 \tag{18.15}$$

possesses nontrivial solutions, that is, A has zero eigenvalues, or

2. Equation (18.15) has no nontrivial solution. In this latter case (which corresponds to $\det A \neq 0$ for finite dimensional matrices), the operator A has a unique inverse A^{-1} (see Section 9.3). For any given v, the solution of (18.14) is uniquely

$$u = A^{-1}v \tag{18.16}$$

and the problem before us has been formally solved.

In case 1, there is no linear operator inverse of A, because (18.15) cannot be inverted (since any operator acting on the null vector produces only the null vector). On the other hand, if in this case (18.14) has a *particular* solution f,

$$Af = v \tag{18.17}$$

then any vector

$$u = u' + f \tag{18.18}$$

will automatically also be a solution. By letting u' symbolize *all* solutions of the homogeneous equation (18.15), u of (18.18) represents all solutions of the inhomogeneous equation, and it is called the *general solution* of (18.14).

In this case, however, the existence of a solution depends on a further necessary (and sufficient) condition: v must have no component in the subspace spanned by the eigenvectors u' that correspond to zero eigenvalue, as seen from the equations

$$(u', v) = (u', Au) = (Au', u) = 0 \tag{18.19}$$

If P denotes the projection operator that projects every vector into the subspace spanned by the eigenvectors u', condition (18.19) can be written as

$$\boxed{Pv = 0} \tag{18.20}$$

We will make essential use of this condition.

Although guessing a particular solution f is perfectly acceptable, we can construct one by the following systematic procedure. By condition (18.20), we may write the equation to be solved, (18.17), as

$$Af = (1 - P)v \tag{18.21}$$

which is seemingly a needless complication of the simple equation (18.17). Yet, although A, which has zero eigenvalues, has no inverse, there exists a Hermitian operator K such that

$$\boxed{AK = 1 - P} \tag{18.22}$$

The operator K is thus the inverse of A in the orthocomplement of the space of the vectors u'. Infinitely many operators K satisfy this equation, because to any solution we may add an operator PB (B arbitrary) and still have a solution of (18.22). We remove this ambiguity and select a unique solution of (18.22) by imposing the subsidiary condition

$$Ku' = 0 \tag{18.23}$$

The unique operator that is defined by (18.22) and (18.23) may symbolically be written as

$$K \equiv \frac{1 - P}{A} \tag{18.24}$$

This expression is not intended to imply that A has an inverse, and it is not permissible to apply the distributive law to it and write it as $A^{-1} - A^{-1}P$, or the like. In the context of the linear differential equations of mathematical physics, an operator like K can be represented by a *Green's* function, $G(\mathbf{r}, \mathbf{r}') = \langle \mathbf{r} | K | \mathbf{r}' \rangle$ (see Section 13.3 and examples in the next section).

The chain of equalities,

$$A(Kv) = (AK)v = (1 - P)v = v$$

shows that

$$f = Kv \tag{18.25}$$

is a particular solution of (18.14). We thus conclude that if $Au' = 0$ possesses nontrivial solutions, then $Au = v$ has the general solution

$$\boxed{u = u' + Kv} \tag{18.26}$$

provided that $Pv = 0$. We will now apply these results to perturbation theory.

Exercise 18.1. Show that the condition (18.23) implies that

$$PK = KP = 0 \tag{18.27}$$

Exercise 18.2. For case 1 show that $f = Kv$ is the unique particular solution of (18.14) that is orthogonal to the subspace of the u'.

3. ***Solution of the Perturbation Equations.*** The first approximation of the Rayleigh-Schrödinger perturbation procedure has led us to the inhomogeneous equation (18.13) or:

$$(E_n^{(0)} - H_0)\Psi_n^{(1)} = -(E_n^{(1)} - V)\Psi_n^{(0)} \tag{18.28}$$

Identifying the operator $E_n^{(0)} - H_0 = A$, we see that the homogeneous equation has indeed nontrivial solutions $\Psi_n^{(0)}$, which we now assume *normalized to unity*. We have case 1 of the last section before us. We fix our attention on a definite nondegenerate unperturbed state with a particular value of n, and we denote by P_n the projection operator for the direction $\Psi_n^{(0)}$. Condition (18.20) appears now as the equation

$$P_n(E_n^{(1)} - V)\Psi_n^{(0)} = 0$$

But

$$P_n\Psi_n^{(0)} = \Psi_n^{(0)} \tag{18.29}$$

Hence,

$$E_n^{(1)}\Psi_n^{(0)} = P_nV\Psi_n^{(0)} \tag{18.30}$$

and, with the assumed normalization of $\Psi_n^{(0)}$,

$$E_n^{(1)} = (\Psi_n^{(0)}, P_nV\Psi_n^{(0)})$$

Since P_n is Hermitian, this first-order correction to the energy is thus (omitting g)

$$\boxed{E_n^{(1)} = (\Psi_n^{(0)}, V\Psi_n^{(0)}) = V_{nn}} \tag{18.31}$$

where V_{nn} is a diagonal matrix element in the representation spanned by the unperturbed energy eigenstates. It is the mean value of the perturbation in the unperturbed state, in agreement with the result of Section 8.3 and with (18.9).

By inserting (18.31) into (18.28), we can solve the latter by comparing it with the general solution (18.26). Since $\Psi_n^{(0)}$ is assumed to be the unperturbed eigenvector, it is evident that the solution (18.28) is

$$\Psi_n^{(1)} = C_n^{(1)}\Psi_n^{(0)} - K_n(E_n^{(1)} - V)\Psi_n^{(0)}$$

where $C_n^{(1)}$ is an arbitrary constant. We must remember that the operator K_n is defined by the conditions (18.22) and (18.23), or

$$\boxed{(E_n^{(0)} - H_0)K_n = 1 - P_n \quad \text{and} \quad K_n\Psi_n^{(0)} = 0} \tag{18.32}$$

Therefore, we have

$$\Psi_n^{(1)} = C_n^{(1)}\Psi_n^{(0)} + K_nV\Psi_n^{(0)} \tag{18.33}$$

It is simplest to set $C_n^{(1)}$ and all other arbitrary constants $C_n^{(k)}$, which multiply $\Psi_n^{(0)}$ in the kth approximation, equal to zero.

With $C_n^{(1)} = 0$, i.e., $P_n\Psi_n^{(1)} = 0$, we have

$$\boxed{\Psi_n^{(1)} = K_n V \Psi_n^{(0)}} \tag{18.34}$$

The operator K_n is obtained by solving equations (18.32); many techniques are available for this task. If (18.34) is used in $\Psi_n^{(0)} + g\Psi_n^{(1)}$, this approximate eigenvector is normalized to unity up to first order in g.

Exercise 18.3. If $\Psi_n^{(1)}$ is given by (18.34), prove that

$$(\Psi_n^{(0)} + g\Psi_n^{(1)}, \Psi_n^{(0)} + g\Psi_n^{(1)}) = 1 + O(g^2)$$

The same procedure can be continued systematically as the perturbation theory is carried to higher orders. In second order, we must solve (18.12), or

$$(E_n^{(0)} - H_0)\Psi_n^{(2)} = -(E_n^{(1)} - V)\Psi_n^{(1)} - E_n^{(2)}\Psi_n^{(0)} \tag{18.35}$$

which is again of the inhomogeneous type. The homogeneous equation, obtained by replacing the right-hand side by zero, again has nontrivial solutions (case 1), $\Psi_n^{(0)}$. Hence, we must require that the inhomogeneity have no component in the direction of $\Psi_n^{(0)}$ or that its inner product with $\Psi_n^{(0)}$ vanish:

$$(\Psi_n^{(0)}, (E_n^{(1)} - V)\Psi_n^{(1)} + E_n^{(2)}\Psi_n^{(0)}) = 0$$

Since $\Psi_n^{(0)}$ is orthogonal to $\Psi_n^{(1)}$, and $\Psi_n^{(0)}$ is normalized to unity, we obtain the simple relation

$$\boxed{E_n^{(2)} = (\Psi_n^{(0)}, V\Psi_n^{(1)})} \tag{18.36}$$

Armed with knowledge of $E_n^{(2)}$, we can now calculate the eigenvector correction $\Psi_n^{(2)}$ by applying the same method as before. We find

$$\Psi_n^{(2)} = C_n^{(2)}\Psi_n^{(0)} - K_n(E_n^{(1)} - V)\Psi_n^{(1)}$$

Again we choose $C_n^{(2)} = 0$, i.e., $P_n\Psi_n^{(2)} = 0$. Substituting (18.34) for $\Psi_n^{(1)}$ and (18.31) to replace $E_n^{(1)}$, we are finally able to write $\Psi_n^{(2)}$ in terms of *unperturbed* eigenstates as

$$\Psi_n^{(2)} = -E_n^{(1)}K_n^2 V\Psi_n^{(0)} + K_n V K_n V\Psi_n^{(0)} = -K_n^2 V P_n V\Psi_n^{(0)} + K_n V K_n V\Psi_n^{(0)} \tag{18.37}$$

The perturbation theory can be further developed in this way to any desired order by successive approximation. The equations rapidly become lengthy and cumbersome, but diagrammatic techniques and algebraic computer programs can lighten the work appreciably.

If we choose all eigenstate corrections $\Psi_m^{(k)}$ for $k > 0$ to be orthogonal to the imperturbed $\Psi_m^{(0)}$, or

$$(\Psi_n^{(0)}, \Psi_n^{(k)}) = \delta_{k0} \tag{18.38}$$

the normalization

$$\boxed{(\Psi_n^{(0)}, \Psi_n) = 1} \tag{18.39}$$

holds to all orders in g. Obviously, with these choices the perturbed state Ψ_n is generally *not* normalized to unity.

According to Eqs. (18.10)–(18.12) and their higher-order sequels, the orthogonality condition (18.38) implies further than

$$\boxed{E_n^{(k)} = (\Psi_n^{(0)}, V\Psi_n^{(k-1)})} \tag{18.40}$$

The challenge of perturbation theory is to determine the operator K_n from the imperturbed Hamiltonian H_0.

Exercise 18.4. Renormalize $\sqrt{Z_n}\Psi_n$ to unity and calculate to second order in g,

$$Z_n = (\Psi_n, \Psi_n)^{-1} = |(\Psi_n^{(0)}, Z_n^{1/2}\Psi_n)|^2 \tag{18.41}$$

which is the probability of finding the perturbed system in the unperturbed energy eigenstate. Show to this order that the value of Z does not depend on the choice of the arbitrary constants $C_n^{(k)}$.

Exercise 18.5. With the choice of $\Psi_n^{(1)}$ and $\Psi_n^{(2)}$ adopted in (18.34) and (18.37) ($C_n^{(1)} = C_n^{(2)} = 0$), show that

$$(\Psi_n, V\Psi_n) = E_n^{(1)} + 2gE_n^{(2)} + O(g^2) \tag{18.42}$$

Also evaluate $(\Psi_n, H_0\Psi_n)$ to second order in g. Does the sum of $(\Psi_n, H_0\Psi_n)$ and $(\Psi_n, gV\Psi_n)$ give E_n to second order in g? Does $Z_n[(\Psi_n, H_0\Psi_n) + (\Psi_n, gV\Psi_n)]$ give E_n to second order in g?

Exercise 18.6. Obtain expressions for the third-order correction to the energy and the eigenvectors in terms of the lower-order corrections.

To make this more concrete, we specialize to a particle moving in one dimension with unperturbed potential $V_0(x)$ and appropriate boundary conditions, perturbed by the potential $V(x)$ (with $g = 1$). In the coordinate representation, the inhomogeneous equation (18.28) becomes

$$\left(\frac{d^2}{dx^2} - \frac{2m}{\hbar^2}V_0(x) + k_n^2\right)\psi_n^{(1)}(x) = \frac{2m}{\hbar^2}[V(x) - E_m^{(1)}]\psi_n^{(0)}(x) \tag{18.43}$$

where we have set $E_n^{(0)} = \hbar^2 k_n^2/2m$. The operator K_n is represented by the Green's function

$$\boxed{G_n(x, x') = \frac{\hbar^2}{2m}\langle x|K_n|x'\rangle} \tag{18.44}$$

which, by (18.32), satisfies the differential equation[1]

$$\left(\frac{d^2}{dx^2} - \frac{2m}{\hbar^2}V_0(x) + k_n^2\right)G_n(x, x') = \delta(x - x') - \psi_n^{(0)}(x)\psi_n^{(0)*}(x') \tag{18.45}$$

This equation must be solved subject to the same boundary conditions as the original Schrödinger equation and the orthogonality constraint (18.23), or

$$\int_{-\infty}^{+\infty} \psi_n^{(0)*}(x)G_n(x, x')\,dx = 0 \tag{18.46}$$

[1]See Courant and Hilbert (1953), Section V.14.

The delta function on the right-hand side of (18.45) implies a discontinuity of the first derivative of $G(x, x')$ at $x = x'$, but the Green's function itself is required to be continuous. Equation (18.34) then turns into the integral expression

$$\psi^{(1)}(x) = \int_{-\infty}^{+\infty} G_n(x, x') \frac{2m}{\hbar^2} V(x')\psi^{(0)}(x')\, dx' \tag{18.47}$$

Exercise 18.7. Consider a particle in a one-dimensional box with $V_0(x) = 0$ in the domain $0 \leq x \leq L$. The unperturbed eigenfunctions are $\psi_n^{(0)}(x) = \sqrt{2/L} \sin k_n x$, and the eigenvalues are $k_n L = n\pi$ $(n = 1, 2, \ldots)$. The Green's function $G_n(x, x') = G_n(x', x)$ is[2]

$$G_n(x, x') = \frac{1}{k_n L}\left[x \cos k_n x \sin k_n x' - (L - x') \cos k_n x' \sin k_n x - \frac{1}{2k} \sin k_n x \sin k_n x'\right]$$

for $x' \geq x$. If the perturbation is a delta-function potential at the center of the box, $V(x) = -g\delta(x - L/2)$, (a) show that only eigenstates with $n =$ odd are perturbed, (b) calculate the first-order correction to the energy eigenfunction at $x = L/2$, and derive the perturbation expansion for the energy,

$$E_n = \frac{\hbar^2 n^2 \pi^2}{2mL^2} - \frac{2g}{L} - \frac{2mg^2}{\hbar^2 n^2 \pi^2} + O(g^3) \tag{18.48}$$

if n is an odd integer.

Exercise 18.8. Treat the problem of the particle in a one-dimensional box of length L, centered at $x = 0$ with perturbation $V(x) = -g\delta(x)$, exactly. Define $E_n = \hbar^2 k_n'^2/2m$ and show that for the even-parity energy eigenfunctions the eigenvalues k_n' are the roots of the transcendental equation

$$\frac{k_n' L}{2} = \arctan\left(\frac{\hbar^2 k_n'}{mg}\right) \tag{18.49}$$

(a) Discuss the solutions of this equation graphically and, by use of the power series expansion of arctan x in inverse powers of x, show that, to second order in g, the energy eigenvalues agree with (18.48).

(b) For an attractive delta-function potential ($g > 0$), show that if $gL \geq 2\hbar^2/m$, an additional bound state with $E < 0$ arises as the root of

$$\tanh \frac{\kappa L}{2} = \frac{\hbar^2 \kappa}{mg} \tag{18.50}$$

where $E = -\hbar^2 \kappa^2/2m$. (Compare with the calculation in Section 6.4.)

If K_n cannot be determined directly, Eq. (18.32) can be solved by the eigenexpansion,

$$\boxed{K_n = \sum_{\ell \neq n} \frac{1}{E_n^{(0)} - E_\ell^{(0)}} P_\ell} \tag{18.51}$$

[2]Morse and Feshbach (1953), Section 5.2, reviews the theory of ordinary inhomogeneous linear differential equations.

as is verified by substitution and noting that $H_0 P_\ell = E_\ell^{(0)} P_\ell$. From

$$P_k V \Psi_n^{(0)} = \Psi_k^{(0)} (\Psi_k^{(0)}, V\Psi_n^{(0)}) = \Psi_k^{(0)} V_{kn}$$

and Eqs. (18.34) and (18.51), it follows that

$$\boxed{\Psi_n^{(1)} = \sum_{k \neq n} \Psi_k^{(0)} \frac{V_{kn}}{E_n^{(0)} - E_k^{(0)}}} \tag{18.52}$$

This is the first-order correction to the nth eigenvector in terms of unperturbed eigenvectors. We use the concise notation V_{kn} for the matrix element of V with respect to the unperturbed energy eigenstates.

Substituting (18.34) or (18.52) into formula (18.36), we get for the second-order correction to the energy:

$$\boxed{E_n^{(2)} = (\Psi_n^{(0)}, VK_n V\Psi_n^{(0)}) = \sum_{k \neq n} \frac{V_{nk} V_{kn}}{E_n^{(0)} - E_k^{(0)}} = \sum_{k \neq n} \frac{|V_{nk}|^2}{E_n^{(0)} - E_k^{(0)}}} \tag{18.53}$$

It is evident from (18.52) and (18.53) that for the Rayleigh-Schrödinger perturbation expansion to converge rapidly it is necessary that

$$\left| \frac{V_{nk}}{E_n^{(0)} - E_k^{(0)}} \right| \ll 1 \tag{18.54}$$

The differences of unperturbed energies that make their appearance in the denominators of all of these expressions signal to us once more that degenerate unperturbed energy levels must be treated separately—unless the matrix element of V connecting two such states happens to vanish.

Exercise 18.9. Calculate the perturbed energies up to the second-order correction $E_n^{(2)}$ for the delta-function perturbation located in the center of a one-dimensional box (Exercises 18.7 and 18.8). Show that the equality of $E_n^{(2)}$ to the earlier result is assured by the identity[3]

$$\sum_{\ell=1(\ell \neq n)}^{\infty} \frac{1}{n^2 - \ell^2} = -\frac{1}{4n^2} \qquad (n \text{ and } \ell = \text{odd})$$

Exercise 18.10. Express (18.37) in terms of unperturbed energies and matrix elements of V in the representation of the unperturbed energy eigenstates.

4. ***Electrostatic Polarization and the Dipole Moment.*** As an example, we consider an electron bound in an atom and placed in a weak uniform constant external electric field **E**. The field can be derived from an electrostatic potential

$$\phi(\mathbf{r}) = -\mathbf{E} \cdot \mathbf{r}$$

where the coordinate origin is most conveniently chosen at the position of the nucleus and the perturbation potential is

$$V = -e\phi = e\mathbf{E} \cdot \mathbf{r}$$

[3]This follows from formula 0.237 (No. 4) in Gradshteyn and Ryzhik (1965).

Unless the Green's function $G_n(\mathbf{r}, \mathbf{r}')$ corresponding to the unperturbed energy $E_n^{(0)}$ can be obtained in closed form, the energy of the system to second order is usually expressed in terms of the expansion formula of the last section as

$$E_n = E_n^{(0)} + e\mathbf{E} \cdot \mathbf{r}_{nn} + e^2 \sum_{k \neq n} \frac{(\mathbf{E} \cdot \mathbf{r}_{nk})(\mathbf{E} \cdot \mathbf{r}_{kn})}{E_n^{(0)} - E_k^{(0)}} \tag{18.55}$$

where all matrix elements of the operator $\mathbf{r}$ are to be taken with respect to the unperturbed states.

The shift of energy levels in an electric field is known as the *Stark effect*. The first few terms of the perturbation expansion give accurate results for applied fields, which are small compared to the internal electric field of the atom. Since the latter is in order of magnitude given by $|\mathbf{E}| \approx 10^{11}$ V/m, in practice this condition is well satisfied for applied static fields. Successive terms in the perturbation expansion decrease rapidly and uniformly, except that some terms may vanish owing to certain symmetry properties of the system.

If the unperturbed electron is in a central field, H_0 is invariant under coordinate inversion through the center of force, and the energy eigenstates may be taken to have definite parity. We saw in Section 17.9 that the expectation value of the operator $\mathbf{r}$, which is odd under reflection, vanishes for states of definite parity. Hence, the external electric field can, in general, produce no first-order, or linear, Stark effect. An exception arises if the central field is a pure Coulomb field (hydrogenic atoms) because the excited states of such atoms exhibit degeneracy of energy eigenstates with opposite parity. Superposition of such states yields energy eigenstates that have no definite parity, and the expectation value of $\mathbf{r}$ need no longer vanish. We will resume discussion of the *linear Stark effect* of the first excited energy level of hydrogen as an example of degenerate perturbation theory in Section 18.6. The inevitable degeneracy of the magnetic substates for states of nonzero angular momentum, on the other hand, does not affect our conclusion concerning the absence of the linear Stark effect, because all these substates have the same parity (Section 17.9).

Usually, then, the first-order term in (18.55) is absent. The second-order term gives rise to the so-called *quadratic Stark effect*. If the electric field is along the z axis, the quadratic Stark shift is given by the formula

$$E_n = E_n^{(0)} + e^2\mathbf{E}^2 \sum_{k \neq n} \frac{|z_{nk}|^2}{E_n^{(0)} - E_k^{(0)}} \tag{18.56}$$

Once again, the parity selection rule ensures that the trivial degeneracy of magnetic substates does not interfere with the applicability of this formula, because two states that differ only by their magnetic quantum number have the same parity. By inspection of the work of Section 18.3, we see that such apparently indeterminate (0/0) terms are simply omitted from the sum in (18.56).

Perturbation theory may also be used to calculate the expectation value of the static electric dipole moment, $-e\mathbf{r}$, in a stationary state of the one-electron atom. In the lowest approximation,

$$\boxed{\mathbf{p}_0 = -e\mathbf{r}_{nn} = -e \int \mathbf{r} |\psi_n^{(0)}(\mathbf{r})|^2 \, d^3r} \tag{18.57}$$

This is called the *permanent electric dipole moment* of the system because it represents a vector which is determined by the unperturbed state of the system and is

entirely independent of the applied field. It vanishes for all states that possess definite parity. Conversely, its presence, as determined by measurement, signals either the occurrence of degeneracy or a breakdown of reflection symmetry.

A better approximation is obtained by using the correction

$$\begin{aligned}\rho &\approx |\psi_n^{(0)} + g\psi_n^{(1)}|^2 \\ &\approx |\psi_n^{(0)}|^2 + e\psi_n^{(0)*} \sum_{k\neq n} \psi_k^{(0)} \frac{\mathbf{E}\cdot\mathbf{r}_{kn}}{E_n^{(0)} - E_k^{(0)}} + e\psi_n^{(0)} \sum_{k\neq n} \psi_k^{(0)*} \frac{\mathbf{E}\cdot\mathbf{r}_{nk}}{E_n^{(0)} - E_k^{(0)}}\end{aligned} \tag{18.58}$$

The last two terms describe the electric polarization of the atom by the applied field. In this approximation, we obtain for the dipole moment of the one-electron atom:

$$\mathbf{p} = -e\int \rho\mathbf{r}\, d^3r = \mathbf{p}_0 - e^2 \sum_{k\neq n} \frac{\mathbf{r}_{nk}\mathbf{r}_{kn} + \mathbf{r}_{kn}\mathbf{r}_{nk}}{E_n^{(0)} - E_k^{(0)}} \cdot \mathbf{E} \tag{18.59}$$

where the last term represents the *induced dipole moment* in the state n,

$$\mathbf{p}_1 = -e^2 \sum_{k\neq n} \frac{\mathbf{r}_{nk}\mathbf{r}_{kn} + \mathbf{r}_{kn}\mathbf{r}_{nk}}{E_n^{(0)} - E_k^{(0)}} \cdot \mathbf{E} = \boldsymbol{\alpha}\cdot\mathbf{E} \tag{18.60}$$

This equation defines a *tensor* (or dyadic) of *polarizability* for the state n,

$$\boxed{\boldsymbol{\alpha} = e^2 \sum_{k\neq n} \frac{\mathbf{r}_{nk}\mathbf{r}_{kn} + \mathbf{r}_{kn}\mathbf{r}_{nk}}{E_k^{(0)} - E_n^{(0)}}} \tag{18.61}$$

The polarizability tensor is symmetric, i.e., $\alpha_{ij} = \alpha_{ji}$. In many applications, we find $\alpha_{xy} = \alpha_{yz} = \alpha_{zx} = 0$, and $\alpha_{xx} = \alpha_{yy} = \alpha_{zz}$, so that the polarizability is a scalar.

Exercise 18.11. Calculate the polarizability of an isotropic harmonic oscillator from (18.61), and verify that the result agrees with an exact calculation of the induced dipole moment.

Note that

$$\mathbf{E}\cdot\mathbf{p} = -\int \psi_n^* gV\psi_n\, d^3r = \mathbf{E}\cdot\mathbf{p}_0 + \mathbf{E}\cdot\boldsymbol{\alpha}\cdot\mathbf{E} + O(\mathbf{E}^3) \tag{18.62}$$

Comparing this with (18.42), we obtain

$$E_n = E_n^{(0)} - \mathbf{E}\cdot\mathbf{p}_0 - \tfrac{1}{2}\mathbf{E}\cdot\boldsymbol{\alpha}\cdot\mathbf{E} + \ldots \tag{18.63}$$

which upon substitution of $\mathbf{p}_0$ and $\boldsymbol{\alpha}$ can easily be seen to be identical with expression (18.55). The factor 1/2 that appears in the energy (18.63) owing to the induced dipole moment is the same as that customarily found when "stress" ($\mathbf{E}$) and "strain" ($\mathbf{p}_1$) are proportional, as is the case in the approximation leading to (18.60).

The accurate evaluation of the sums over unperturbed states, which are encountered in higher-order perturbation calculations, is usually a difficult problem. Sometimes special techniques allow us to avoid the use of the unperturbed eigenstate expansion. As an example, consider again the first-order perturbation equation (18.13),

$$(E_n^{(0)} - H_0)\Psi_n^{(1)} = (V - E_n^{(1)})\Psi_n^{(0)}$$

and assume that it is possible to find, by whatever procedure and perhaps just guessing, an operator F_n that satisfies the equation

$$(E_n^{(0)} - H_0)F_n\Psi_n^{(0)} = [F_n, H_0]\Psi_n^{(0)} = (V - E_n^{(1)})\Psi_n^{(0)} \tag{18.64}$$

By inspection, we see that the particular solution $\Psi_n^{(1)}$ of (18.13) that is orthogonal to $\Psi_n^{(0)}$, and thus of paramount interest, is

$$\Psi_n^{(1)} = (1 - P_n)F_n\Psi_n^{(0)} \tag{18.65}$$

Correspondingly, the second-order energy shift is

$$E_n^{(2)} = (\Psi_n^{(0)}, V(1 - P_n)F_n\Psi_n^{(0)}) = \langle n|VF_n|n\rangle - \langle n|V|n\rangle\langle n|F_n|n\rangle \tag{18.66}$$

if $\Psi_n^{(0)} = |n\rangle$ is normalized to unity and the bra-ket notation is employed. The usefulness of this method hinges on our ability to determine the operator F_n.

We illustrate this method by applying it to the level shift caused by a uniform electric field in the ground state $|0\rangle$ of the hydrogen atom. Since $E_0^{(1)} = \langle 0|V|0\rangle = 0$ in this case (no static dipole moment), it is easy to determine an operator F_0 from the equation

$$[F_0, H_0]|0\rangle = e|\mathbf{E}|z|0\rangle \tag{18.67}$$

as a function of spatial coordinates only (and not the momenta). In the coordinate representation, this equation for F_0 becomes a differential equation that may be solved by separation of variables, with the result:

$$F_0 = -\frac{e|\mathbf{E}|ma}{\hbar^2}\left(\frac{r}{2} + a\right)z \tag{18.68}$$

where a is the Bohr radius.

The remaining expectation value in (18.66) is easily calculated by noting that for a spherically symmetric state (S state):

$$\langle 0|f(r)z^2|0\rangle = \frac{1}{3}\langle 0|f(r)r^2|0\rangle$$

Hence,

$$E_0^{(2)} = e|\mathbf{E}|\langle 0|zF_0|0\rangle = -e^2|\mathbf{E}|^2\frac{ma}{3\hbar^2}\left[\frac{1}{2}\langle 0|r^3|0\rangle + a\langle 0|r^2|0\rangle\right]$$

But

$$\langle 0|r^n|0\rangle = \frac{1}{\pi a^3}\int d\Omega\int_0^\infty r^{n+2}\exp\left(-\frac{2r}{a}\right)dr = \frac{a^n}{2^{n+1}}(n + 2)! \tag{18.69}$$

Thus, we finally obtain

$$\boxed{E_0 \approx -\frac{e^2}{2a} - \frac{9}{4}a^3|\mathbf{E}|^2} \tag{18.70}$$

for the ground state of the hydrogen atom to second order in the applied electric field. An alternative method of deriving this result consists of using parabolic coordinates, in which the Schrödinger equation is separable even in the presence of a uniform electric field.[4] The presence of the field causes a lowering of the energy. This was expected from the application of formula (18.53) for $E_n^{(2)}$, since every term in the sum may be regarded as a repulsion of the nth level by the kth level.

[4]Bethe and Salpeter (1957), Sections 51–53.

Exercise 18.12. Verify that F_0 in (18.68) is a solution of Eq. (18.67) and, using F_0, derive a closed-form expression for the first-order correction, $\Psi_0^{(1)}(\mathbf{r})$, to the ground state wave function for a hydrogen atom in a uniform electric field.

5. *Degenerate Perturbation Theory.* We must now extend our perturbation methods by admitting the possibility of degeneracy among the unperturbed energies. Two or more unperturbed energies may be equal, or nearly equal, usually as a result of certain symmetries. Thus, if the system is a particle in a central-force field, the magnetic substates of a given angular momentum all have the same energy, owing to rotational symmetry. If this symmetry is disturbed, as by the application of a magnetic field, the degeneracy is usually removed.

The perturbation procedure developed in Sections 18.1 through 18.3 cannot be applied without modification, because the expansion (18.8) of the perturbed energy eigenvector Ψ_n was based on the assumption that we know into which unperturbed eigenvector $\Psi_n^{(0)}$ collapses as g approaches zero continuously. While symmetry properties can frequently be used to avoid this ambiguity, in general we have no prior knowledge that would allow us to predict what particular linear combination of the given degenerate eigenstates the eigenvector Ψ_n will converge to as $g \to 0$. Here we set ourselves the task of finding the "correct" unperturbed states on which a perturbation expansion can be built.

The breakdown of the simple Rayleigh-Schrödinger theory in the case of degenerate unperturbed states formally appears as the vanishing of some of the energy denominators in formulas such as (18.51), (18.52), and (18.53). When this happens, the operator K and the corresponding Green's function are undefined. The perturbation expansion becomes meaningless, unless, as a result of some symmetry, the matrix element of V in the numerator vanishes as well. For practical applications, it is important to realize that these difficulties arise not only if the unperturbed eigenstates are strictly degenerate, but also if they are merely so close in energy or linked by such large matrix elements of the perturbation V that $|V_{kn}/(E_n^{(0)} - E_k^{(0)})|$ is large, violating condition (18.54) and causing unusually large mixing of unperturbed states in (18.52).

In order to keep the notation uncluttered, we suppose that partitioning the Hamiltonian H into an unperturbed operator H_0 and a perturbation V (with $g = 1$) has produced just *two* unperturbed energies, $E_1^{(0)}$ and $E_2^{(0)}$, which are close to each other (by comparison with the magnitude of V_{12}). For the special case of strict degeneracy, $E_1^{(0)} = E_2^{(0)}$. We take the corresponding eigenvectors $\Psi_1^{(0)} \equiv |1\rangle$ and $\Psi_2^{(0)} \equiv |2\rangle$ to be orthonormal.

The problem of degenerate or nearly degenerate perturbation theory is to replace these two basis states by a new set of orthonormal states (the "correct" basis states) which may serve as zero-order approximation to the actual perturbed energy eigenvectors. In essence, we attempt a *partial* diagonalization of the matrix representing the total Hamiltonian in the unperturbed eigenvector representation. This is the familiar problem of matrix mechanics, except that it arises here in the context of matrices that, instead of having infinitely many rows and columns, have the dimensionality d of the degree of degeneracy of the nth level—two in our example.

To derive the new "correct" basis vectors appropriate for perturbation theory, we construct the projection operator

$$P_{12} = P_1 + P_2 = |1\rangle\langle 1| + |2\rangle\langle 2| \tag{18.71}$$

which projects any state vector into the two-dimensional subspace of the degenerate or nearly degenerate unperturbed eigenvectors. In this subspace, we define the truncated Hamiltonian

$$P_{12}HP_{12} = E_1^{(0)} |1\rangle\langle 1| + E_2^{(0)} |2\rangle\langle 2| + P_{12}VP_{12} \tag{18.72}$$

and then proceed to calculate its eigenvalues $\overline{E}_1^{(0)}$ and $\overline{E}_2^{(0)}$ exactly. The corresponding "correct" orthonormal eigenvectors, $|\overline{1}\rangle$ and $|\overline{2}\rangle$, are determined as linear combinations of the two original basis vectors $|1\rangle$ and $|2\rangle$, so that

$$P_{12}HP_{12} = \overline{E}_1^{(0)} |\overline{1}\rangle\langle\overline{1}| + \overline{E}_2^{(0)} |\overline{2}\rangle\langle\overline{2}| \tag{18.73}$$

We now define a new unperturbed Hamiltonian,

$$\begin{aligned} \overline{H}_0 &= H_0 - E_1^{(0)} |1\rangle\langle 1| - E_2^{(0)} |2\rangle\langle 2| + \overline{E}_1^{(0)} |\overline{1}\rangle\langle\overline{1}| + \overline{E}_2^{(0)} |\overline{2}\rangle\langle\overline{2}| \\ &= H_0 - P_{12}H_0P_{12} + P_{12}HP_{12} = H_0 + P_{12}VP_{12} \end{aligned} \tag{18.74}$$

and a new perturbation operator

$$\overline{V} = H - \overline{H}_0 = V + H_0 - \overline{H}_0 = V - P_{12}VP_{12} \tag{18.75}$$

Since both H and $\overline{H}_0$ have zero off-diagonal matrix elements between the two new basis states $|\overline{1}\rangle$ and $|\overline{2}\rangle$, and since the diagonal elements $\langle\bar{i}|H|\bar{i}\rangle = \langle\bar{i}|\overline{H}_0|\bar{i}\rangle$, the matrix of the new perturbation $\overline{V}$ in this two-dimensional subspace is zero. Thus, the offending terms with vanishing denominators have been eliminated, and a new set, $\overline{E}_1^{(0)}$ and $\overline{E}_2^{(0)}$, of "unperturbed" energy eigenvalues has been found. These differ from the original eigenvalues of H_0 by the approximate level shifts, for $i = 1, 2$,

$$\Delta E_i \approx \overline{E}_i^{(0)} - E_i^{(0)} \tag{18.76}$$

The techniques of perturbation theory developed in Section 18.3 can now be applied to the Hamiltonian partitioned according to

$$H = \overline{H}_0 + \overline{V} \tag{18.77}$$

The explicit solution of the eigenvalue problem of the auxiliary Hamiltonian $P_{12}HP_{12}$ defined in (18.72), starts with construction of the matrix of $P_{12}HP_{12}$ in the representation spanned by the original unperturbed energy eigenvectors $|1\rangle$ and $|2\rangle$:

$$P_{12}HP_{12} \rightarrow \begin{pmatrix} E_1^{(0)} + \langle 1|V|1\rangle & \langle 1|V|2\rangle \\ \langle 2|V|1\rangle & E_2^{(0)} + \langle 2|V|2\rangle \end{pmatrix} \tag{18.78}$$

The eigenvectors $|\overline{1}\rangle$ and $|\overline{2}\rangle$ and the eigenvalues $\overline{E}_1^{(0)}$ and $\overline{E}_2^{(0)}$ are determined by the set of simultaneous homogeneous linear equations:

$$\begin{pmatrix} E_1^{(0)} + \langle 1|V|1\rangle & \langle 1V|2\rangle \\ \langle 2|V|1\rangle & E_2^{(0)} + \langle 2|V|2\rangle \end{pmatrix} \begin{pmatrix} \langle 1|\bar{i}\rangle \\ \langle 2|\bar{i}\rangle \end{pmatrix} = \overline{E}_i^{(0)} \begin{pmatrix} \langle 1|\bar{i}\rangle \\ \langle 2|\bar{i}\rangle \end{pmatrix} \tag{18.79}$$

These equations yield the correct linear combination of the unperturbed zero-order energy eigenvectors and the perturbation corrections to the degenerate (or nearly degenerate) zero-order energies. The secular equation that must be satisfied if (18.79) is to have nontrivial solutions is familiar from Section 8.3, where essentially the same approximate energy eigenvalues and eigenvectors for degenerate or nearly degenerate unperturbed energy levels were derived by applying the variational Rayleigh-Ritz method. Equation (18.79) is the same as Eq. (8.36), if we identify the root E of the secular equation there with $\overline{E}_i^{(0)}$ here. One of the two roots of this secular equations is $\overline{E}_1^{(0)}$, and the other is $\overline{E}_2^{(0)}$. Explicitly, they are given by

formula (8.37), which shows that the corrected initial energies depend generally in a complicated way on the unperturbed energies and on the strength of the perturbation V.

Summarizing and generalizing: If $E_i^{(0)}$ $(i = 1, 2, \ldots, d)$ is a set of d degenerate or nearly degenerate unperturbed levels, we construct the $d \times d$ matrix of the Hamiltonian H with respect to the d-dimensional manifold of orthonormal unperturbed energy eigenstates, $|E_i^{(0)}\rangle$. The d roots of the corresponding secular equation are the perturbed energies up to the first order. They are often much more accurate than this statement implies, because the interactions between the sets of d states are taken into account to *all* orders.

Exercise 18.13. Expanding the perturbed energies of two exactly degenerate unperturbed energy eigenstates $(i = 1, 2)$ in powers of V as

$$E_i = \overline{E}_i^{(0)} + \ldots = E_i^{(0)} + E_i^{(1)} + E_i^{(2)} + \ldots$$

derive the first order corrections $E_i^{(1)}$ and show that the degeneracy is removed at this level of approximation unless $\langle 1|V|1\rangle = \langle 2|V|2\rangle$ and $\langle 1|V|2\rangle = \langle 2|V|1\rangle = 0$.

If the Hamiltonian $H(\lambda)$ of a system depends on a variable real parameter λ, such as an external field, the energy eigenvalues are functions of λ. The theory of this section can be used to discuss the conditions under which *level crossing* as a function of λ may occur. Let us assume that for some value of the parameter, $\lambda = \lambda_0$, two energy eigenvalues, $E_1(\lambda_0)$ and $E_2(\lambda_0)$, are close to each other (and both are far from all other energy eigenvalues). Can we, by changing λ, cause the energy levels $E_i(\lambda)$ to cross? Is there a nearby value of $\lambda = \lambda_1$ at which strict degeneracy occurs? Near the unperturbed $\lambda = \lambda_0$, we expand the Hamiltonian to first order as

$$H(\lambda) = H(\lambda_0) + (\lambda - \lambda_0)H'(\lambda_0) \tag{18.80}$$

We identify an unperturbed Hamiltonian, $H(\lambda_0) = H_0$, and a perturbation $(\lambda - \lambda_0)H'(\lambda_0) = V(\lambda)$, such that $V(\lambda) \to 0$ as $\lambda \to \lambda_0$. For the two perturbed energies to be equal at $\lambda = \lambda_1$, the coincidence of the two roots requires that the discriminant in (8.37), adapted to the present notation, must vanish:

$$\begin{aligned}[E_2(\lambda_0) + \langle 2, \lambda_0|V|2, \lambda_0\rangle - E_1(\lambda_0) - \langle 1, \lambda_0|V|1, \lambda_0\rangle]^2 \\ + 4|\langle 1, \lambda_0|V|2, \lambda_0\rangle|^2 = 0\end{aligned} \tag{18.81}$$

Hence, it is necessary that

$$\langle 1, \lambda_0|V(\lambda_1)|2, \lambda_0\rangle = 0 \quad \text{or} \quad \langle 1, \lambda_0|H'(\lambda_0)|2, \lambda_0\rangle = 0 \tag{18.82}$$

Generally, condition (18.82), which requires the vanishing of both the real and imaginary parts of the off-diagonal matrix element, cannot be satisfied by any real value λ_0. However, symmetry properties often cause (18.82) to hold identically, independent of the value of λ_0. Then the two energy levels may be made to intersect, but otherwise a *no-crossing rule* applies. This behavior of energies as a function of a parameter will be illustrated in Section 18.6.

Exercise 18.14. If $H(\lambda)$ depends on a parameter λ and if $\Psi_n(\lambda)$ is its eigenvector that is normalized to unity, prove the *Hellmann-Feynman theorem*,

$$\frac{dE_n(\lambda)}{d\lambda} = \left(\Psi_n(\lambda), \frac{\partial H(\lambda)}{\partial \lambda}\Psi_n(\lambda)\right) \tag{18.83}$$

Show that if condition (18.82) is satisfied, the energy levels cross approximately at

$$\lambda_1 = \lambda_0 - \left[\frac{E_1(\lambda) - E_2(\lambda)}{E_1'(\lambda) - E_1'(\lambda)}\right]_{\lambda=\lambda_0}$$

A fuller explanation is in order to clarify the effects of symmetries in perturbation theory. In general, the perturbation calculation for a d-dimensional manifold of degenerate or nearly degenerate unperturbed eigenstates requires us to compute the eigenvalues of a $d \times d$ matrix. Although it is possible to solve secular equations of high dimensionality, we can save labor and gain some physical insight if we take full advantage of the symmetry properties of the system. If certain symmetries of the unperturbed system survive in the presence of the perturbation, there exist commuting Hermitian operators A, such as generators of the symmetry operations, which commute both with H_0 and H, hence also with their difference, the perturbation V:

$$[H_0, A] = [H, A] = [V, A] = 0 \tag{18.84}$$

But V need not commute with H_0. For example, if an electron moves in a central field and is perturbed by a uniform external magnetic or electric field that is parallel to the z direction, the rotational symmetry about the z axis is preserved. As a consequence, L_z commutes with H_0, H, and V. (L_x or L_y also commutes with H_0, but not with V or H.) If an operator A satisfying (18.84) exists, the eigenvectors of H_0 and H may be arranged to be also eigenvectors of A. In constructing the correct linear combinations of unperturbed energy eigenvectors, we need to include in the calculation only those eigenvectors that belong to the same eigenvalue, A' of A. Formally, the simplification comes about because if the constant of the motion A and the perturbation V commute,

$$\langle A' | V | A'' \rangle = 0 \tag{18.85}$$

unless $A' = A''$. Hence, the perturbation matrix has been partially diagonalized and the computational burden reduced, often significantly.

The *selection rule* embodied in Eq. (18.85) can be expressed in the language of group theory by saying that a perturbation obeying a certain symmetry does not mix states, which under the operations of the relevant symmetry group transform according to (or "belong" to) different irreducible representations, labeled by the eigenvalues of A.

We conclude this section with some general remarks about the perturbation expansion. The Rayleigh-Schrödinger scheme is simple, but its practical applicability is in many cases limited by poor convergence, except when the perturbation is quite weak. It has therefore been found to be most useful when the perturbation V is produced by an external field whose magnitude can be controlled experimentally. For instance, by adjusting an applied electric or magnetic field, we can conduct an experiment at a perturbation level sufficiently low to warrant applicability of the power series expansion.

Frequently, however, perturbation methods are called for in the treatment of systems with interactions V over which we have no control. For example, the Coulomb interaction between electrons in an atom or the force between nucleons in a nucleus may or may not be weak, but we certainly cannot influence it. Everywhere in many-body physics we encounter such inherent perturbations affecting the behavior of the unperturbed constituents.

Once the restricted d-dimensional eigenvalue problem for a set of degenerate or nearly degenerate unperturbed eigenstates has been solved exactly, we can pro-

ceed with the Rayleigh-Schrödinger formalism to the desired order. An alternative perturbation method is based on treating the right-hand side of

$$(E_n - H_0)\Psi_n = V\Psi_n \tag{18.86}$$

as a known quantity. If the perturbation is assumed to produce a finite level shift, $\Delta E_n = E_n - E_n^{(0)}$, the homogeneous equation, obtained by setting the right-hand side of (18.86) equal to zero, has no nontrivial solution, and we may write (case 2 in Section 18.2)

$$\Psi_n = \frac{1}{E_n - H_0} V\Psi_n \tag{18.87}$$

This equation is projected onto the orthocomplement of the d-dimensional subspace of unperturbed eigenstates corresponding to eigenvalue $E_n^{(0)}$ of H_0 by multiplying both sides of the equation by $1 - P_n$. The result is

$$\boxed{\Psi_n = P_n\Psi_n + (1 - P_n)\frac{1}{E_n - H_0} V\Psi_n} \tag{18.88}$$

effecting a separation of the unperturbed eigenstate from a perturbed component that is orthogonal to the unperturbed state similar to the procedure used in the Rayleigh-Schrödinger form of perturbation theory. In the coordinate representation, (18.88) again takes the form of an integral equation. It is important to observe, however, that on the right-hand side of (18.88) there is not only the unknown state Ψ_n, but also the unknown energy E_n (rather than the known unperturbed energy $E_n^{(0)}$). Equation (18.88) is supplemented by the condition

$$\boxed{\Delta E_n = E_n - E_n^{(0)} = \frac{(\Psi_n, P_n V\Psi_n)}{(\Psi_n, P_n\Psi_n)}} \tag{18.89}$$

which follows from (18.86) and is a generalization of (18.6). Together, (18.88) and (18.89) can be used to launch a perturbation expansion by successive iteration. This approximation scheme is known as *Brillouin-Wigner perturbation theory*.

For the elementary applications of lowest-order perturbation theory to atoms in the next section, there is no distinction between the Rayleigh-Schrödinger and Brillouin-Wigner schemes.

Exercise 18.15. By repeated iteration of Eq. (18.88), derive a perturbation series for Ψ_n for "small" V and show that this is formally much simpler than the corresponding expansion in the Rayleigh-Schrödinger theory.

6. ***Applications to Atoms.*** We are now in a position to solve a great many realistic physical problems, and this section contains a sampling from atomic spectroscopy.

The theory of the *linear Stark effect* in the hydrogen atom can be used to illustrate degenerate perturbation theory. The nonrelativistic Hamiltonian is

$$H = \frac{\mathbf{p}^2}{2m} - \frac{e^2}{r} + e|\mathbf{E}|z \tag{18.90}$$

where the last term is regarded as a perturbation. Since the spin of the electron is irrelevant at this level of approximation and therefore ignored, the ground state of hydrogen ($1S$) is nondegenerate and has even parity. Hence, no linear Stark effect occurs, and there is no permanent electric dipole moment. The situation is different for the excited states, of which we will treat only the lowest ones. The $2S$ state and

the three $2P$ states are degenerate, the former being of even parity, whereas the latter three are odd. The perturbation itself is odd and thus can mix S and P states. At first sight, degenerate perturbation theory requires that we allow the correct unperturbed energy eigenfunctions to be linear combinations of all four unperturbed eigenstates. However, H shares cylindrical (or "axial") symmetry with H_0, and L_z commutes with H. Hence, the perturbed eigenstates can still be required to be eigenstates of L_z and characterized by the eigenvalues $m\hbar$, where $m = -1$, 0, or $+1$. This quantum number classifies the one-dimensional irreducible representations of the rotation group in two dimensions, $R(2)$, appropriate to the cylindrical symmetry of the problem. The selection rule for the matrix elements of the perturbation $V = e|\mathbf{E}|z$ and of H is $\Delta m = 0$.

If we denote the several eigenstates of the unperturbed energy level n by $|n\ell m\rangle$, we see that for $n = 2$ only the states $|2, 0, 0\rangle$ and $|2, 1, 0\rangle$ are mixed by the perturbation. The states $|2, 1, 1\rangle$ and $|2, 1, -1\rangle$ are single and remain so. They exhibit no linear Start effect because they have definite parity. What remains to solve is the simple secular equation,

$$\begin{vmatrix} e|\mathbf{E}|\langle 2,0,0|z|2,0,0\rangle - \Delta E_{\pm} & e|\mathbf{E}|\langle 2,0,0|z|2,1,0\rangle \\ e|\mathbf{E}|\langle 2,1,0|z|2,0,0\rangle & e|\mathbf{E}|\langle 2,1,0|z|2,1,0\rangle - \Delta E_{\pm} \end{vmatrix} = 0 \tag{18.91}$$

where ΔE is the energy shift between the perturbed and the unperturbed degenerate states. Because of conservation of parity, the diagonal elements $\langle 2, 0, 0|z|2, 0, 0\rangle$ and $\langle 2, 1, 0|z|2, 1, 0\rangle$ vanish. Hence, the first-order change in energy is

$$\Delta E_{\pm} = \pm e|\mathbf{E}\|\langle 2,0,0|z|2,1,0\rangle| \tag{18.92}$$

and only one matrix element has to be evaluated. For this purpose, it is necessary to use the unperturbed energy eigenfunctions explicitly. They are (see Chapter 12):

$$2S(m=0)\colon \psi_{2S}^{(0)} = \frac{1}{\sqrt{4\pi}}\left(\frac{1}{2a}\right)^{3/2}\left(2 - \frac{r}{a}\right)\exp\left(-\frac{r}{2a}\right)$$

$$2P(m=0)\colon \psi_{2P}^{(0)} = \frac{1}{\sqrt{4\pi}}\left(\frac{1}{2a}\right)^{3/2}\frac{r}{a}\exp\left(-\frac{r}{2a}\right)\cos\theta$$

We calculate

$$\langle 2,0,0|z|2,1,0\rangle = \langle 2,0,0|r\cos\theta|2,1,0\rangle$$
$$= \frac{1}{4\pi}\left(\frac{1}{2a}\right)^3\frac{1}{a}\int_0^{\infty}\int_0^{\pi}\int_0^{2\pi} r^4\left(2 - \frac{r}{a}\right)\exp\left(-\frac{r}{a}\right)\cos^2\theta\sin\theta\, dr\, d\theta\, d\varphi = -3a$$

Hence, the linear Stark effect splits the degenerate $m = 0$ level into two components, the shifts being

$$\Delta E_{\pm} = \pm 3ae\,|\mathbf{E}| \tag{18.93}$$

The corresponding eigenfunctions are easily seen to be

$$\psi_{\pm} = \frac{1}{\sqrt{2}}(\psi_{2S}^{(0)} \mp \psi_{2P}^{(0)}) \tag{18.94}$$

mixing the two different parity components in equal proportions.

The degeneracy of the $2S$ and $2P$ states in the hydrogen atom is broken by any perturbation of the pure Coulomb field. Thus, the $2s$ and $2p$ levels are no longer degenerate in either the lithium atom or the heavy atoms, where they constitute components of the inner L-shell. In both cases, there are central screening fields that

modify the pure Coulomb field. On the other hand, the degeneracy of the $2\ell + 1$ unperturbed magnetic substates survives the addition of such perturbations. But in an external field along the z direction, the energy eigenstates are eigenstates of L_z because of cylindrical symmetry. This fact allows us to avoid the use of degenerate perturbation theory, once levels of different ℓ are split appreciably. The strict $2S$-$2P$ degeneracy in the hydrogen atom also disappears when the electron is treated as a relativistic particle with spin and its interaction with the electromagnetic field is taken fully into account (Lamb shift).

Exercise 18.16. Calculate the linear Stark effect for the $n = 3$ levels of hydrogen.

Next, we will deal more accurately than before with the motion of an electron in a central field with the inclusion of spin effects. We saw in Chapter 16 that a *spin-orbit interaction* proportional to the operator $\mathbf{L} \cdot \mathbf{S}$ must be included in the Hamiltonian. Its strength is given by a coefficient $2W(r)/\hbar$ about which more will be said later. The Schrödinger equation for the spinning electron in an atom is thus of the two-component spinor form

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(r) + W(r)\mathbf{L} \cdot \boldsymbol{\sigma}\right]\begin{pmatrix}\psi_+\\ \psi_-\end{pmatrix} = E\begin{pmatrix}\psi_+\\ \psi_-\end{pmatrix} \tag{18.95}$$

Actually, Eq. (18.95) is not adequate to yield precise values for the fine structure of the hydrogen energy levels because these are affected by relativistic corrections to the kinetic operator, $-(\hbar^2/2m)\nabla^2$, as much as by the specific spin-orbit interaction. Indeed, there is no advantage in using a fundamentally nonrelativistic Schrödinger equation like (18.95) for the hydrogen atom. This simple system requires and can be accorded a much more accurate treatment by the use of the fully relativistic Dirac equation of the electron. Such a treatment will be given in Section 24.9.

Equation (18.95), however, can still be regarded as a useful guide to understanding the electron spin's general role in effective one-electron systems. For example, an investigation of (18.95) can give us insight into the qualitative features of the alkali spectra. An alkali atom may indeed, for many purposes, be regarded as a one-electron atom, with all electrons in closed shells contributing merely to an effective screening of the electrostatic potential in which the single valence electron moves.

The unperturbed Hamiltonian,

$$H_0 = -\frac{\hbar^2}{2m}\nabla^2 + V(r) \tag{18.96}$$

represents a familiar central-force problem. We know from Chapter 12 that its eigenfunctions may be assumed also to be eigenfunctions of L_z and $\mathbf{L}^2$. The operator H_0, which contains no reference to the spin variable, commutes with S_z and $\mathbf{S}^2$, and it is thus possible to express the unperturbed eigenfunctions of H_0 as

$$R_{n\ell}(r)Y_\ell^m(\theta, \varphi)\alpha \quad \text{and} \quad R_{n\ell}(r)Y_\ell^m(\theta, \varphi)\beta$$

Here n denotes a radial quantum number that characterizes the unperturbed energy eigenvalues, which are in general $2(2\ell + 1)$-fold degenerate. Straightforward application of the methods of Section 18.5 would require us first to diagonalize the submatrix of the perturbation matrix with respect to these degenerate eigenstates. However, a great simplification results if it is recognized that, although neither $\mathbf{L}$ nor $\mathbf{S}$ commutes with the complete Hamiltonian $H = H_0 + W(r)\mathbf{L} \cdot \boldsymbol{\sigma}$, the *total*

angular momentum vector operator, $\mathbf{J} = \mathbf{L} + \mathbf{S}$, does commute with H, as explained in Section 16.4. The simultaneous eigenstates of the mutually commuting coupled operators $\mathbf{L}^2$, $\mathbf{S}^2$, $\mathbf{J}^2$, and J_z (instead of the uncoupled operators $\mathbf{L}^2$, L_z, $\mathbf{S}^2$, S_z) make up a suitable set of correct unperturbed eigenfunctions, since the perturbation matrix is automatically diagonal with respect to them. We have already calculated these eigenstates, and they are given explicitly in formula (17.64). However, because of the identity $2\mathbf{L} \cdot \mathbf{S} = \mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2$, we do not need to know the detailed form of the eigenfunctions if we only wish to evaluate the first-order change in energy due to the spin-orbit interaction:

$$\begin{aligned} \Delta E &= \frac{1}{\hbar} \langle E_n^{(0)},\ \ell\ \tfrac{1}{2}\ jm | W(r)(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2) | E_n^{(0)},\ \ell\ \tfrac{1}{2}\ jm \rangle \\ &= [j(j+1) - \ell(\ell+1) - \tfrac{3}{4}]\hbar \int_0^\infty |R_{n\ell}(r)|^2\ W(r) r^2\ dr \end{aligned} \tag{18.97}$$

or

$$\Delta E = \begin{Bmatrix} \ell\hbar \\ -(\ell+1)\hbar \end{Bmatrix} \int_0^\infty |R_{n\ell}(r)|^2\ W(r) r^2\ dr \quad \text{for} \quad \begin{matrix} j = \ell + \frac{1}{2} \\ j = \ell - \frac{1}{2} \end{matrix} \tag{18.98}$$

From this formula (*Landé's interval rule*) the so-called *fine structure splitting* can be evaluated if $W(r)$ is known.

As was pointed out in Section 16.1, a spin-orbit interaction in an atom arises from the interaction between the Coulomb field of the nucleus and the intrinsic magnetic moment of the orbiting electron. The energy associated with this interaction is

$$H_{\text{magn}} = \boldsymbol{\mu} \cdot \frac{\mathbf{v}}{c} \times \mathbf{E} = \frac{e}{mc^2} \mathbf{S} \cdot \mathbf{v} \times \nabla\phi \tag{18.99}$$

The potential energy due to the central force is $V(r) = -e\phi$. Hence,

$$e\mathbf{E} = -e\nabla\phi = \frac{\mathbf{r}}{r} \frac{dV}{dr}$$

and

$$H_{\text{magn}} = \frac{1}{m^2c^2} \frac{1}{r} \frac{dV}{dr} \mathbf{L} \cdot \mathbf{S} \tag{18.100}$$

When the actual calculation is made with the proper Lorentz transformations for the fields, it is found that owing to purely kinematic effects we must add a term to the energy, which has the same form as (18.100) but a different coefficient. Known as the *Thomas term*, this contribution to the Hamiltonian is

$$H_{\text{Thomas}} = -\frac{1}{2} H_{\text{magn}} \tag{18.101}$$

Such a term is expected whether or not the potential is of electromagnetic origin, and would be present even if the magnetic moment of the electron were zero.[5] The total spin-orbit interaction in atoms is the sum of expressions (18.100) and (18.101),

$$H_{\text{spin-orbit}} = \frac{1}{2m^2c^2} \frac{1}{r} \frac{dV}{dr} \mathbf{L} \cdot \mathbf{S} \tag{18.102}$$

[5]Jackson (1975), Section 11.8.

Hence, the splitting of a level of given ℓ and n into two components with $j = \ell + 1/2$ and $j = \ell - 1/2$ according to (18.98) and (18.102) is determined by the interaction potential

$$W(r) = \frac{\hbar}{4m^2c^2}\frac{1}{r}\frac{dV}{dr} \tag{18.103}$$

For an attractive potential, this quantity is positive; consequently, the level with $j = \ell - 1/2$ lies below that with $j = \ell + 1/2$.

In order to get an idea of the magnitudes involved, we note that if $V(r)$ were a pure Coulomb potential,

$$\langle W(r)\rangle = \frac{\hbar}{4m^2c^2} Ze^2\left\langle\frac{1}{r^3}\right\rangle$$

The evaluation of the fine structure splitting ΔE requires special care for S states ($\ell = 0$), because (18.98) becomes indeterminate ($0 \times \infty$).[6] Since these difficulties will be avoided by a proper calculation with the relativistic Dirac equation in Chapter 24, it is for our present purposes sufficient to estimate for the nth orbit:

$$\left\langle\frac{1}{r^3}\right\rangle \cong \frac{Z^3}{n^3a^3} = \left(\frac{Ze^2m}{\hbar^2 n}\right)^3$$

where a is the Bohr radius of hydrogen. Hence, owing to the Bohr energy formula (12.68),

$$\langle \hbar W\rangle \cong \left(\frac{e^2}{\hbar c}\right)^2 \frac{Z^2}{n}\left|E_n^{(0)}\right|$$

The dimensionless constant,

$$\boxed{\alpha = \frac{e^2}{\hbar c} \approx \frac{1}{137.036}} \tag{18.104}$$

is called the *fine structure constant*. Its small value is responsible for the effectiveness of perturbation methods in atomic physics and for the smallness of the fine structure splitting, $\Delta E_{\ell+1/2} - \Delta E_{\ell-1/2}$, compared with the gross structure energies, $E_n^{(0)}$:

$$\frac{\Delta E_{\ell+1/2} - \Delta E_{\ell-1/2}}{\left|E_n^{(0)}\right|} \cong \frac{(Z\alpha)^2}{n}$$

The fine structure constant α is of much more general significance than its introduction in the present special context suggests. As a pure number that is independent of the units chosen, it has an absolute meaning, measuring the relative magnitudes of the Bohr radius, the Compton wavelength, and the classical radius of the electron:

$$\frac{\hbar^2}{me^2} :: \frac{\hbar}{mc} :: \frac{e^2}{mc^2} = 1 :: \alpha :: \alpha^2$$

If $\hbar$ and c are considered more fundamental than e, the fine structure constant becomes a measure of the electronic charge. This point of view is dominant in quantum

[6]Bethe and Salpeter (1957), Section 13.

field theory where α plays a role of the *coupling constant* between charged matter and electromagnetic fields.

When an atom is exposed to a constant uniform magnetic field, the energy levels split further. If we neglect terms that are quadratic in the field strength (which induce small effects such as diamagnetism) and choose the z axis conveniently in the direction of the external field $\mathbf{B}$, the central-force Hamiltonian (18.96) is perturbed by an interaction energy

$$H' = \frac{2}{\hbar} W(r)\mathbf{L} \cdot \mathbf{S} + \frac{eB}{2mc}(g_L L_z + g_S S_z) \tag{18.105}$$

Equation (17.101) has been used for the magnetic moment of the atom, with $g_L = 1$ and $g_S = 2$ for a one-electron atom.

If the matrix elements of H' are small compared with the energy-level separations of H_0, first-order perturbation theory requires, according to Section 18.5, that we solve the eigenvalue problem of H' in a space of degenerate eigenstates of H_0. For a general atom, the zero-order energies are labeled by a set of quantum numbers of which only ℓ and s are relevant for the present purpose, because the corresponding operators $\mathbf{L}^2$ and $\mathbf{S}^2$ commute with H'. (Unlike an electric field, a magnetic field does not mix states of different ℓ.)

In the $(2\ell + 1)(2s + 1)$-dimensional vector space characterized by the quantum numbers ℓ and s, the perturbation may be written in the form

$$H' = \lambda \mathbf{L} \cdot \mathbf{S} + \mu L_z + \nu S_z \tag{18.106}$$

which, appropriately interpreted, is sufficiently general to render the present discussion applicable to a large class of problems involving complex as well as one-electron atoms, and a variety of other types of angular momentum couplings, such as hyperfine interactions. The coefficients λ, μ, ν in (18.106) are real constants. Their values are assumed to have been computed separately.

The eigenvalue problem posed by the effective Hamiltonian H' may be solved in any representation, but a judicious choice of basis can save much labor. In order to minimize the work of computing matrix elements, it is desirable to use as many constants of the motion as possible for specifying the complete set of operators whose eigenvectors are to serve as basis. We thus see the advisability of selecting $\mathbf{L}^2$, $\mathbf{S}^2$, $\mathbf{J}^2$, J_z as the appropriate complete set spanning the coupled representation $|\ell sjm\rangle$. Only $\mathbf{J}^2$ fails to commute with H', and the only nonvanishing off-diagonal matrix elements are, therefore, of the form $\langle \ell sj'm|H'|\ell sjm\rangle$. In fact, since H' can be rewritten as

$$H' = \frac{\lambda}{2}(\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2) + \mu J_z + (\nu - \mu)S_z \tag{18.107}$$

we only have to evaluate $\langle \ell sj'm|S_z|\ell sjm\rangle$.

The evaluation of this matrix element of S_z is accomplished by using the defining relation (17.52) of the C-G coefficients. We obtain

$$\langle \ell sj'm|S_z|\ell sjm\rangle = \hbar \sum_{m_s} m_s \langle m - m_s, m_s|j'm\rangle\langle m - m_s, m_s|jm\rangle \tag{18.108}$$

We now specialize to the case $s = 1/2$, which includes one-electron atoms. Using Eqs. (17.63) in (18.108), we obtain

$$\langle \ell\, \tfrac{1}{2}, \ell \pm \tfrac{1}{2}, m|S_z|\ell\, \tfrac{1}{2}, \ell \pm \tfrac{1}{2}, m\rangle = \pm\frac{m\hbar}{2\ell + 1} \tag{18.109}$$

for the diagonal matrix elements, and

$$\langle \ell\ \tfrac{1}{2},\ \ell \pm \tfrac{1}{2},\ m | S_z | \ell\ \tfrac{1}{2},\ \ell \mp \tfrac{1}{2},\ m \rangle = -\frac{\hbar}{2\ell + 1}\sqrt{(\ell + \tfrac{1}{2})^2 - m^2} \quad (18.110)$$

for the off-diagonal matrix elements.

As an important example, we finally consider a ^2P state ($\ell = 1$) with total angular momentum $j = 3/2$ and $j = 1/2$. The manifold of substates now is six-dimensional. Using the symbolic notation of Eq. (17.71), we write the equivalence of the uncoupled and coupled basis in the six-dimensional space as

$$\mathbf{3} \otimes \mathbf{2} = \mathbf{4} \oplus \mathbf{2} \quad (18.111)$$

The operator S_z has nonvanishing off-diagonal matrix elements only between the basis states $|jm\rangle = |\frac{1}{2}, \frac{1}{2}\rangle$ and $|\frac{3}{2}, \frac{1}{2}\rangle$ and between the states $|jm\rangle = |\frac{1}{2}, -\frac{1}{2}\rangle$ and $|\frac{3}{2}, -\frac{1}{2}\rangle$. The states $|jm\rangle = |\frac{3}{2}, \frac{3}{2}\rangle$ and $|jm\rangle = |\frac{3}{2}, -\frac{3}{2}\rangle$ are eigenvectors of H', and the remaining eigenvectors are of the form

$$a|\tfrac{3}{2}, \tfrac{1}{2}\rangle + b|\tfrac{1}{2}, \tfrac{1}{2}\rangle \quad \text{and} \quad c|\tfrac{3}{2}, -\tfrac{1}{2}\rangle + d|\tfrac{1}{2}, -\tfrac{1}{2}\rangle$$

The eigenvalues of H' are corrections to the unperturbed energy, which take the interactions among the six basis states of the ^2P manifold exactly into account, while totally ignoring the interactions involving all other basis states. Denoted by E_m, the perturbed energies are

$$E_{\pm 3/2} = \frac{\lambda}{2}\hbar^2 \pm \left(\mu + \frac{\nu}{2}\right)\hbar \quad (18.112)$$

and the roots of the secular equation

$$\begin{vmatrix} \frac{\lambda}{2}\hbar^2 \pm \frac{1}{6}(2\mu + \nu)\hbar - E_{\pm 1/2} & \frac{\sqrt{2}}{3}(\mu - \nu)\hbar \\ \frac{\sqrt{2}}{3}(\mu - \nu)\hbar & -\lambda\hbar^2 \pm \frac{1}{6}(4\mu - \nu)\hbar - E_{\pm 1/2} \end{vmatrix} = 0 \quad (18.113)$$

There are four roots, two for $m = 1/2$ and two for $m = -1/2$.

Exercise 18.17. Check the eigenvalue Eqs. (18.112) and (18.113).

For the case of a one-electron atom ($g_L = 1$, $g_S = 2$), the perturbed ^2P energies are plotted in Figure 18.1 as a function of the applied magnetic field of strength B. The resulting splitting of spectral lines is known as the *Zeeman effect*. The operator J_z commutes with H' for all values of the magnetic field, and m is thus a "*good quantum number*" throughout. In conformity with the no-crossing rule discussed in the last section, only levels belonging to different values of m are seen to cross for finite B.

Exercise 18.18. Carry out the calculation of the Zeeman effect in detail. Show that, if the external field is either very small or very large compared with the internal magnetic field of the atom, the energy depends linearly on B. Show that the internal field which produces the fine structure is of the order 1 T or 10^4 gauss.

7. *The Variational Method and Perturbation Theory.* In many applications of quantum mechanics, the Schrödinger equation cannot be solved rigorously, nor can

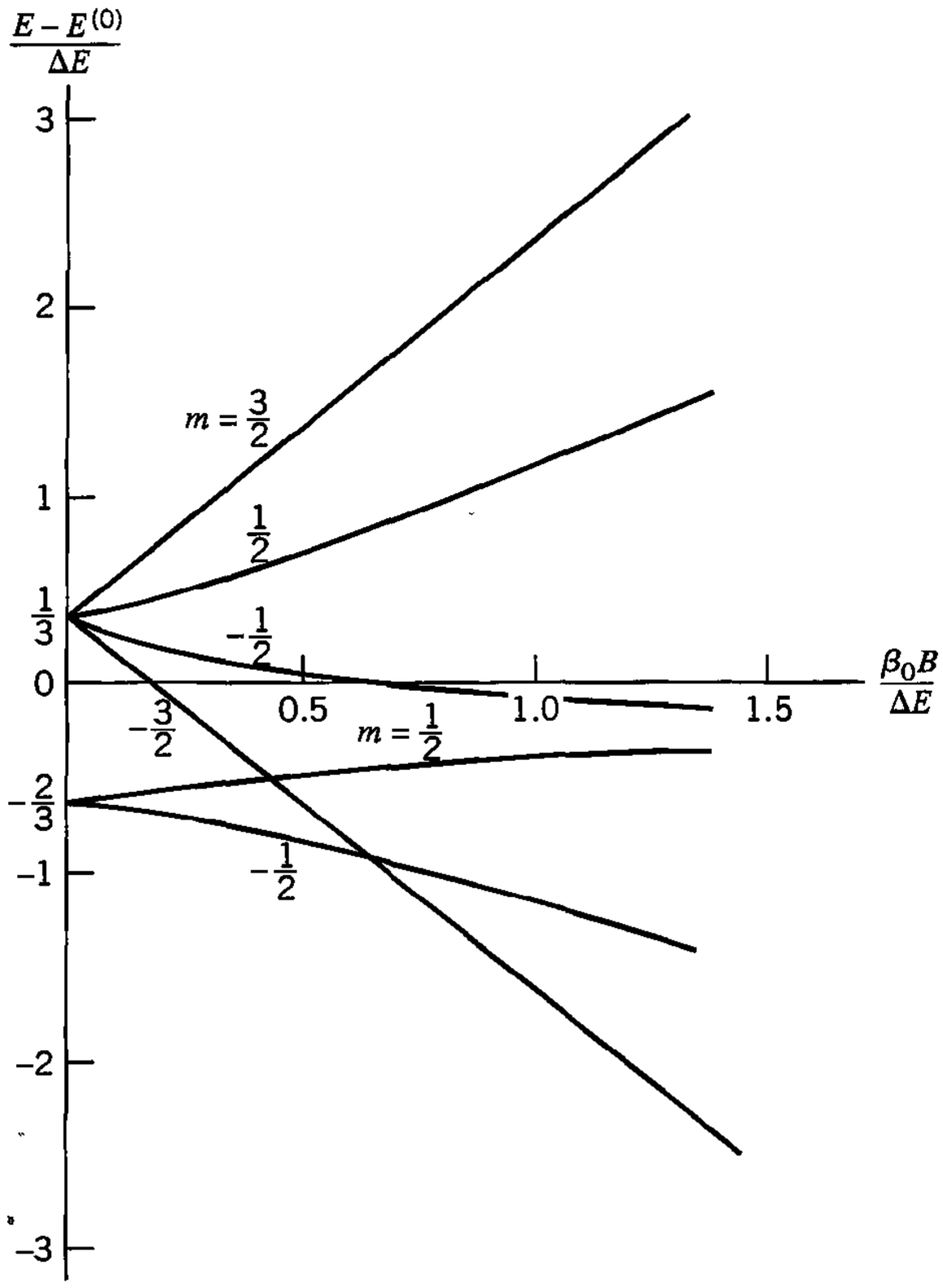


Figure 18.1. Splitting of a ^2P energy level of hydrogen in a magnetic field **B** (Zeeman effect). The unperturbed energy is $E^{(0)}$, and the fine structure splitting is ΔE.

a neighboring unperturbed Hamiltonian be found which affords a good approximation and a suitable starting point for a perturbation treatment. One then often falls back on the equivalence of the eigenvalue problem with a variational principle for H. For the standard one-particle Schrödinger equation, this equivalence was demonstrated and illustrated in Chapter 8. A more general proof was the subject of Section 10.3. Here, we repeat a shorthand derivation of this fundamental proposition.

First, we show that the eigenvectors of H are the vectors for which the variation of the expectation value

$$\langle H \rangle = \frac{(\Psi, H\Psi)}{(\Psi, \Psi)} \tag{18.114}$$

vanishes. Generalizing the argument in Section 8.1, we have for a small change $\delta\Psi$,

$$\begin{aligned}(\Psi, \Psi)^2 \delta\langle H \rangle = {} & (\Psi, \Psi)[(\delta\Psi, H\Psi) + (\Psi, H\delta\Psi)] \\ & -(\Psi, H\Psi)[(\delta\Psi, \Psi) + (\Psi, \delta\Psi)] + O[(\delta\Psi)^2]\end{aligned} \tag{18.115}$$

Since H is Hermitian, the vectors $\Psi = \Psi_n$ which make $\delta\langle H \rangle = 0$ to first order in $\delta\Psi$ must satisfy an equation

$$H\Psi_n = E_n\Psi_n \tag{18.116}$$

Conversely, every eigenvector of H makes the variation of $\langle H \rangle$ zero. The stationary values of $\langle H \rangle$ are the eigenvalues E_n.

As an almost trivial application of the variational method, consider a Rayleigh-Schrödinger perturbation calculation of an energy eigenvalue E_n of a system with Hamiltoniann H. Using the notation of Section 18.1 (with the expansion parameter g), we observe that the kth approximation to the eigenstate,

$$\Psi_n \approx \Psi_n^{(0)} + g\Psi_n^{(1)} + g^2\Psi_n^{(2)} + \cdots + g^k\Psi_n^{(k)} \tag{18.117}$$

differs from the correct eigenstate Ψ_n by terms of order g^{k+1}. If the approximate vector (18.117) is used as a trial vector in (18.114), we see from (18.115) that the value of $\langle H \rangle$ will differ from E_n by terms of order g^{2k+2}. It follows that $\langle H \rangle$ computed with the trial vector (18.117) gives E_n correctly to order g^{2k+1}. This is a strong and useful result.

Exercise 18.19. Illustrate the connection between the variational and the Rayleigh-Schrödinger perturbation methods by applying a trial vector

$$\Psi_n = \Psi_n^{(0)} + g\Psi_n^{(1)}$$

to approximate an energy eigenstate of the perturbed Hamiltonian $H = H_0 + gV$. Compare with the results of Exercise 18.6.

In practice, variational estimates are most easily and reliably made for the ground state of a system. Since we cannot extract infinite amounts of energy from any physical system, the expectation value of the energy must have a lower bound. Hence, if H represents the energy, the absolute minimum of $\langle H \rangle$ is the lowest energy eigenvalue E_0, and this value is reached if the trial vector used to estimate $\langle H \rangle$ is an eigenvector corresponding to E_0.

In principle, the variational method can be used for energy eigenstates other than the ground state. We already saw in Section 10.3 how it can be applied to eigenstates of successively higher energy by gradually narrowing the vector space in which H operates. If the ground state Ψ_0 has been obtained, its subspace is split off and the variational procedure is repeated in the remaining subspace of all vectors that are orthogonal to Ψ_0. Obviously, the method will give more and more inaccurate results if Ψ_0 is known only approximately. This difficulty can be circumvented if it is possible to construct a subset of trial vectors that are known to be rigorously orthogonal to the exact Ψ_0. For instance, in the case of a particle in a central field, where the ground state is an S state, we restrict ourselves to the use of spherically symmetric trial functions to estimate E_0. Similarly, by using P-state wave functions, we can obtain an upper limit for the lowest P state, and so on.

If we know some constant of the motion A, such that

$$[H, A] = 0 \tag{18.118}$$

the projection operator $P_{A_i'}$, which projects any vector into the subspace spanned by the eigenvectors of A with eigenvalue A_i' is also a constant of the motion. Hence,

$$[H, P_{A_i'}] = 0 \tag{18.119}$$

From this commutation relation, it follows that if Ψ_n is an eigenvector of H, $P_{A_i'}\Psi_n$ (if it is not zero) must also be an eigenvector of H, corresponding to the same

eigenvalue, E_n. By limiting ourselves to variations in the subspace characterized by A_i', we thus establish that the condition

$$\delta \frac{(P_{A_i}\Psi, HP_{A_i}\Psi)}{(P_{A_i}\Psi, P_{A_i}\Psi)} = 0 \tag{18.120}$$

generates the energy eigenvalues and eigenvectors that carry the label A_i'. In particular, the absolute minimum of $\langle H\rangle$ in the subspace of the vectors $P_{A_i'}\Psi$ gives us the lowest energy eigenvalue with the property A_i'.

A generalization of the variational principle can be formulated if we vary not only the state vector Ψ but also the Hamiltonian H. Denoting such generalized variations by the symbol Δ, and neglecting terms that are quadratic in Δ, we have by the usual rule concerning the variation of a product:

$$\Delta\langle H\rangle = \Delta \frac{(\Psi, H\Psi)}{(\Psi, \Psi)} = \frac{(\Psi, \Delta H\Psi)}{(\Psi, \Psi)} + \delta \frac{(\Psi, H\Psi)}{(\Psi, \Psi)} \tag{18.121}$$

where the δ-variation is one in which the operator is kept fixed, while Ψ is varied. If the variation is performed near an eigenvector Ψ_n, the last term vanishes, and we are left with

$$\boxed{\Delta \frac{(\Psi_n, H\Psi_n)}{(\Psi_n, \Psi_n)} = \frac{(\Psi_n, \Delta H\Psi_n)}{(\Psi_n, \Psi_n)}} \tag{18.122}$$

Conversely, any vector Ψ_n for which (18.122) holds is an eigenvector of H.

So far, the variations of Ψ and H have been independent. If we now restrict them by requiring that Ψ remain an eigenvector of H during the variation, $\Delta\langle H\rangle$ becomes the change ΔE_n of the eigenvalue E_n, and we get

$$\boxed{\Delta E_n = \frac{(\Psi_n, \Delta H\Psi_n)}{(\Psi_n, \Psi_n)}} \tag{18.123}$$

which, in slightly different notation, is identical with the energy shift (18.9) in first-order perturbation theory.

From (18.122) and (18.123), the variational theorem may be re-expressed in a generalized form as follows: Ψ is an eigenvector of H if and only if $\Delta\langle H\rangle$ depends only on the change in H, and not on the change in Ψ. When condition (18.122) is achieved, the value of $\Delta\langle H\rangle$ is, to first order, equal to the change of the corresponding energy eigenvalue.

Exercise 18.20. Prove the Hellmann-Feynman theorem (18.83), using a variational argument.

Exercise 18.21. Combining the virial theorem and conservation of energy with the Hellmann-Feynman theorem, prove that the energy of a bound state for the potential $V = Ar^n$ scales as $A^{2/(n+2)}$. Also show that E scales with the particle mass as $m^{-n/(n+2)}$, if the potential is independent of m. Check these results for the Coulomb potential, the linear potential, and the harmonic oscillator.

8. ***The Helium Atom.*** The neutral helium atom with a fixed nucleus is described by the Schrödinger equation in configuration space:

$$\left[-\frac{\hbar^2}{2m}(\nabla'^2 + \nabla''^2) - \frac{2e^2}{r'} - \frac{2e^2}{r''} + \frac{e^2}{|\mathbf{r}' - \mathbf{r}''|}\right]\psi(\mathbf{r}', \mathbf{r}'') = E\psi(\mathbf{r}', \mathbf{r}'') \tag{18.124}$$

The coordinates of the two electrons are labeled $\mathbf{r}'$ and $\mathbf{r}''$ under the provisional assumption that the particles are in principle distinguishable. We know this assumption to be false but, since electrons have spin one-half with two possible substates, and to the extent that spin-orbit interactions can be neglected, the fiction of two (but no more!) distinguishable electrons can be maintained in solving Eq. (18.124). We may think of one electron as having spin "up" and the other spin "down," so that the two particles belong in effect to two different, distinguishable electron species. We will see that with this assumption we can obtain the entire energy spectrum (but not all the states) of a two-electron system. The corrections to this picture, imposed by the identity of the two electrons in the same spin state and by the Pauli exclusion principle, will be found in Chapter 22.

Owing to the symmetry of the differential operator in (18.124), the solutions of the Schrödinger equation

$$H(\mathbf{r}', \mathbf{r}'')\psi(\mathbf{r}', \mathbf{r}'') = E\psi(\mathbf{r}', \mathbf{r}'') \tag{18.125}$$

fall naturally into two classes: Every solution can be assumed to be either symmetric or antisymmetric in the space coordinates:

$$\psi(\mathbf{r}'', \mathbf{r}') = \pm\psi(\mathbf{r}', \mathbf{r}'') \tag{18.126}$$

If a particular solution $f(\mathbf{r}', \mathbf{r}'')$ fails to have this property, we observe that $f(\mathbf{r}'', \mathbf{r}')$ is also a solution—a phenomenon that is referred to as *exchange degeneracy*, because both solutions belong to the same energy eigenvalue. Hence, owing to the linearity of the Schrödinger equation, the linear combinations

$$\psi_{\pm}(\mathbf{r}', \mathbf{r}'') = \frac{1}{2}\left[f(\mathbf{r}', \mathbf{r}'') \pm f(\mathbf{r}'', \mathbf{r}')\right]$$

are also solutions. Since conversely every solution can be written as a superposition of solutions with definite exchange symmetry,

$$f(\mathbf{r}', \mathbf{r}'') = \psi_+(\mathbf{r}', \mathbf{r}'') + \psi_-(\mathbf{r}', \mathbf{r}'')$$

we lose nothing by confining our attention to the *symmetric* and *antisymmetric* solutions of the two-electron Schrödinger equation (18.124).

It is tempting to neglect in lowest approximation the repulsive interaction between the electrons and treat the neglected part of the potential energy as a perturbation. Since the repulsive term is clearly of the same order of magnitude as the Coulomb attraction due to the nucleus, there is no *a priori* justification for this procedure other than the simplification which results from it, providing a qualitative understanding of the level scheme of helium.

The simplification results because the approximate zero-order Schrödinger equation,

$$\left[-\frac{\hbar^2}{2m}(\nabla'^2 + \nabla''^2) - \frac{2e^2}{r'} - \frac{2e^2}{r''}\right]\psi^{(0)}(\mathbf{r}', \mathbf{r}'') = E^{(0)}\psi^{(0)}(\mathbf{r}', \mathbf{r}'')$$

is further separable, if we assume that $E^{(0)} = E' + E''$ and that

$$\psi^{(0)}(\mathbf{r}', \mathbf{r}'') = g(\mathbf{r}')h(\mathbf{r}'')$$

The separated equations are both hydrogenic:

$$\left(-\frac{\hbar^2}{2m}\nabla'^2 - \frac{2e^2}{r'}\right)g(\mathbf{r}') = E'g(\mathbf{r}') \qquad \left(-\frac{\hbar^2}{2m}\nabla''^2 - \frac{2e^2}{r''}\right)h(\mathbf{r}'') = E''h(\mathbf{r}'')$$

The lowest unperturbed energy level of the atom ($Z = 2$) corresponds to $E' = E'' = -Z^2 \times 13.6 \text{ eV} = -4 \times 13.6 \text{ eV} = -54.4 \text{ eV}$. The corresponding unperturbed ground state energy eigenfunction is

$$\psi_+^{(0)} = \frac{Z^3}{\pi a^3} \exp\left[-\frac{Z(r' + r'')}{a}\right] \tag{18.127}$$

which is a product of two hydrogenic $1s$ wave functions. This is symmetric under exchange.

The first-order correction to the energy is

$$\Delta E = \int\int |\psi_+^{(0)}(\mathbf{r}', \mathbf{r}'')|^2 \frac{e^2}{|\mathbf{r}' - \mathbf{r}''|} d^3r' \, d^3r'' = \frac{5e^2}{4a} = \frac{5}{2} \times 13.6 = 34 \text{ eV}$$

The total ground state energy in first approximation is then

$$E_{\text{cal}} = E^{(0)} + \langle V(\mathbf{r}', \mathbf{r}'')\rangle = -2 \times 54.4 + 34 = -74.8 \text{ eV}$$

The measured ionization potential of helium is 24.46 V. Hence, the total energy of the atom is

$$E_{\text{obs}} = -54.4 - 24.5 = -78.9 \text{ eV}$$

$E_{\text{cal}} > E_{\text{obs}}$ in agreement with the variational principle.

The agreement between theory and experiment can be improved by using a better trial function than (18.127). The design of suitable trial functions has been the subject of ceaseless investigation ever since quantum mechanics was invented. The problem offers a challenge to the theoretician because the two-electron system is mathematically manageable, although solutions in closed form cannot be found. Notable efforts have been made in devising better and better variational trial functions for the ground state of helium, the test of their quality being how closely $\langle H\rangle$ lies above the precisely measured energy.

We content ourselves with describing the simplest variational method. It uses trial functions of the form (18.127) but with Z replaced by an effective nuclear charge, $Z_{\text{eff}} = Z - \sigma$. This is physically reasonable, because each electron is partially screened from exposure to the full charge of the nucleus by the presence of the other electron. If σ is left arbitrary, we can calculate the expectation value of the energy

$$E(\sigma) = \langle H\rangle$$

where H is the complete Hamiltonian of the helium atom, including the interaction between the electrons. According to the variational theorem, $E(\sigma) \geq E$ for any value of σ. The best estimate with a trial function of the form (18.127) is therefore the lowest upper bound, obtained by minimizing $E(\sigma)$ with respect to σ. The result, for any two-electron atom with nuclear charge Z, is

$$E(\sigma) = -\frac{e^2}{a}\left(Z^2 - \frac{5}{8}Z + \frac{5}{8}\sigma - \sigma^2\right)$$

Exercise 18.22. Work out this result. Show that the minimum of $E(\sigma)$ is obtained for $\sigma = 5/16$, suggesting that in the ground state the screening effect from one electron amounts to one-third of an electron charge e. Show that the minimum value is

$$E\left(\frac{5}{16}\right) = -\left(Z - \frac{5}{16}\right)^2 \frac{e^2}{a} \tag{18.128}$$

For helium, how much of an improvement is this over the earlier crude estimate?

Exercise 18.23. Show that if an unperturbed Hamiltonian

$$H_0 = -\frac{\hbar^2}{2m}(\nabla'^2 + \nabla''^2) - \frac{(2 - 5/16)e^2}{r'} - \frac{(2 - 5/16)e^2}{r''} \tag{18.129}$$

is chosen, the first-order perturbation energy of the ground state of helium vanishes.

In lowest approximation, the excited states of helium are obtained by lifting one electron to an excited hydrogenic level and leaving the other electron in the $1s$ level. The corresponding zero-order wave functions are

$$\psi_{\pm}^{(0)} = \frac{1}{\sqrt{2}}[\psi_{100}(\mathbf{r}')\psi_{n\ell m_\ell}(\mathbf{r}'') \pm \psi_{n\ell m_\ell}(\mathbf{r}')\psi_{100}(\mathbf{r}'')] \tag{18.130}$$

with an energy $E^{(0)} = E_0 + E_n$. The unperturbed energy levels are degenerate exactly as in the hydrogen atom, but the interaction between the two electrons removes the ℓ degeneracy. The unperturbed eigenfunctions (18.130) are just the correct linear combinations to begin a degenerate perturbation treatment: The perturbation potential is diagonal with respect to any two states of the type (18.130), because total orbital angular momentum $\mathbf{L}$ is a constant of the motion and because two states of different exchange symmetry cannot mix.

The excited states thus split into components according to their symmetry under exchange of the particle coordinates. The energy correction due to the symmetric perturbation $V(\mathbf{r}', \mathbf{r}'')$ is

$$\Delta E_{\pm} = \int\int \psi_{\pm}^{(0)*}(\mathbf{r}', \mathbf{r}'')V(\mathbf{r}', \mathbf{r}'')\psi_{\pm}^{(0)}(\mathbf{r}', \mathbf{r}'')\, d^3r'\, d^3r'' = I \pm J \tag{18.131}$$

Here the so-called *direct integral*,

$$I = \int\int \psi_{100}^*(\mathbf{r}')\psi_{n\ell m_\ell}^*(\mathbf{r}'')V(\mathbf{r}', \mathbf{r}'')\psi_{100}(\mathbf{r}')\psi_{n\ell m_\ell}(\mathbf{r}'')\, d^3r'\, d^3r'' \tag{18.132}$$

is the expectation value of the interaction energy, which would be the entire first-order perturbation correction if exchange degeneracy were absent from the unperturbed Hamiltonian, or if ψ_{100} and $\psi_{n\ell m}$ did not overlap. The other term,

$$J = \int\int \psi_{100}^*(\mathbf{r}')\psi_{n\ell m_\ell}^*(\mathbf{r}'')V(\mathbf{r}', \mathbf{r}'')\psi_{n\ell m_\ell}(\mathbf{r}')\psi_{100}(\mathbf{r}'')\, d^3r'\, d^3r'' \tag{18.133}$$

bears the descriptive name *exchange integral.*

Exercise 18.24. Verify explicitly that

$$\Delta E_{\pm} = I \pm J$$

Exercise 18.25. To give an interpretation of the two nonstationary states

$$\psi_{100}(\mathbf{r}')\psi_{n\ell m_\ell}(\mathbf{r}'') \quad \text{and} \quad \psi_{n\ell m_\ell}(\mathbf{r}')\psi_{100}(\mathbf{r}'')$$

show that if the atom is initially in one of these, it shuttles back and forth with an exchange frequency $J/\hbar$. [Compare (8.81).]

For our repulsive interaction potential, $V(\mathbf{r}', \mathbf{r}'') = e^2/|\mathbf{r}' - \mathbf{r}''|$, the direct integral I is positive. Calculation shows that the exchange integral J has a small positive value in this case. Hence, all (spatially symmetric) singlet terms lie slightly higher than the corresponding (spatially antisymmetric) triplet terms. This behavior

is understandable, since in the spatially antisymmetric state the probability is small for the electrons to be found near each other, while in the spatially symmetric state they have a greater opportunity to repel each other, thereby raising the energy.

Since, owing to the different spatial symmetries, helium atoms in singlet states are physically quite different from those in triplet states, a special terminology is often used. Helium atoms in singlet states are said to form *parahelium*, whereas those in triplet states form *orthohelium*. The lowest energy level of parahelium is the normal ground state.

Problems

1. The Hamiltonian of a rigid rotator in a magnetic field perpendicular to the x axis is of the form $A\mathbf{L}^2 + BL_z + CL_y$, if the term that is quadratic in the field is neglected. Obtain the exact energy eigenvalues and eigenfunctions of the Hamiltonian. Then, assuming $B \gg C$, use second-order perturbation theory to get approximate eigenvalues and compare these with the exact answers.

2. A charged particle is constrained to move on a spherical shell in a weak uniform electric field. Obtain the energy spectrum to second order in the field strength.

3. Apply perturbation theory to the elastically coupled harmonic oscillators of Problems 6 and 7 in Chapter 15, assuming that the interaction between the two particles is weak, and compare with the rigorous solutions.

4. Use second-order perturbation theory to calculate the change in energy of a linear harmonic oscillator when a constant force is added, and compare with the exact result.

5. A slightly anisotropic three-dimensional harmonic oscillator has $\omega_z \approx \omega_x = \omega_y$. A charged particle moves in a field of this oscillator and is at the same time exposed to a uniform magnetic filed in the x direction. Assuming that the Zeeman splitting is comparable to the splitting produced by the anisotropy, but small compared to $\hbar\omega$, calculate to first order the energies of the components of the first excited state. Discuss various limiting cases.

6. Prove that if $\psi_0 = e^{-\varphi_0}$ is a positive bounded function satisfying appropriate boundary conditions, it represents the ground state for a particle moving in a potential

$$V = \frac{\hbar^2}{2m}[(\nabla\varphi_0)^2 - \nabla^2\varphi_0] + E_0$$

and the corresponding energy is E_0. Verify the theorem for (a) the isotropic harmonic oscillator and (b) the hydrogen atom.

7. Apply the theorem proven in Problem 6 to a particle of mass m in one dimension, with the assumed ground state,

$$\psi_0(x) = \frac{C}{\cosh \kappa x}$$

and a corresponding energy eigenvalue $E_0 = -\hbar^2\kappa^2/2m$. Plot the resulting potential.

8. Prove that the trace of the direct product of two matrices equals the product of the traces of the matrices. Apply this result to show that the ''center of gravity'' of a multiplet split by the spin-orbit interaction is at the position of the unperturbed energy level.

9. Obtain the relativistic correction $\propto p^4$ to the nonrelativistic kinetic energy of an electron, and, using first-order perturbation theory, evaluate the energy shift that it produces in the ground state of hydrogen.

10. Using the Hamiltonian for an atomic electron in a magnetic field, determine, for a state of zero angular momentum, the energy change to order $\mathbf{B}^2$, if the system is in a uniform magnetic field represented by the vector potential $\mathbf{A} = \mathbf{B} \times \mathbf{r}/2$.

Defining the atomic diamagnetic susceptibility χ by $E = -\chi \mathbf{B}^2/2$, calculate χ for a helium atom in the ground state and compare the result with the measured value.

11. Apply second-order perturbation theory to a one-dimensional periodic perturbing potential

$$V(x) = \sum_{n=-\infty}^{+\infty} V_n e^{2\pi i n x/\xi}$$

with period ξ. To enforce closely spaced discrete energies, assume that the entire "crystal" has length $L = N\xi$ and use the periodic boundary condition $\Psi(x + L/2) = \Psi(x - L/2)$, where N is a large even number. Assume that the zero-order Hamiltonian is that of a free particle subject to these boundary conditions. Show that nondegenerate perturbation theory breaks down at the band edges and that the forbidden energy gaps are proportional to the Fourier coefficients of the potential.

Show that in the limit $L \to \infty$, carried out as in Section 1 of the Appendix, the unperturbed (unnormalized) energy eigenfunctions are plane waves, e^{ikx}, and derive the second-order approximation to the dispersion function $E(k)$.

12. Assuming that the zero-order free particle energy is large compared with g/ξ, apply the results of Problem 11 to the special example of a Kronig-Penney potential with attractive δ-functions of strength g. In the middle of the valence band at $k\xi = \pi/2$, verify the result of perturbation theory by expanding the exact eigenvalue condition (see Exercise 8.29).

13. Let ψ be a variational trial function for the ground state ψ_0 of a system with nondegenerate energy eigenvalues. Assume that ψ and ψ_0 are real, normalized to unity, and that $\int \psi\psi_0 \, d^3r$ is positive. Show that

$$\frac{1}{2}\int (\psi - \psi_0)^2 \, d^3r \leq 1 - \left(1 - \frac{\langle H\rangle - E_0}{E_1 - E_0}\right)^{1/2} \approx \frac{1}{2}\frac{\langle H\rangle - E_0}{E_1 - E_0}$$

where E_0 and E_1 are the exact energies of the ground and first excited states, and $\langle H\rangle$ is the expectation value of the Hamiltonian in the state ψ. Estimate the accuracy of the trial functions used in Section 17.8 for the ground state of the helium atom.

14. A rotator whose orientation is specified by the angular coordinates θ and φ performs a *hindered rotation* described by the Hamiltonian

$$H = A\mathbf{L}^2 + B\hbar^2 \cos 2\varphi$$

with $A \gg B$. Calculate the S, P, and D energy levels of this system in first-order perturbation theory, and work out the corresponding unperturbed energy eigenfunctions.

CHAPTER 19

Time-Dependent Perturbation Theory

When the nature of a physical system and its Hamiltonian is known, we are often confronted with the problem of predicting the time development of the system from its initial condition. The general formalism of quantum dynamics was established in Chapter 14 and applied in Chapter 15 to some problems solvable in closed form. Commonly, however, one must resort to perturbation methods. The interaction picture lends itself naturally to the formulation of perturbative schemes. While deferring a systematic treatment of collisions, which are typical time-dependent processes, to Chapter 20, here we study the relatively weak interaction between electromagnetic fields and charged-particle systems as illustration of time-dependent perturbation theories. In the examples, we usually do not go beyond the first nonvanishing order. Fermi's Golden Rule for constant transition rates and its relation to exponential decay emerges as a central result.

1. ***The Equation of Motion in the Interaction Picture.*** If H is a time-independent Hamiltonian and all solutions of the eigenvalue problem

$$H|\Psi_n\rangle = E_n|\Psi_n\rangle \tag{19.1}$$

are known, it is straightforward to construct the solution of the equation of motion,

$$i\hbar \frac{d|\Psi(t)}{dt} = H|\Psi(t)\rangle \tag{19.2}$$

which coincides at $t = t_0$ with the given initial state $|\Psi(t_0)\rangle$. Under these circumstances, the state vector

$$|\Psi(t)\rangle = e^{-(i/\hbar)H(t-t_0)}|\Psi(t_0)\rangle = \sum_n e^{-(i/\hbar)E_n(t-t_0)}|\Psi_n\rangle\langle\Psi_n|\Psi(t_0)\rangle \tag{19.3}$$

represents the general solution of the equation of motion, (19.2). For example, in Section 13.2 this procedure was applied to the scattering of a wave packet by a fixed potential.

Usually, H is a complicated operator, and rigorous solutions of (19.1) are not available except in special cases. Stationary state perturbation theory could be used to obtain approximate eigenvalues and eigenstates of H for substitution in (19.3). Although this procedure is in principle perfectly suitable if H is time-independent, in this chapter we will take up a mathematically simpler approach, which has the additional advantage of lending itself to immediate physical interpretation.

Since our perturbation method will be applied directly to the *time-dependent* equation (19.2), H need not be a time-independent operator. Indeed, perturbations that are explicitly dependent on the time are of considerable importance, as illustrated by the forced linear harmonic oscillator. The excitation or ionization of an

atom, molecule, or nucleus by a varying electric field is another obvious example, and we will work out the cross sections for excitation by a passing charged particle and for absorption of light. When H depends on t, there are no strictly stationary states, and the energy of the system alone is not conserved. Hence, the simple formula (19.3) is inapplicable.

It is characteristic of the physical processes mentioned—scattering of a particle by a fixed potential and absorption of energy by a system from a passing disturbance—that the perturbing interaction is limited in space and time. The system is unperturbed before it enters the scattering region or before it is hit by a wave, and it is again free from the perturbation after sufficient time has elapsed. This sequence of events suggests that the total Hamiltonian again be considered as the sum of two terms:

$$H = H_0 + V \tag{19.4}$$

where the time-independent operator H_0 describes the unperturbed system, and V is the perturbation that may be explicitly time-dependent (as in the case of excitation by a transient electromagnetic field) or not (as in the case of potential scattering). In either case, the initial state is usually, at least approximately, an eigenstate of H_0. The forced linear harmonic oscillator, for which exact solutions were worked out in Section 14.6, may serve to test any approximation method to be developed.

The basic idea of time-dependent perturbation theory is simple. It assumes knowledge of the solutions of the unperturbed eigenvalue problem,

$$H_0|n\rangle = E_n|n\rangle \tag{19.5}$$

Our objective is to calculate transition amplitudes between the relevant unperturbed eigenstates, owing to the presence of the perturbation V.

The approximations are most readily carried out in the interaction picture. We therefore recapitulate the essential formulas from Section 14.2, modifying the notation slightly for our present purposes. The equation of motion in the Schrödinger picture,

$$i\hbar \frac{d|\Psi(t)\rangle}{dt} = (H_0 + V)|\Psi(t)\rangle \tag{19.6}$$

is transformed to the interaction picture (indicated as in Chapter 14 by a tilde over the state vectors and operators) by a time-dependent unitary operator:

$$|\tilde{\Psi}(t)\rangle = \exp\left(\frac{i}{\hbar} H_0 t\right)|\Psi(t)\rangle \tag{19.7}$$

In the interaction picture, the equation of motion for the state is

$$i\hbar \frac{d|\tilde{\Psi}(t)\rangle}{dt} = \tilde{V}(t)|\tilde{\Psi}(t)\rangle \tag{19.8}$$

where the new interaction operator is given by

$$\tilde{V}(t) = \exp\left(\frac{i}{\hbar} H_0 t\right) V \exp\left(-\frac{i}{\hbar} H_0 t\right) \tag{19.9}$$

In order to derive general perturbation theory expressions, independent of the particular choice of an initial state, we introduce the time development operator $\tilde{T}(t, t_0)$ such that

$$|\tilde{\Psi}(t)\rangle = \tilde{T}(t, t_0)|\tilde{\Psi}(t_0)\rangle \tag{19.10}$$

The time development operator satisfies the integral equation

$$\tilde{T}(t, t_0) = I - \frac{i}{\hbar}\int_{t_0}^{t} \tilde{V}(t')\tilde{T}(t', t_0)\, dt' \tag{19.11}$$

This operator is related to the time development operator, $T(t, t_0)$, in the Schrödinger picture by

$$\tilde{T}(t, t_0) = e^{(i/\hbar)H_0 t}T(t, t_0)e^{-(i/\hbar)H_0 t_0} \tag{19.12}$$

We are interested in the transition matrix elements of these time development operators between the eigenstates of the unperturbed Hamiltonian. In terms of matrix elements, (19.11) becomes

$$\langle k|\tilde{T}(t, t_0)|s\rangle = \delta_{ks} - \frac{i}{\hbar}\sum_n \int_{t_0}^{t} \exp\left[\frac{i}{\hbar}(E_k - E_n)t'\right]\langle k|V(t')|n\rangle\langle n|\tilde{T}(t', t_0)|s\rangle\, dt' \tag{19.13}$$

This exact equation of motion is a suitable starting point for approximations. The possibility that V may depend on the time has been explicitly allowed for. Introducing the transition frequency by

$$\hbar\omega_{kn} = E_k - E_n \tag{19.14}$$

we obtain from (19.13),

$$\boxed{\langle k|\tilde{T}(t, t_0)|s\rangle = \delta_{ks} - \frac{i}{\hbar}\sum_n \int_{t_0}^{t} e^{i\omega_{kn}t'}\langle k|V(t')|n\rangle\langle n|\tilde{T}(t', t_0)|s\rangle\, dt'} \tag{19.15}$$

We note that the matrix elements of the time development operator defined here are simply related to the transition amplitudes defined in Eq. (14.51):

$$\langle k|T(t, t_0)|s\rangle = \langle \tilde{k}, t|\tilde{T}(t, t_0)|\tilde{s}, t_0\rangle = e^{-(i/\hbar)E_k t}\langle k|\tilde{T}(t, t_0)|s\rangle e^{(i/\hbar)E_s t_0} \tag{19.16}$$

The probabilities deduced from these two kinds of amplitudes are obviously the same.

Before proceeding, we finally recall from Section 14.6 a compact, though not generally useful, formula for the solution of the equation of motion (19.11):

$$\tilde{T}(t, t_0) = \mathrm{T}\exp\left[-\frac{i}{\hbar}\int_{t_0}^{t} \tilde{V}(t')\, dt'\right] \tag{19.17}$$

where T stands for the time ordering instruction (14.129).

The integral equations (19.15) are equivalent to a system of simultaneous linear homogeneous differential equations. Omitting for brevity the time arguments of the $\tilde{T}$ and V operators, we can write this system in matrix notation as

$$i\hbar \frac{d}{dt}\begin{pmatrix} \langle 1|\tilde{T}|1\rangle & \langle 1|\tilde{T}|2\rangle & \langle 1|\tilde{T}|3\rangle & . & . \\ \langle 2|\tilde{T}|1\rangle & \langle 2|\tilde{T}|2\rangle & \langle 2|\tilde{T}|3\rangle & . & . \\ \langle 3|\tilde{T}|1\rangle & \langle 3|\tilde{T}|2\rangle & \langle 3|\tilde{T}|3\rangle & . & . \\ . & . & . & . & . \\ . & . & . & . & . \end{pmatrix} =$$

$$\begin{pmatrix} \langle 1|V|1\rangle & \langle 1|V|2\rangle e^{i\omega_{12}t} & \langle 1|V|3\rangle e^{i\omega_{13}t} & . & . \\ \langle 2|V|1\rangle e^{-i\omega_{12}t} & \langle 2|V|2\rangle & \langle 2|V|3\rangle e^{i\omega_{23}t} & . & . \\ \langle 3|V|1\rangle e^{-i\omega_{13}t} & \langle 3|V|2\rangle e^{-i\omega_{23}t} & \langle 3|V|3\rangle & . & . \\ . & . & . & . & . \\ & & & . & . \end{pmatrix} \begin{pmatrix} \langle 1|\tilde{T}|1\rangle & \langle 1|\tilde{T}|2\rangle & \langle 1|\tilde{T}|3\rangle & . \\ \langle 2|\tilde{T}|1\rangle & \langle 2|\tilde{T}|2\rangle & \langle 2|\tilde{T}|3\rangle & . \\ \langle 3|\tilde{T}|1\rangle & \langle 3|\tilde{T}|2\rangle & \langle 3|\tilde{T}|3\rangle & . \\ . & . & . & . \\ . & & . & . \end{pmatrix} \tag{19.18}$$

It is in the solution of this complicated linear system that the perturbation approximation is invoked. The method takes its name, *time-dependent perturbation theory*, from the fact that it solves directly the time-dependent equations of motion rather than the time-independent Schrödinger equation. The perturbation V may or may not depend on the time.

2. ***The Perturbation Method.*** The strategy of time-dependent perturbation theory is to assume that V is small and proceed with iterating Eq. (19.11). First, $\tilde{T} = 1$ is substituted under the integral; the approximate $\tilde{T}$ so obtained is substituted again in the integrand, and so on. The result of this repeated iteration, already carried out earlier in Section 14.6 with a slightly different notation, appears as a power series in terms of $\tilde{V}$:

$$\tilde{T}(t, t_0) = 1 - \frac{i}{\hbar}\int_{t_0}^{t} \tilde{V}(t')\,dt' + \left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^{t} \tilde{V}(t')\,dt' \int_{t_0}^{t'} \tilde{V}(t'')\,dt'' + \ldots \tag{19.19}$$

Approximate solutions of the equation of motion are obtained if this power series is arbitrarily terminated.

Substitution of (19.19) into the integrand of (19.13) produces the desired perturbation expansion of the transition amplitude,

$$\langle k|\tilde{T}(t, t_0)|s\rangle = \delta_{ks} - \frac{i}{\hbar}\int_{t_0}^{t} e^{i\omega_{ks}t'}\langle k|V(t')|s\rangle\,dt' + \left(-\frac{i}{\hbar}\right)^2 \sum_n \int_{t_0}^{t} dt'\, e^{i\omega_{kn}t'}\langle k|V(t')|n\rangle \int_{t_0}^{t'} dt''\, e^{i\omega_{ns}t''}\langle n|V(t'')|s\rangle + \ldots \tag{19.20}$$

If $k \neq s$, the transition amplitude in first-order perturbation theory is given by the simple and much-used formula,

$$\boxed{\langle k|\tilde{T}(t, t_0)|s\rangle = e^{(i/\hbar)E_k t}\langle k|T(t, t_0)|s\rangle e^{-(i/\hbar)E_s t_0} = -\frac{i}{\hbar}\int_{t_0}^{t} \langle k|V(t')|s\rangle e^{i\omega_{ks}t'}\,dt'} \tag{19.21}$$

If the initial state of the system is an eigenstate of the unperturbed Hamiltonian H_0, corresponding to the unperturbed energy E_s, which is generally not an eigenstate of the perturbed Hamiltonian, the time development is not simply oscillatory and the

perturbation induces transitions between unperturbed energy eigenstates. How this kind of an initial state is adapted to the real physical conditions that prevail in an experiment will be shown later in several examples and applications. If the perturbation is sufficiently small during the period from t_0 to t, the probability for the system to be found in a different eigenstate of H_0 with energy E_k is

$$P_{k\leftarrow s}(t) = \left| -\frac{i}{\hbar} \int_{t_0}^{t} \langle k|V(t')|s\rangle e^{i\omega_{ks}t'}\, dt' \right|^2 \tag{19.22}$$

A necessary condition for this approximate result to be valid is that the time interval from t_0 to t be short enough for the transition probabilities to remain small compared to unity.

If the perturbation is transient and sufficiently weak in magnitude, the transition amplitudes may remain small and the first-order approximation valid throughout all times. In this case, the time development operator $\tilde{T}(t, t_0)$ is independent of t_0 and t, provided that t_0 is a time preceding the onset of the perturbation, and t is a time after the perturbation ceases. We may then with impunity take the limits $t_0 \to -\infty$ and $t \to +\infty$ and calculate the transition amplitude from one unperturbed energy eigenstate to another by the Fourier integral,

$$\langle k|\tilde{T}(+\infty, -\infty)|s\rangle = -\frac{i}{\hbar} \int_{-\infty}^{+\infty} \langle k|V(t)|s\rangle e^{i\omega_{ks}t}\, dt \tag{19.23}$$

The transition probability $P_{k\leftarrow s}(+\infty)$ is thus proportional to the square of the absolute value of the Fourier component of the perturbation matrix element $\langle k|V(t)|s\rangle$ evaluated at the transition frequency ω_{ks}.

Exercise 19.1. From the integral equation (19.15) and from (19.21), derive the corresponding linear differential equations for the transition amplitudes, and compare the perturbation version to the exact equations.

Exercise 19.2. Apply first-order perturbation theory to a forced linear harmonic oscillator which is initially in the ground state, and compare the transition probability with the exact result in Eq. (14.146). Calculate the energy transfer to the oscillator exactly and also in perturbation theory. Explain the agreement. Show that classical mechanics gives the same result for an oscillator that is initially at rest!

Exercise 19.3. Using the first-order formula (19.22) for the transitions from a state s to states $k \neq s$, calculate the probability that the system remains in state s. Show that the result agrees with $P_{s\leftarrow s}(t)$ calculated from $\langle s|\tilde{T}(t, t_0)|s\rangle$ only if the second-order terms are retained in the perturbation expansion of this diagonal matrix element.

If the approximation (19.21) is too drastic to be useful, it can be significantly improved by taking the interactions of a selected number of states more accurately into account. This procedure is analogous to the treatment of the nearly degenerate unperturbed state in Section 18.5. As in all perturbation theories, the chief criterion of success is the consistency of the solution with the simplifying assumptions made in arriving at the approximate equations. If a proper choice of *closely coupled states* is made, the perturbation amplitudes may turn out to be sufficiently accurate for times long enough to have physical relevance.

If, for example, r and s label two unperturbed energy eigenstates that interact strongly with each other but only weakly with all other states k, and if the initial state $|\Psi(t_0)\rangle$ is a linear superposition of $|r\rangle$ and $|s\rangle$, Eqs. (19.21) are replaced most conveniently by a set of first-order perturbation equations in differential form. They are simplest when the transition amplitudes are expressed in terms of matrix elements of the time development operator in the interaction picture. In the spirit of perturbation theory, we neglect on the right-hand side any transition amplitudes that link the initial state $|\Psi(t_0)\rangle$ to any unperturbed energy eigenstates other than $|r\rangle$ and $|s\rangle$, and write:

$$\begin{aligned}
i\hbar \frac{d}{dt}\langle k|\tilde{T}(t, t_0)|\tilde{\Psi}(t_0)\rangle &= \langle k|V(t)|r\rangle e^{i\omega_{kr}t}\langle r|\tilde{T}(t, t_0)|\tilde{\Psi}(t_0)\rangle \\
&\quad + \langle k|V(t)|s\rangle e^{i\omega_{ks}t}\langle s|\tilde{T}(t, t_0)|\tilde{\Psi}(t_0)\rangle \quad (k \neq r, s) \\
i\hbar \frac{d}{dt}\langle r|\tilde{T}(t, t_0)|\tilde{\Psi}(t_0)\rangle &= \langle r|V(t)|r\rangle\langle r|\tilde{T}(t, t_0)|\tilde{\Psi}(t_0)\rangle \\
&\quad + \langle r|V(t)|s\rangle e^{i\omega_{rs}t}\langle s|\tilde{T}(t, t_0)|\tilde{\Psi}(t_0)\rangle \\
i\hbar \frac{d}{dt}\langle s|\tilde{T}(t, t_0)|\tilde{\Psi}(t_0)\rangle &= \langle s|V(t)|r\rangle e^{i\omega_{sr}t}\langle r|\tilde{T}(t, t_0)|\tilde{\Psi}(t_0)\rangle \\
&\quad + \langle s|V(t)|s\rangle\langle s|\tilde{T}(t, t_0)|\tilde{\Psi}(t_0)\rangle
\end{aligned} \tag{19.24}$$

An example of the use of these equations will be given in Section 19.5.

3. ***Coulomb Excitation and Sum Rules.*** As an illustration of first-order time-dependent perturbation theory, we now study the effect of a transient Coulomb field on a compact bound system of charged particles, which we may call the *target*.

As a concrete example, consider the excitation of a *nucleus* by a particle with charge Z_1e which sweeps by. When charged particles, e.g., alpha particles of moderate energy, bombard a nucleus of atomic number Z_2, the particles are deflected by the mutual Coulomb repulsion, and the familiar Rutherford scattering takes place (Section 13.8). In the process, the nucleus sees a rapidly varying electric field that can cause transitions from the ground state to some excited nuclear state. When this happens, we speak of *Coulomb excitation* in contrast to other inelastic processes that occur if the projectile is hadronic and fast enough to allow penetration of the Coulomb barrier, so that strong interactions through forces of origin other than electromagnetic can occur.

Here we will consider another example: the excitation and ionization of an *atom* by the passage of a charged particle. This is the dominant mechanism by which charged particles are slowed down in matter.

In lowest approximation, it is frequently permissible to treat the projectile as a classical particle moving on a trajectory with a definite initial momentum and impact parameter in the field of the target atom. The perturbation interaction is simply of the form

$$V = \pm Z_1e^2 \sum_{i=1}^{Z_2} \frac{1}{|\mathbf{r}_i - \mathbf{R}(t)|} \tag{19.25}$$

where $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_{Z_2}$ denote the position operators for the Z_2 electrons in the target, and $\mathbf{R}(t)$ is the position vector of the projectile with charge $\mp Z_1e$. It is convenient to choose the atomic nucleus as the coordinate origin.

If the projectile never enters the target, an electric multipole expansion of the form

$$\frac{1}{|\mathbf{r}_i - \mathbf{R}|} = \sum_{\ell=0}^{\infty} \frac{r_i^\ell}{R^{\ell+1}} P_\ell(\hat{\mathbf{r}}_i \cdot \hat{\mathbf{R}}) \tag{19.26}$$

is appropriate. Substituting this into (19.23), we obtain the transition amplitude,

$$\langle k|\tilde{T}(+\infty, -\infty)|s\rangle = \mp\frac{iZ_1e^2}{\hbar} \sum_{\ell=0}^{\infty} \int_{-\infty}^{+\infty} e^{i\omega_{ks}t} \frac{\langle k| \sum_{i=1}^{Z_2} r_i^\ell P_\ell(\hat{\mathbf{r}}_i \cdot \hat{\mathbf{R}}(t))|s\rangle}{[R(t)]^{\ell+1}} dt \tag{19.27}$$

If the distance between the projectile and the target atom is always large compared with the size of the atom, the multipole expansion converges rapidly, and only the first nonvanishing term in the power series in ℓ in (19.27) need be retained. For a given initial state s and a final state k, there is always a smallest value of ℓ, for which the matrix element $\langle k| \sum_{i=1}^{Z_2} r_i^\ell P_\ell(\hat{\mathbf{r}}_i \cdot \hat{\mathbf{R}}(t))|s\rangle$ will not vanish. Since $P_0 = 1$, and $|k\rangle$ and $|s\rangle$ are orthogonal states, the matrix elements for $\ell = 0$ (monopole transitions) are zero. The next term, $\ell = 1$, corresponds to electric dipole or $E1$ transitions, $\ell = 2$ to quadrupole or $E2$ transitions, and generally we speak of 2^ℓ-pole or $E\ell$ transitions. Higher multipole transitions are important in nuclear Coulomb excitation, but in the atomic case *electric dipole* ($\ell = 1$) transitions strongly dominate, and we limit the discussion to them.

If the plane of the orbit is taken to be the xz-coordinate plane, we have ($\varphi = 0$)

$$\dot{r}_i P_1(\hat{\mathbf{r}}_i \cdot \hat{\mathbf{R}}(t)) = \mathbf{r}_i \cdot \hat{\mathbf{R}}(t) = x_i \sin\theta(t) + z_i \cos\theta(t) \tag{19.28}$$

where $\theta(t)$ is the instantaneous polar angle of the projectile (see Figure 19.1). If we assume that the projectile moves in a central field originating from the target, conservation of angular momentum may be invoked in the classical description of this motion, and we have

$$[R(t)]^2 \frac{d\theta}{dt} = -bv \tag{19.29}$$

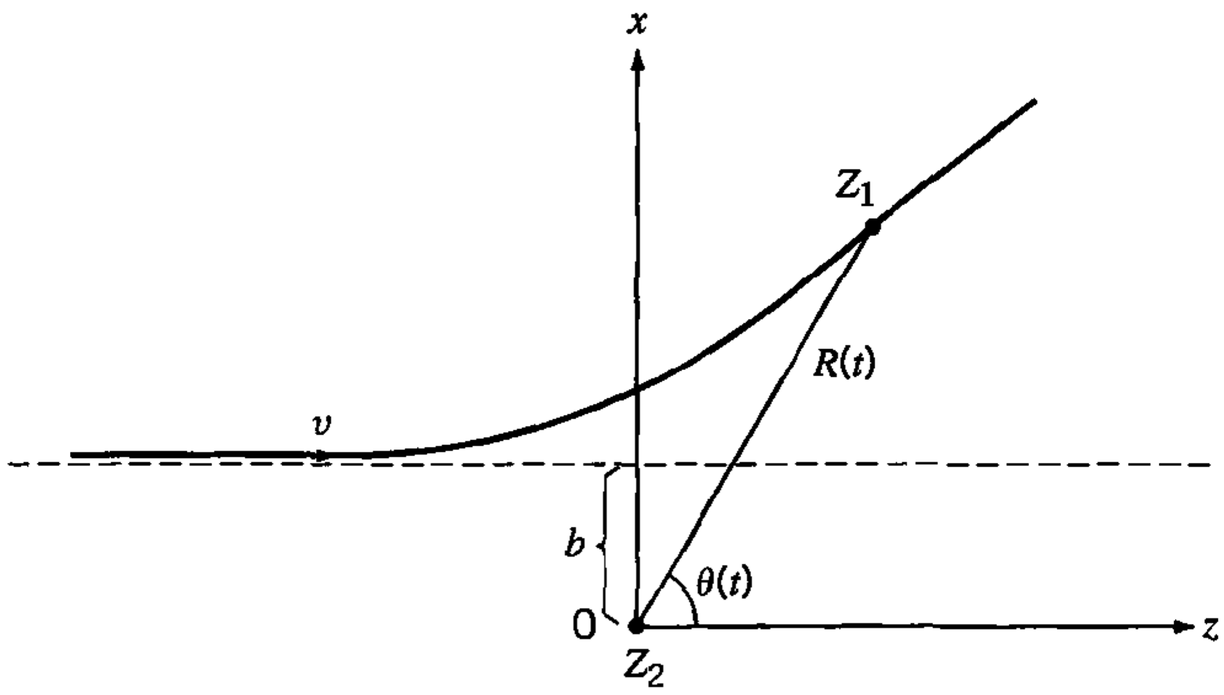


Figure 19.1. The Coulomb field of a charged particle Z_1 passing near a charged target Z_2 causes internal excitation. The impact parameter is b, and the coordinates $\theta(t)$ and $R(t)$ determine the instantaneous position of the projectile.

where v is the initial velocity of the incident particle and b its impact parameter. Substituting (19.28) and (19.29) into (19.27), we obtain for the dipole transition amplitude

$$\langle k|\tilde{T}(+\infty, -\infty)|s\rangle = \mp\frac{iZ_1e^2}{\hbar vb}\int_{\pi}^{\theta(b)} e^{i\omega_{ks}t}(\langle k|x|s\rangle \sin\theta + \langle k|z|s\rangle \cos\theta)\, d\theta \quad (19.30)$$

where the components of the *dipole matrix elements* of the target atom are defined by

$$\langle k|\mathbf{r}|s\rangle = \langle k|\sum_{i=1}^{Z_2}\mathbf{r}_i|s\rangle \quad (19.31)$$

The integration in (19.29) extends to the final scattering angle $\theta(b)$, which corresponds to the prescribed impact parameter, b. Although the integrand in (19.29) looks simple, it should be remembered that t is usually a complicated function of θ, and the integral can be evaluated only if the classical motion of the projectile is known.

For some purposes, it is a good approximation to assume that the exponent $\omega_{ks}t$ in the integrand of (19.30) is negligible. This approximation is valid if the collision time, during which the projectile is close enough to the atom to be effective in exciting the latter, is short compared with $1/\omega_{ks}$, which is a typical period associated with the proper oscillations of the atomic electrons. In the limit $\omega_{ks}t \to 0$, (19.30) can be integrated easily and yields

$$\langle k|\tilde{T}(+\infty, -\infty)|s\rangle = \mp\frac{iZ_1e^2}{\hbar vb}\{\langle k|x|s\rangle[\cos\theta(b) + 1] - \langle k|z|s\rangle \sin\theta(b)\} \quad (19.32)$$

Exercise 19.4. If the transition amplitude is given by the approximation (19.32), show by using closure that the total transition probability from an initial spherically symmetric atomic state s is

$$\sum_{k\neq s} P_{k\leftarrow s}(+\infty) = \frac{4Z_1^2e^4}{3\hbar^2v^2b^2}\langle s|r^2|s\rangle \cos^2\theta(b)/2 \quad (19.33)$$

Make an estimate of the magnitude of this transition probability for 1 MeV protons incident on hydrogen atoms at an impact parameter b comparable to the Bohr radius.

The average energy transfer ΔE in inelastic collisions is defined as

$$\Delta E = \sum_k (E_k - E_s)P_{s\leftarrow k}(+\infty) = \sum_k (E_k - E_s)|\langle k|\tilde{T}(+\infty, -\infty)|s\rangle|^2 \quad (19.34)$$

where the sum can be taken over *all* states of the target atom (since the term $k = s$ contributes nothing). When (19.32) is substituted in (19.34), we are typically faced with the need to evaluate sums involving dipole matrix elements, such as

$$S_{ij} = \frac{m}{\hbar^2}\sum_k (E_k - E_s)[\langle s|x_i|k\rangle\langle k|x_j|s\rangle + \langle s|x_j|k\rangle\langle k|x_i|s\rangle] \quad (19.35)$$

This can be written as

$$S_{ij} = \frac{m}{\hbar^2}\langle s|x_iH_0x_j + x_jH_0x_i - H_0x_ix_j - x_jx_iH_0|s\rangle = \frac{m}{\hbar^2}\langle s|[[x_i, H_0], x_j]|s\rangle$$

If the unperturbed atomic Hamiltonian H_0 contains the momentum only in the kinetic energy operator, we have

$$[x_i, H_0] = \frac{i\hbar}{m} p_i \quad \text{and} \quad [[x_i, H_0], x_j] = \frac{\hbar^2}{m} \delta_{ij}$$

hence, the useful identity[1]

$$S_{ij} = \delta_{ij} \tag{19.36}$$

Applying this relation to a component of the dipole matrix elements of a one-electron atom ($Z_2 = 1$), we thus obtain

$$\frac{2m}{\hbar^2} \sum_k (E_k - E_s)|\langle k|x|s\rangle|^2 = \frac{2m}{\hbar} \sum_k \omega_{ks}|\langle k|x|s\rangle|^2 = 1 \tag{19.37}$$

For a multi-electron atom with dipole moment (19.31), this identity can be generalized to the dipole sum rule,

$$\boxed{\frac{2m}{\hbar} \sum_k \omega_{ks}|\langle k|x|s\rangle|^2 = Z_2} \tag{19.38}$$

The dimensionless quantity

$$\boxed{f_{ks} = \frac{2m}{\hbar} \omega_{ks}|\langle k|x|s\rangle|^2} \tag{19.39}$$

is a useful number in spectroscopy. If the transition $s \rightarrow k$ were between the two lowest energy levels of a harmonic oscillator, with level spacing $\hbar\omega_{ks}$, we would find $f_{ks} = \pm 1$, the sign depending on which of the two levels is higher. Therefore, f_{ks} is known as the *oscillator strength* for the transition between s and k with frequency ω_{ks}, and (19.39) is the *Thomas-Reiche-Kuhn sum rule* for the oscillator strengths:

$$\sum_k f_{ks} = Z_2 \tag{19.40}$$

Exercise 19.5. For a hydrogen atom that is initially in the $1S$ ground state, compute the oscillator strengths to the lowest three excited states. What fraction of the total do these oscillator strengths add up to? (Use the z coordinate, rather than x, to simplify the calculations.)

The approximate average energy transferred by a fast charged particle in a distant (large-impact-parameter) collision to an atom is, according to (19.32), (19.34), and (19.38), given by

$$\Delta E = \frac{2Z_1^2 Z_2\, e^4}{mv^2 b^2} \cos^2 \frac{\theta(b)}{2}$$

which does not contain Planck's constant and is essentially the same result as a classical calculation would give. The concepts (oscillator strength, dipole sum rules)

[1]For more information on atomic sum rules see Bethe and Salpeter (1957), Sections 61–63, and Bethe and Jackiw (1968), Chapter 11.

introduced in this section will find useful application when we turn our attention to the interaction of radiation with matter.

4. ***The Atom in a Radiation Field.*** As a further illustration of the use of (19.23), it is instructive to calculate the absorption cross section of an atom for light. The electron spin will be neglected, and for simplicity it will be assumed that only one atomic electron is involved in the interaction with the incident radiation and that the nucleus is infinitely heavy. The results can be generalized to other systems (molecules, nuclei) that can absorb radiation.

The Hamiltonian of the atomic electron (charge $-e$) in an electromagnetic field is given by (4.93) as

$$H = \frac{1}{2m}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 - e\phi \tag{19.41}$$

In the coordinate representation, this becomes the operator

$$\begin{aligned} H = &-\frac{\hbar^2}{2m}\nabla^2 - e\phi(\mathbf{r}, t) + \frac{e}{mc}\mathbf{A}(\mathbf{r}, t) \cdot \frac{\hbar}{i}\nabla \\ &- \frac{ie\hbar}{2mc}[\nabla \cdot \mathbf{A}(\mathbf{r}, t)] + \frac{e^2}{2mc^2}[\mathbf{A}(\mathbf{r}, t)]^2 \end{aligned} \tag{19.42}$$

We will apply this Hamiltonian only under conditions that allow the $\mathbf{A}^2$ term to be neglected.

It is convenient to separate the electromagnetic field that represents the incident light wave from all other contributions to the Hamiltonian, whether or not their sources are electromagnetic, such as the Coulomb interaction within the atom. Only the effect of the pure radiation will be treated as a perturbation, and all other forces that act on the electron are incorporated into the V_0 term of the unperturbed Hamiltonian,

$$H_0 = \frac{\mathbf{p}^2}{2m} + V_0 \tag{19.43}$$

According to Maxwell's theory, by choosing a suitable gauge the pure radiation field that perturbs the atom can be described in terms of a vector potential $\mathbf{A}$ alone. Since this field has no sources in the vicinity of the atom, it may in the so-called *Coulomb gauge* be derived from potentials which in Cartesian coordinates satisfy the equations

$$\nabla^2\mathbf{A} - \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} = 0, \qquad \phi = 0, \qquad \nabla \cdot \mathbf{A} = 0$$

With this choice, the complete Hamiltonian simplifies to

$$H = \frac{\mathbf{p}^2}{2m} + V_0 + \frac{e}{mc}\mathbf{A} \cdot \mathbf{p} = H_0 + \frac{e}{mc}\mathbf{A} \cdot \mathbf{p} \tag{19.44}$$

The last term will be regarded as an external perturbation dependent on $\mathbf{r}$ and t. If V_0 represents the potential energy of a harmonic oscillator and if the spatial variation of $\mathbf{A}$ can be neglected, (19.44) is the Hamiltonian of a forced harmonic oscillator, which was treated in detail and without approximations in Section 14.6.

Exercise 19.6. Prove that the perturbation term in (19.44) is Hermitian.

It is convenient, and corresponds to actual experimental conditions, to suppose that the incident radiation can be described as a very broad plane wave packet moving in a given direction. Any such wave can be written as a superposition of harmonic plane waves,

$$\mathbf{A}(\mathbf{r}, t) = \int_{-\infty}^{+\infty} \mathbf{A}(\omega) \exp\left[-i\omega\left(t - \frac{\hat{\mathbf{n}} \cdot \mathbf{r}}{c}\right)\right] d\omega \tag{19.45}$$

where $\hat{\mathbf{n}}$ is the unit vector pointing in the direction of propagation. The vector potential must be a real quantity, which requires that

$$\mathbf{A}^*(\omega) = \mathbf{A}(-\omega) \tag{19.46}$$

On the other hand, the solenoidal character of the vector potential ($\nabla \cdot \mathbf{A} = 0$) implies that the wave is transverse:

$$\hat{\mathbf{n}} \cdot \mathbf{A}(\omega) = 0 \tag{19.47}$$

The perturbation,

$$V = \frac{e}{mc} \int_{-\infty}^{+\infty} \exp\left[-i\omega\left(t - \frac{\hat{\mathbf{n}} \cdot \mathbf{r}}{c}\right)\right] \mathbf{A}(\omega) \cdot \mathbf{p}\, d\omega \tag{19.48}$$

is now substituted into (19.23) to give the transition amplitude,

$$\begin{aligned}
&\langle k|\tilde{T}(+\infty, -\infty)|s\rangle \\
&= -\frac{ie}{\hbar mc} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \langle k|\exp\left(i\frac{\omega}{c}\hat{\mathbf{n}} \cdot \mathbf{r}\right)\mathbf{p}|s\rangle \cdot \mathbf{A}(\omega) \exp[i(\omega_{ks} - \omega)t]\, dt\, d\omega \\
&= -\frac{2\pi ie}{\hbar mc} \int_{-\infty}^{+\infty} \langle k|\exp\left(i\frac{\omega}{c}\hat{\mathbf{n}} \cdot \mathbf{r}\right)\mathbf{p}|s\rangle \cdot \mathbf{A}(\omega)\delta(\omega - \omega_{ks})\, d\omega \\
&= -\frac{2\pi ie}{\hbar mc} \langle k|\exp\left(i\frac{\omega_{ks}}{c}\hat{\mathbf{n}} \cdot \mathbf{r}\right)\mathbf{p}|s\rangle \cdot \mathbf{A}(\omega_{ks})
\end{aligned} \tag{19.49}$$

The last expression shows that, to this order, the only Fourier component of the incident radiation that is effective in the absorption processes leading to the final state k corresponds to the frequency

$$\omega_{ks} = \frac{E_k - E_s}{\hbar} \tag{19.50}$$

in agreement with the frequency condition originally postulated by Bohr. The energy difference $\hbar\omega_{ks}$ is given up by the light pulse to the absorbing system, but the change that the radiation field undergoes in this energy transfer can be described only if the electromagnetic field, instead of being treated as an external prescribed driving force, is made part of the dynamical system and thus is itself amenable to quantization. The quantum theory of the radiation field—historically, the platform from which Planck and Einstein launched quantum theory in the first place—will be developed in Chapter 23. It should always be remembered that (19.49) holds only if the light wave is sufficiently weak so that first-order perturbation theory is valid for the entire duration of the pulse and the transition probability is much less than unity. Anticipating the language of the quantized radiation field, we may interpret (19.49)

as the probability amplitude for annihilation and absorption of one *photon* of energy $\hbar\omega_{ks}$.

The light pulse also induces inverse transitions if the atom is initially in state k. The transition amplitude to the state s of lower energy is given by

$$\langle s|\tilde{T}(+\infty, -\infty)|k\rangle = -\frac{i}{\hbar}\int_{-\infty}^{+\infty}\langle s|V(t)|k\rangle e^{i\omega_{sk}}\,dt \tag{19.51}$$

Since V is a Hermitian perturbation operator (see Exercise 19.6) and $\omega_{sk} = -\omega_{ks}$, it follows that in first-order perturbation theory

$$\langle s|\tilde{T}(+\infty, -\infty)|k\rangle = -\langle k|\tilde{T}(+\infty, -\infty)|s\rangle^* \tag{19.52}$$

showing that in this approximation the two transition probabilities are equal. This property is known as the condition of *detailed balancing*. Since $E_k > E_s$, the excess energy $\hbar\omega_{ks}$ is transferred to the radiation field. A proper quantum treatment of the radiation field in Chapter 23 will confirm that this energy appears as light of frequency ω_{ks}. The transition from k to s, induced by the applied radiation pulse, is known as *stimulated emission*. If the radiating system is near thermal equilibrium, stimulated emission between two energy levels is usually much less intense than absorption despite the equality of the transition probabilities, because the initial population of the upper level is generally smaller. In masers and lasers, however, a population inversion is achieved, and stimulated emission predominates. In the language of the quantized radiation field, stimulated emission corresponds to the creation of a photon with energy $\hbar\omega_{ks}$.

For actual calculations of transition probabilities using (19.49), we may assume the pulse (19.45) to be linearly polarized, which implies that all frequency components $\mathbf{A}(\omega)$ have the same direction independent of ω. No loss of generality is involved in this assumption, if the cross section for absorption is calculated, since in an actual experiment a large number of incoherent pulses of the form (19.45) are directed at the system. Incoherence here simply means that the phases of the pulses are arbitrary and independent of each other. According to the wave equation, each Fourier component is elliptically polarized. Each elliptical oscillation can be decomposed into two linear ones directed along two fixed perpendicular axes, both of which are orthogonal to the propagation vector $\hat{\mathbf{n}}$. Because of the uniform randomness of the phases, the interference terms that arise between the two linear oscillations upon squaring the amplitude (19.49) will cancel when an average over many pulses is performed. Hence, the total absorption probability is the sum of two expressions of the form $|\langle k|\tilde{T}(+\infty, -\infty)|s\rangle|^2$, obtained by squaring (19.49), one for each direction of polarization.

If we adopt a definite direction of linear polarization $\hat{\mathbf{e}}$ and write

$$\mathbf{A}(\omega) = A(\omega)\hat{\mathbf{e}}$$

the transition probability becomes

$$\boxed{\begin{aligned} P_{k\leftarrow s} &\equiv |\langle k|\tilde{T}(+\infty, -\infty)|s\rangle|^2 \\ &= \frac{4\pi^2e^2}{\hbar^2m^2c^2}\,|A(\omega_{ks})|^2|\langle k|\exp\left(i\frac{\omega_{ks}}{c}\,\hat{\mathbf{n}}\cdot\mathbf{r}\right)\mathbf{p}\cdot\hat{\mathbf{e}}|s\rangle|^2 \end{aligned}} \tag{19.53}$$

This transition probability can be interpreted as the probability for absorption (or emission) or a photon, or the mean photon number absorbed by the atom from the pulse.

This result can also be related to the intensity of the incident pulse by comparing it with the Poynting vector,

$$\begin{aligned}\mathbf{N} = c\frac{\mathbf{E}\times\mathbf{H}}{4\pi} &= -\frac{1}{4\pi}\frac{\partial\mathbf{A}}{\partial t}\times(\nabla\times\mathbf{A}) \\ &= -\frac{1}{4\pi c}\int_{-\infty}^{+\infty}\omega A(\omega)\,\hat{\mathbf{e}}\exp\left[-i\omega\left(t-\frac{\hat{\mathbf{n}}\cdot\mathbf{r}}{c}\right)\right]d\omega \\ &\times\int_{-\infty}^{+\infty}\omega' A(\omega')(\hat{\mathbf{n}}\times\hat{\mathbf{e}})\exp\left[-i\omega'\left(t-\frac{\hat{\mathbf{n}}\cdot\mathbf{r}}{c}\right)\right]d\omega'\end{aligned} \tag{19.54}$$

The triple product, $\hat{\mathbf{e}}\times(\hat{\mathbf{n}}\times\hat{\mathbf{e}}) = \hat{\mathbf{n}} - \hat{\mathbf{e}}(\hat{\mathbf{e}}\cdot\hat{\mathbf{n}})$, simplifies because the wave is transverse, whence $\hat{\mathbf{e}}\cdot\hat{\mathbf{n}} = 0$, giving the equation

$$\mathbf{N} = -\frac{1}{4\pi c}\hat{\mathbf{n}}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}d\omega\,d\omega'\;\omega\omega'\,A(\omega)A(\omega')\exp\left[-i(\omega+\omega')\left(t-\frac{\hat{\mathbf{n}}\cdot\mathbf{r}}{c}\right)\right] \tag{19.55}$$

It follows that the energy flows in the direction of propagation and that the total energy which the pulse caries through a unit area placed perpendicular to the direction of propagation is given by

$$\begin{aligned}\int_{-\infty}^{+\infty}\mathbf{N}\cdot\hat{\mathbf{n}}\,dt &= \frac{1}{2c}\int_{-\infty}^{+\infty}d\omega\;\omega^2|A(\omega)|^2 \\ &= \frac{1}{c}\int_0^{\infty}d\omega\;\omega^2|A(\omega)|^2 = \int_0^{\infty}N(\omega)\,d\omega\end{aligned} \tag{19.56}$$

Hence, the energy carried through a unit area per unit frequency interval is

$$N(\omega) = \frac{\omega^2}{c}|A(\omega)|^2 \tag{19.57}$$

Combining this result with (19.53), we obtain

$$\boxed{P_{k\leftarrow s} = \frac{4\pi^2e^2}{\hbar^2m^2c\omega_{ks}^2}|\langle k|\exp\left(i\frac{\omega_{ks}}{c}\hat{\mathbf{n}}\cdot\mathbf{r}\right)\mathbf{p}\cdot\hat{\mathbf{e}}|s\rangle|^2N(\omega_{ks})} \tag{19.58}$$

Energy transfer from the incident radiation pulse to the atom takes place only at those frequencies that the system can absorb in accordance with the Bohr frequency condition (19.50). The average energy loss of the pulse at the frequency ω_{ks} is

$$\hbar\omega_{ks}\,P_{k\leftarrow s} = \frac{4\pi^2\alpha}{m^2\omega_{ks}}|\langle k|\exp\left(i\frac{\omega_{ks}}{c}\hat{\mathbf{n}}\cdot\mathbf{r}\right)\mathbf{p}\cdot\hat{\mathbf{e}}|s\rangle|^2N(\omega_{ks}) \tag{19.59}$$

where α denotes the fine structure constant, $\alpha = e^2/\hbar c \approx 1/137$.

It is customary to base the definition of an *absorption cross section* for linearly polarized electromagnetic radiation of a certain frequency on (19.59). [Equivalently, we can use (19.58) and think of the absorption cross section as measuring the atom's propensity for absorbing photons.] As explained in Section 13.1, the absorption cross section is a fictitious area which, when placed perpendicular to the incident pulse, is exposed to and traversed by an amount of radiation energy (or a mean number of

photons) equal to that absorbed by the atom or nucleus. This notion of the cross section is useful only if it is independent of the detailed shape of the incident pulse. Besides the nature and initial condition of the absorbing system, it should depend only on the frequency, the direction of incidence, and the state of polarization of the incident beam. Under the usual experimental conditions, these goals can be achieved and a cross section $\sigma(\omega)$ defined, such that the energy absorbed in the frequency interval $d\omega$ is

$$\sigma(\omega)N(\omega)\,d\omega \tag{19.60}$$

Absorption cross sections corresponding to two diametrically opposite situations will now be evaluated:

1. The initial and final states are discrete, and the absorption spectrum consists of sharp lines (Section 19.5).

2. The initial state is again discrete, but the final state is in the continuum. This is the photoelectric effect, in which an atomic electron is released and $\sigma(\omega)$ is a slowly varying function of the frequency (Section 19.6).

5. ***The Absorption Cross Section.*** If the states of s and k are discrete, radiation can be absorbed only if its frequency lies very close to ω_{ks}. Hence, $\sigma(\omega)$ is a strongly peaked function of ω. The total energy absorbed is obtained by integrating the expression (19.60), and this quantity can then be equated to the right-hand side of (19.59):

$$\int_{\Delta\omega} \sigma(\omega)N(\omega)\,d\omega = \frac{4\pi^2\alpha}{m^2\omega_{ks}}\left|\langle k|\exp\left(i\,\frac{\omega_{ks}}{c}\,\hat{\mathbf{n}}\cdot\mathbf{r}\right)\mathbf{p}\cdot\hat{\mathbf{e}}|s\rangle\right|^2 N(\omega_{ks}) \tag{19.61}$$

Here the integral on the left is to be extended over an interval $\Delta\omega$ that completely contains the spectral line (but no other absorption frequencies). Since $\sigma(\omega)$ has a steep and narrow maximum at $\omega = \omega_{ks}$, a reasonable definition of the cross section can be obtained if $N(\omega)$ varies relatively slowly. If we restrict ourselves to pulses broad in frequency compared with the line, the left-hand side of (19.61) can be written as

$$\int_{\Delta\omega} \sigma(\omega)N(\omega)\,d\omega \approx N(\omega_{ks})\int_{\Delta\omega} \sigma(\omega)\,d\omega$$

and an integrated absorption cross section can be obtained:

$$\boxed{\int_{\Delta\omega} \sigma(\omega)\,d\omega = \frac{4\pi^2\alpha}{m^2\omega_{ks}}\left|\langle k|\exp\left(i\,\frac{\omega_{ks}}{c}\,\hat{\mathbf{n}}\cdot\mathbf{r}\right)\mathbf{p}\cdot\hat{\mathbf{e}}|s\rangle\right|^2} \tag{19.62}$$

Despite its somewhat misleading name, the integrated absorption cross section is not a geometric area; it is the product of an area and a frequency and is measured in the corresponding units.

Further progress in the evaluation of this general cross-section formula depends on specific assumptions about the structure of the unperturbed Hamiltonian H_0 and its eigenstates s and k. The coordinate origin is conveniently chosen at the center of mass of the atom. Generally, it is possible to make an important approximation based on the fact that the wavelength of the incident light, $\lambda = 2\pi c/\omega_{ks}$, is large compared with the linear dimensions of the absorbing system. In this case, the exponential

operator in the matrix element of (19.62) is advantageously expanded in a power series

$$\exp\left(i\frac{\omega_{ks}}{c}\,\hat{\mathbf{n}}\cdot\mathbf{r}\right) = 1 + i\frac{\omega_{ks}}{c}\,\hat{\mathbf{n}}\cdot\mathbf{r} + \ldots \tag{19.63}$$

and only the first nonvanishing term in the corresponding series of matrix elements,

$$\langle k|\exp\left(i\frac{\omega_{ks}}{c}\,\hat{\mathbf{n}}\cdot\mathbf{r}\right)\mathbf{p}\cdot\hat{\mathbf{e}}|s\rangle = \langle k|\mathbf{p}\cdot\hat{\mathbf{e}}|s\rangle + \langle k|i\frac{\omega_{ks}}{c}\,\hat{\mathbf{n}}\cdot\mathbf{r}\,\mathbf{p}\cdot\hat{\mathbf{e}}|s\rangle + \ldots \tag{19.64}$$

is retained. The ratio of two successive nonvanishing terms in this series can be estimated by assuming that r is effectively of the order of magnitude of the radius of the system. For an atom of radius a/Z (a: Bohr radius) the expansion parameter is equal to

$$\frac{\omega a}{cZ} \cong \frac{Ze^2}{\hbar c} = Z\alpha \approx \frac{Z}{137}$$

and thus the series (19.64) converges rapidly for small Z.

Selection rules for the wave functions of the unperturbed states determine which term in the expansion (19.64) is the first nonvanishing matrix element. If the first term in the expansion, $\langle k|\mathbf{p}\cdot\hat{\mathbf{e}}|s\rangle$, is different from zero, the transition is said to be of the *electric dipole* or $E1$ type. Since in this approximation we effectively replace $\exp(i\,\omega\hat{\mathbf{n}}\cdot\mathbf{r}/c)$ by unity, the electric dipole approximation also implies a complete neglect of retardation across the atom. Owing to the inherent weakness of higher multipolarity ("forbidden" transitions), electric dipole (or "allowed") transitions are the most important ones in atoms. In nuclei the ratio $r\omega/c$ is not quite so small, since $\hbar\omega$ is of the order of $\hbar^2/M_nR^2$, where M_n is the nucleon mass and R is the nuclear radius. Hence, the expansion parameter is

$$\frac{\omega}{c}R \cong \frac{\hbar}{M_ncR} = \frac{\text{Compton wavelength of nucleon}}{\text{Nuclear radius}}$$

which for light nuclei can be as large as 1/10. The powerful detection techniques of nuclear physics make it easy to observe transitions of higher multipolarity. When such transitions are considered, it is no longer permissible to neglect the interaction of the electromagnetic field with the spin magnetic moment.

In electric dipole transitions, the matrix element $\langle k|\mathbf{p}|s\rangle$ is the decisive quantity that must be evaluated. This can be related to the matrix element of the position operator $\mathbf{r}$, if the unperturbed Hamiltonian is of the form $H_0 = \mathbf{p}^2/2m + V_0$, and if V_0 commutes with $\mathbf{r}$. Under these conditions

$$[\mathbf{r}, H_0] = \mathbf{r}H_0 - H_0\mathbf{r} = i\hbar\frac{\mathbf{p}}{m} \tag{19.65}$$

Hence, by taking the matrix element of both sides of this equation,

$$\langle k|\mathbf{p}|s\rangle = \frac{m}{i\hbar}(E_s^{(0)} - E_k^{(0)})\langle k|\mathbf{r}|s\rangle = im\omega_{ks}\langle k|\mathbf{r}|s\rangle \tag{19.66}$$

Making the dipole approximation and substituting (19.66) into (19.62), we obtain the integrated absorption cross section,

$$\boxed{\int_{\Delta\omega}\sigma(\omega)\,d\omega = 4\pi^2\alpha\omega_{ks}|\langle k|\mathbf{r}\cdot\hat{\mathbf{e}}|s\rangle|^2} \tag{19.67}$$

For many purposes, it is useful to rewrite the transition matrix element as follows:

$$\langle k|\mathbf{r}\cdot\hat{\mathbf{e}}|s\rangle = \frac{4\pi}{3}\sum_{q=-1}^{+1}(-1)^q\langle k|rY_1^q(\hat{\mathbf{r}})|s\rangle Y_1^{-q}(\hat{\mathbf{e}}) \tag{19.68}$$

If the absorbing system is a one-electron atom, and if the fine structure of the atomic energy levels may be neglected, the quantum numbers $n\ell m$ and $n'\ell'm'$ characterize the initial and final states. Formula (17.74) may then be applied to carry out the integration over the angle coordinates in the evaluation of the matrix element, giving us

$$\langle n'\ell'm'|\mathbf{r}\cdot\hat{\mathbf{e}}|n\ell m\rangle = \sqrt{\frac{4\pi}{3}\frac{2\ell+1}{2\ell'+1}}\sum_{q=-1}^{+1}\langle\ell 1mq|\ell 1\ell'm'\rangle\langle\ell 100|\ell 1\ell' 0\rangle[Y_1^q(\hat{\mathbf{e}})]^* R_{n\ell}^{n'\ell'} \tag{19.69}$$

where

$$R_{n\ell}^{n'\ell'} = \int_0^\infty R_{n'\ell'}(r)R_{n\ell}(r)r^3\,dr \tag{19.70}$$

and $R_{n\ell}(r)$ is the real normalized radial wave function. The electric dipole selection rules

$$\Delta\ell = \ell' - \ell = \pm 1, \qquad \Delta m = m' - m = 0, \pm 1$$

are an immediate consequence of (19.69). The values of the relevant Clebsch-Gordan (C-G) coefficients are

$$\langle\ell 100|\ell 1, \ell+1, 0\rangle = \sqrt{\frac{\ell+1}{2\ell+1}}, \quad \langle\ell 100|\ell 1, \ell-1, 0\rangle = -\sqrt{\frac{\ell}{2\ell+1}} \tag{19.71}$$

Using this information, we may write

$$\begin{aligned}&|\langle n'\ell'm'|\mathbf{r}\cdot\hat{\mathbf{e}}|n\ell m\rangle|^2\\ &\quad = \frac{4\pi}{3}\frac{2\ell+1}{2\ell'+1}\langle\ell 1m, m'-m|\ell 1\ell'm'\rangle^2\langle\ell 100|\ell 1\ell' 0\rangle^2|Y_1^{m'-m}(\hat{\mathbf{e}})|^2(R_{n\ell}^{n'\ell'})^2\end{aligned} \tag{19.72}$$

for the last factor in (19.69).

We may also calculate the average rate of absorption for a system (atom or molecule) that is exposed to a continuous frequency distribution of incoherent electromagnetic radiation from all directions and with random polarization. These conditions may, for example, represent the absorption of blackbody radiation. Usually, only electric dipole transitions are important, and we restrict our formulas to these, although forbidden atomic lines are observed in stellar spectra.

Applying the dipole approximation to (19.61), we find that the energy absorbed from a single polarized unidirectional pulse is

$$N(\omega_{ks})\int_{\Delta\omega}\sigma(\omega)\,d\omega = 4\pi^2\alpha\omega_{ks}|\langle k|\mathbf{r}\cdot\hat{\mathbf{e}}_i|s\rangle|^2 N_i(\omega_{ks}) \tag{19.73}$$

where $\hat{\mathbf{e}}_i$ denotes the polarization vector of the ith pulse, and $N_i(\omega_{ks})$ the intensity of its frequency component ω_{ks}. Expression (19.73) must be summed over the n light pulses which on the average strike the atom per unit time. Since in blackbody radiation there is no correlation between $\hat{\mathbf{e}}_i$ and $N_i(\omega_{ks})$, the average of the product

is the product of the averages; hence, the *rate of absorption* from the radiation field is given by

$$R = 4\pi^2 \alpha \omega_{ks} \overline{|\langle k|\mathbf{r} \cdot \hat{\mathbf{e}}|s\rangle|^2}\, \bar{N}(\omega_{ks}) n \tag{19.74}$$

where the bar denotes averages. The average intensity is related to the energy density per unit frequency interval, $u(\omega)$, by

$$n\bar{N}(\omega) = cu(\omega)$$

The average over the polarization directions $\hat{\mathbf{e}}$ can be performed easily, since for Cartesian components of $\hat{\mathbf{e}}$,

$$\int e_i e_j d\Omega_{\hat{\mathbf{e}}} = \frac{4\pi}{3} \delta_{ij} \tag{19.75}$$

Hence,

$$\overline{|\langle k|\mathbf{r} \cdot \hat{\mathbf{e}}|s\rangle|^2} = \langle k|\mathbf{r}|s\rangle \cdot \overline{\hat{\mathbf{e}}\,\hat{\mathbf{e}}} \cdot \langle k|\mathbf{r}|s\rangle^* = \frac{1}{3} |\langle k|\mathbf{r}|s\rangle|^2 \tag{19.76}$$

and the average rate of absorption is

$$\boxed{R = \frac{4\pi^2}{3} \alpha \omega_{ks}\, c |\langle k|\mathbf{r}|s\rangle|^2\, u(\omega_{ks})} \tag{19.77}$$

Alternatively, the $\hat{\mathbf{e}}$-directional average

$$\overline{|Y_1^{m'-m}(\hat{\mathbf{e}})|^2} = \frac{1}{4\pi}$$

may be used to calculate from (19.72) and (19.74) the average rate of absorption by initial states of sharp angular momentum. The total transition probability from a given state $n\ell m$ to all the degenerate m-substates of the level $n'\ell'$ is obtained by summing over the allowed values of m'. Since the C-G coefficients can be shown to satisfy the relation

$$\sum_{m'} \langle \ell k m, m'-m|\ell k \ell' m'\rangle^2 = \frac{2\ell' + 1}{2\ell + 1} \tag{19.78}$$

it follows from (19.72) that for $\ell' = \ell + 1$,

$$\sum_{m'} \overline{|\langle n'\ell'm'|\mathbf{r} \cdot \hat{\mathbf{e}}|n\ell m\rangle|^2} = \frac{1}{3} \frac{\ell + 1}{2\ell + 1} (R_{n\ell}^{n'\ell'})^2 \tag{19.79}$$

and for $\ell' = \ell - 1$,

$$\sum_{m'} \overline{|\langle n'\ell'm'|\mathbf{r} \cdot \hat{\mathbf{e}}|n\ell m\rangle|^2} = \frac{1}{3} \frac{\ell}{2\ell + 1} (R_{n\ell}^{n'\ell'})^2 \tag{19.80}$$

Substituting these values in (19.74), we obtain the absorption rates or integrated absorption cross sections. Because the Clebsch-Gordan sum (19.78) is independent of the quantum number m, the average absorption rates do not depend on the particular initial magnetic substate. This consequence of rotational symmetry is traditionally known as the *principle of spectroscopic stability*.

Exercise 19.7. For the $1S \rightarrow 2P$ (Lyman α) transitions in the hydrogen atom, evaluate the total integrated absorption cross section (see Exercise 19.5).

A cross section, $\sigma_{\text{s.e.}}(\omega)$, for stimulated emission (s.e.) can be defined in analogy with the absorption cross section, $\sigma_{\text{abs}}(\omega)$, by equating to $N(\omega)\sigma_{\text{s.e.}}(\omega)\,d\omega$ the energy given off by an atom as a consequence of the passage of the light wave in the frequency interval $d\omega$ around ω. Since the transition probabilities for stimulated emission and absorption are the same, we see from (19.67) that for any initial state s,

$$\int_0^\infty [\sigma_{\text{abs}}(\omega) - \sigma_{\text{s.e.}}(\omega)]\,d\omega = 4\pi^2\alpha\left[\sum_{E_k>E_s}\omega_{ks}|\langle k|\mathbf{r}\cdot\hat{\mathbf{e}}|s\rangle|^2 - \sum_{E_k<E_s}\omega_{sk}|\langle s|\mathbf{r}\cdot\hat{\mathbf{e}}|k\rangle|^2\right]$$

provided that the sum is taken over all eigenstates of H_0. From the sum rule (19.37), we then deduce for electric dipole transitions in one-electron atoms the general formula

$$\int_0^\infty [\sigma_{\text{abs}}(\omega) - \sigma_{\text{s.e.}}(\omega)]\,d\omega = 4\pi^2\alpha\,\frac{\hbar}{2m} = 2\pi^2\,\frac{e^2}{mc^2}\,c = 2\pi^2 r_0 c \qquad (19.81)$$

where r_0 is the classical electron radius. This relation is independent of the direction of polarization and of the particular initial state and holds for any system for which the Thomas-Reiche-Kuhn sum rule is valid.

For an atom with Z electrons, the sum rule (19.38) holds and gives

$$\int_0^\infty [\sigma_{\text{abs}}(\omega) - \sigma_{\text{s.e.}}(\omega)]\,d\omega = 2\pi^2\,Zr_0c \qquad (19.82)$$

provided that all possible electric dipole transitions of the atom are included. Equation (19.82) is identical with the result of a similar classical calculation.[2]

Sum rules for electric dipole cross sections can also be derived for nuclei that are exposed to gamma radiation, but care is required because nuclei are composed of charged and neutral particles. Since electric dipole absorption comes about as a result of the relative displacement of the center of charge of the nucleus and its center of mass, the dipole sum rule for nuclei depends critically on the relative number of protons and neutrons.[3]

So far only the integrated cross section $\int\sigma(\omega)\,d\omega$ has been discussed; the detailed frequency dependence of $\sigma(\omega)$ was left unspecified. All that could be said about it was that it has a very pronounced peak at $\omega = \omega_{ks}$. In the approximation used in this chapter, no absorption occurs unless the exact frequency component ω_{ks} is represented in the incident beam. Hence, $\sigma(\omega)$ must be proportional to a delta function and according to (19.61) must be given by

$$\sigma(\omega) = \frac{4\pi^2\alpha}{m^2\omega_{ks}}|\langle k|\exp\left(i\,\frac{\omega_{ks}}{c}\,\hat{\mathbf{n}}\cdot\mathbf{r}\right)\mathbf{p}\cdot\hat{\mathbf{e}}|s\rangle|^2\delta(\omega - \omega_{ks}) \qquad (19.83)$$

Equation (19.83) is, of course, an approximation. Observed spectral absorption lines are not infinitely sharp, but possess a finite width that the simple theory neglects. In actual fact, the excited state k is not, as was tacitly assumed, a stable and therefore strictly discrete energy level. Rather, it decays with a characteristic lifetime τ back to the original level s, and possibly to other lower lying levels if such are available. Besides, an atomic level is depleted not only by emission of radiation but also, and often predominantly, by collisions with other atoms. In a more accurate form of

[2]Jackson (1975), Section 17.8.
[3]Blatt and Weisskopf (1952), pp. 640–643.

perturbation theory, these interactions, which also cause transitions, must be included. Thus, the shape of the absorption line with its finite width can be predicted only if the subsequent return of the atom to its ground state is taken into account. Section 19.8 will show how interactions lead to the exponential decay of excited states and to a more realistic expression for the line shape than (19.83).

In retrospect, it may be well to stress some tacit assumptions that underlie the treatment of absorption given in this section. We have assumed that the incident radiation is weak, so that no atom is ever hit by several light pulses simultaneously; in fact, we have assumed that the atom could settle down to an undisturbed existence for comparatively long periods between pulses. A sufficient number of collisions in these intervals ensure that almost certainly the atom is in the ground state when the next pulse arrives, or that at least thermal equilibrium has been restored, and each atom is in a definite unperturbed energy eigenstate. We have also assumed that each pulse is sufficiently weak so that the probability of absorption per single pulse is numerically small. If, in violation of these assumptions, the incident radiation were very intense, we would have to use a more accurate form of perturbation theory with the result that, in the case of such *saturation* of the absorption line, additional broadening of the line would be observed. This broadening occurs because the incident radiation stimulates emissive transitions appreciably, and furthermore there is a measurable depletion of the ground state. The additional broadening of the absorption line depends on the intensity of the incident radiation.

Lasers and masers can produce conditions that are totally incompatible with the assumptions outlined above. The radiation in these devices is coherent, sharply monochromatic, and extremely intense, and it is essential for their operation that excited atomic states of the active substance are appreciably populated. If, under these circumstances, only two levels r and s of an atomic system, which interacts with the radiation, are important, Eqs. (19.24) may be employed with the perturbation

$$V(t) = \frac{e}{mc}\,(e^{i\omega t}\mathbf{A}_0 + e^{-i\omega t}\,\mathbf{A}_0^*)\cdot\mathbf{p} \tag{19.84}$$

Here, $\mathbf{A}_0$ is a complex-valued constant vector, and the electric dipole approximation has been made. If selection rules cause the diagonal elements of V in the states r and s to vanish and if, furthermore, the off-diagonal matrix element $\langle r|\mathbf{A}_0\cdot\mathbf{p}|s\rangle \neq 0$, Eqs. (19.24) assume the simple form

$$i\hbar\,\frac{d}{dt}\begin{pmatrix} c_r \\ c_s \end{pmatrix} = \begin{pmatrix} 0 & \alpha e^{i[(\omega_{rs}-\omega)t+\delta]} \\ \alpha e^{-i[(\omega_{rs}-\omega)t+\delta]} & 0 \end{pmatrix}\begin{pmatrix} c_r \\ c_s \end{pmatrix} \tag{19.85}$$

where

$$c_r = \langle r|\tilde{T}(t, t_0)|\Psi(t_0)\rangle \quad \text{and} \quad c_s = \langle s|\tilde{T}(t, t_0)|\Psi(t_0)\rangle$$

denote transition amplitudes. The positive constant α is proportional to the strength of the electromagnetic field and to the magnitude of the dipole matrix element. Since we plan to apply the solutions of (19.85) near resonance, $\omega \approx \omega_{rs}$, in setting up this equation we have neglected the rapidly oscillating matrix elements that are proportional to $e^{\pm i(\omega_{rs}+\omega)t}$.

Exercise 19.8. Equation (19.85) is formally the same as the equation of motion for a spin one-half particle with magnetic moment in a rotating magnetic field. Apply

the technique of solving the dynamics of the particle with spin to the problem of a two-level atom interacting with a classical monochromatic electromagnetic wave in the limit of long wavelength. Derive the analogue of Rabi's formula, discuss the resonance behavior of the transition probability, and examine the dependence of the resonance width on the physical parameters.

Finally, it should be remembered that in this chapter the field is regarded as a prescribed external agent, perturbing the quantum system such as an atom but itself incapable of being reacted upon. Evidently, this approximation is quite restrictive and ignores all effects in the interaction between light and matter which go beyond simple absorption and stimulated emission, notably the scattering of light. Even the process of *spontaneous emission*, i.e., emission from an excited atom in the absence of all applied electromagnetic fields, cannot be described properly without including the field as a dynamical system that can be found in various quantum states of excitation, quite similar to the material system with which it interacts. The elements of the *quantum theory of radiation* will be taken up in Chapter 23.

6. ***The Photoelectric Effect.*** In the photoelectric effect, the absorption of light by an atom (or nucleus) leads to the emission of an electron (or nucleon) into a continuum state. In contradistinction to the sharp-line behavior of $\sigma(\omega)$ for jumps from discrete to discrete levels, the cross section for the photoelectric effect is a smooth function of the frequency. Hence, by comparing (19.59) with (19.60) and adapting the notation to the transitions into the continuum of final states, we obtain

$$d\sigma(\omega) = \frac{4\pi^2\alpha}{m^2\omega}\left|\langle k|\exp\left(i\frac{\omega}{c}\,\hat{\mathbf{n}}\cdot\mathbf{r}\right)\mathbf{p}\cdot\hat{\mathbf{e}}|s\rangle\right|^2\frac{\Delta n}{\Delta\omega} \tag{19.86}$$

Here, ω is the frequency of the incident radiation and of the quantum jump. As before, s is the discrete initial state, but k, the final state, now lies in the continuum. The quantity $d\sigma(\omega)$ denotes the differential cross section corresponding to emission of the electron into a sharply defined solid angle $d\Omega$ with respect to the linearly polarized beam. Finally, Δn is the number of electron eigenstates in the frequency interval $\Delta\omega$, with an average energy $E_k = \hbar\omega + E_s$ and corresponding to the given solid angle.

To exemplify the evaluation of (19.86), the extreme assumption will be made that $\hbar\omega$ is much larger than the ionizing potential of the atom—so large in fact that the final electron states can be satisfactorily approximated by plane waves instead of the correct positive-energy atomic electron wave functions (e.g., screened Coulomb wave functions). If we use plane waves that are normalized in a very large cube of length L containing the atom, we have

$$\langle\mathbf{r}|k\rangle = \frac{1}{L^{3/2}}\exp(i\mathbf{k}\cdot\mathbf{r}), \qquad E_k = \frac{\hbar^2k^2}{2m}$$

and according to (4.57) the density of electron energy eigenstates,

$$\frac{\Delta n}{\Delta\omega} = \frac{d\Omega}{4\pi}\,\frac{m^{3/2}\sqrt{E_k}\,L^3}{\sqrt{2}\pi^2\hbar^2}$$

If the electron is initially in the K shell, ejection from which is comparatively probable, the initial wave function is given by (12.98) as

$$\langle\mathbf{r}|s\rangle = \frac{1}{\sqrt{\pi}}\left(\frac{Z}{a}\right)^{3/2}\exp\left(-\frac{Zr}{a}\right)$$

Substituting all these quantities into (19.86), we obtain for the differential photoemission cross section

$$\frac{d\sigma(\omega)}{d\Omega} = \frac{e^2k}{2mc\hbar^2\pi^2\omega}\frac{Z^3}{a^3}\left|\int \exp(-i\mathbf{k}\cdot\mathbf{r})\exp\left(i\frac{\omega}{c}\hat{\mathbf{n}}\cdot\mathbf{r}\right)\hat{\mathbf{e}}\cdot\frac{\hbar}{i}\nabla\exp\left(-\frac{Zr}{a}\right)d^3r\right|^2$$

Owing to the Hermitian property of the momentum operator, and the transverse character of the light waves ($\hat{\mathbf{e}}\cdot\hat{\mathbf{n}} = 0$), we can evaluate the integral:

$$\begin{aligned}
&\int \exp\left(-i\mathbf{k}\cdot\mathbf{r} + i\frac{\omega}{c}\hat{\mathbf{n}}\cdot\mathbf{r}\right)\hat{\mathbf{e}}\cdot\frac{\hbar}{i}\nabla\exp\left(-\frac{Zr}{a}\right)d^3r \\
&= \hat{\mathbf{e}}\cdot\int\left\{\frac{\hbar}{i}\nabla\exp\left[i\left(\mathbf{k}-\frac{\omega}{c}\hat{\mathbf{n}}\right)\cdot\mathbf{r}\right]\right\}^*\exp\left(-\frac{Zr}{a}\right)d^3r \\
&= \hbar\hat{\mathbf{e}}\cdot\mathbf{k}\int\exp\left[i\left(\frac{\omega}{c}\hat{\mathbf{n}}-\mathbf{k}\right)\cdot\mathbf{r}-\frac{Zr}{a}\right]d^3r = \hat{\mathbf{e}}\cdot\mathbf{k}\,\frac{8\pi Z\hbar}{a}\left(\frac{Z^2}{a^2}+q^2\right)^{-2}
\end{aligned}$$

where $\mathbf{q}$ is the momentum transfer in units of $\hbar$, i.e.,

$$\mathbf{q} = \mathbf{k} - \frac{\omega}{c}\hat{\mathbf{n}}$$

Hence, finally,

$$\boxed{\frac{d\sigma(\omega)}{d\Omega} = \frac{32e^2k(\hat{\mathbf{e}}\cdot\mathbf{k})^2}{mc\omega}\frac{Z^5}{a^5}\left(\frac{Z^2}{a^2}+q^2\right)^{-4}} \tag{19.87}$$

which is in qualitative agreement with experimental observations.

If $\hbar\omega$ is comparable with the ionization potential, a better approximation to the final state electron wave function must be used. Ideally, exact continuum eigenfunctions of the atomic Hamiltonian could be used which represent the emission of an electron into a state approaching asymptotically the plane wave $e^{i\mathbf{k}\cdot\mathbf{r}}$. Such a state is $\psi_{\mathbf{k}}^{(-)}(\mathbf{r})$ defined by (13.39). The state $\psi_{\mathbf{k}}^{(-)}(\mathbf{r})$ with asymptotically *incoming* waves rather than $\psi_{\mathbf{k}}^{(+)}(\mathbf{r})$ with *outgoing* waves must be chosen, because a wave packet of approximate momentum $\mathbf{k}$ moving *away* from the origin is constructed as a superposition of neighboring functions $\psi_{\mathbf{k}}^{(-)}(\mathbf{r})$, just as an incident wave packet moving *toward* the origin was shown in Chapter 13 to be a simple superposition of $\psi_{\mathbf{k}}^{(+)}(\mathbf{r})$ functions. Since after inclusion of the bound states both $\psi_{\mathbf{k}}^{(+)}(\mathbf{r})$ and $\psi_{\mathbf{k}}^{(-)}(\mathbf{r})$ are separately complete sets, a nearly plane wave packet moving away from the origin could in principle be expanded in terms of $\psi_{\mathbf{k}}^{(+)}(\mathbf{r})$ as well as $\psi_{\mathbf{k}}^{(-)}(\mathbf{r})$. In the former case, momenta $\mathbf{k}$ of all directions would be needed in the superposition in order to mimic the desired wave packet by judicious destructive interference. In this approach, the cross section calculation would be far more complicated than if the sharply defined set of $\psi_{\mathbf{k}}^{(-)}(\mathbf{r})$ eigenfunctions are used.[4]

Exercise 19.9. Derive the hydrogenic ground state wave function in momentum space, and show that it can be used in the calculation of the cross section for the photoelectric effect in a one-electron atom with nuclear charge Ze.

[4]See Gottfried (1966), Section 58.

7. *The Golden Rule for Constant Transition Rates.* In many cases of practical interest, the perturbation V does not itself depend on the time, but the time development of the system with a Hamiltonian $H = H_0 + V$ is nevertheless conveniently described in terms of transitions between eigenstates of the unperturbed Hamiltonian H_0. For example, a fixed potential scatters a beam of particles from an initial momentum state into various possible final momentum states, thus causing transitions between the eigenstates of the free particle Hamiltonian. Scattering will be discussed from this point of view in Chapter 20.

The theory is again based on the approximate first-order equation (19.21), but we now choose $t_0 = 0$:

$$\langle k|\tilde{T}(t, 0)|s\rangle = -\frac{i}{\hbar}\int_0^t \langle k|V(t')|s\rangle e^{i\omega_{ks}t'}\,dt' \tag{19.88}$$

If V does not depend on the time, the integration is trivial. We get for $k \neq s$,

$$\langle k|\tilde{T}(t, 0)|s\rangle = \frac{\langle k|V|s\rangle}{E_k - E_s}(1 - e^{i\omega_{ks}t}) \tag{19.89}$$

At $t = 0$, the system is known to have been in the initial discrete state s. The transition probability that at time t it will be in the unperturbed eigenstate $k \neq s$ is given by

$$\boxed{P_{k\leftarrow s}(t) = |\langle k|\tilde{T}(t, 0)|s\rangle|^2 = 2|\langle k|V|s\rangle|^2 \frac{1 - \cos\,\omega_{ks}t}{(E_k - E_s)^2}} \tag{19.90}$$

This is an oscillating function of t with a period $\dfrac{2\pi}{|\omega_{ks}|}$, except at or near $\omega_{ks} = 0$. Its derivation from (19.21) presupposes that the time development operator remains close to its initial value, unity, in the interval $(0, t)$. During such times, the probability $P_{s\leftarrow s}(t) = |\langle s|\tilde{T}(t, 0)|s\rangle|^2 \approx 1$, and the transition probabilities to states $k \neq s$ must stay small. We know from the Rayleigh-Schrödinger perturbation theory that the eigenstate $|s\rangle$ of H_0 is very nearly an eigenstate also of the complete Hamiltonian H if the perturbation is weak enough so that

$$\left|\frac{\langle k|V|s\rangle}{E_k - E_s}\right| \ll 1 \tag{19.91}$$

This same inequality ensures that the transition probabilities to states with an unperturbed energy appreciably different from the initial value can never become large. On the other hand, transitions to states with energies that are nearly equal to E_s can be very important. If such states exist, there may be a good chance of finding the system in them after some time has elapsed. Since transitions to states with an E_k radically different from E_s are rare, we may say that if the perturbation is weak, the unperturbed energy is nearly conserved. Of course, the *total energy* is strictly conserved, and $\langle H\rangle$ is a rigorous constant of the motion.

From (19.88) we see that as long as

$$|\omega_{ks}t| \ll 1 \tag{19.92}$$

transitions to states with $E_k \approx E_s$ have amplitudes proportional to t, and (19.90) becomes approximately

$$P_{k \leftarrow s}(t) \approx \frac{1}{\hbar^2} |\langle k|V|s\rangle|^2 t^2 \tag{19.93}$$

The transition probability increases quadratically with time. This circumstance has important consequences if many states k are available having an energy $E_k \approx E_s$, as will happen if H_0 has a near-continuum of eigenstates in the vicinity of E_s.

As a concrete example on which we can fix our ideas, let us consider the helium atom again, regarding the electrostatic repulsion between the electrons as a perturbation (and neglecting radiative transitions altogether):

$$H_0 = \frac{\mathbf{p}_1^2}{2m} + \frac{\mathbf{p}_2^2}{2m} - \frac{2e^2}{r_1} - \frac{2e^2}{r_2}, \qquad V = \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{19.94}$$

The excited state in which both electrons occupy the hydrogenic $2s$ level has an unperturbed energy of $6 \times 13.6 = 81.6$ eV above the ground state. Since this energy is greater than the energy required (54.4 eV) to lift one electron into the continuum with zero kinetic energy, the $(2s)^2$ level coincides in energy with a continuum state formed by placing one electron in the bound $1s$ state and the other electron in the continuum. But the wave functions of these two types of states are very different. The first is a discontinuous function of E, and the second is a continuous one. Figure 19.2 presents an energy level diagram for the unperturbed two-electron system. If at $t = 0$ the atom has been excited to the $(2s)^2$ state, the electrostatic Coulomb repulsion V between the electrons will cause *radiationless* (autoionizing or Auger) transitions to the continuum. Radiative transitions to lower levels with the emission of light also take place but are much less probable. The effects of radiationless transitions from autoionizing states, like the $(2s)^2$ state in the helium atom, are observable in electron-atom collisions. Auger transitions are observed in heavy atoms that have been excited by the removal of an electron from an inner shell. From such states the emission into the continuum of so-called *Auger electrons* competes with the emission of X rays. The internal conversion of nuclear gamma rays is another example of radiationless transitions.

Since the formulas in this section were developed for quadratically integrable eigenstates of H_0, it is convenient to imagine the entire system to be enclosed in a very large box, or better still, to enforce discreteness of the unperturbed eigenstates by imposing on the unbound states periodic boundary conditions in such a large box. (In our example, these states are approximately plane waves, although at the lower energies they must be taken to be appropriately screened Coulomb wave functions.) In order to be able to replace sums by integrals, the number of these *quasi-continuum* states *per unit energy interval* is introduced. This density of final unperturbed states is denoted by $\rho_f(E)$.

According to (19.90), the total transition probability to all final states under consideration, labeled collectively by f, is given by

$$\sum_{k \in f} P_{k \leftarrow s}(t) = 2 \int |\langle k|V|s\rangle|^2 \frac{1 - \cos \omega_{ks} t}{(E_k - E_s)^2} \rho_f(E_k)\, dE_k \tag{19.95}$$

In usual practice, both $|\langle k|V|s\rangle|^2$ and $\rho_f(E_k)$ are sensibly constant over an energy range ΔE in the neighborhood of E_s. In this same interval, as a function of E_k the factor $(1 - \cos \omega_{ks} t)/(E_k - E_s)^2$ varies rapidly for all t satisfying

$$t \Delta E \gg \hbar \tag{19.96}$$

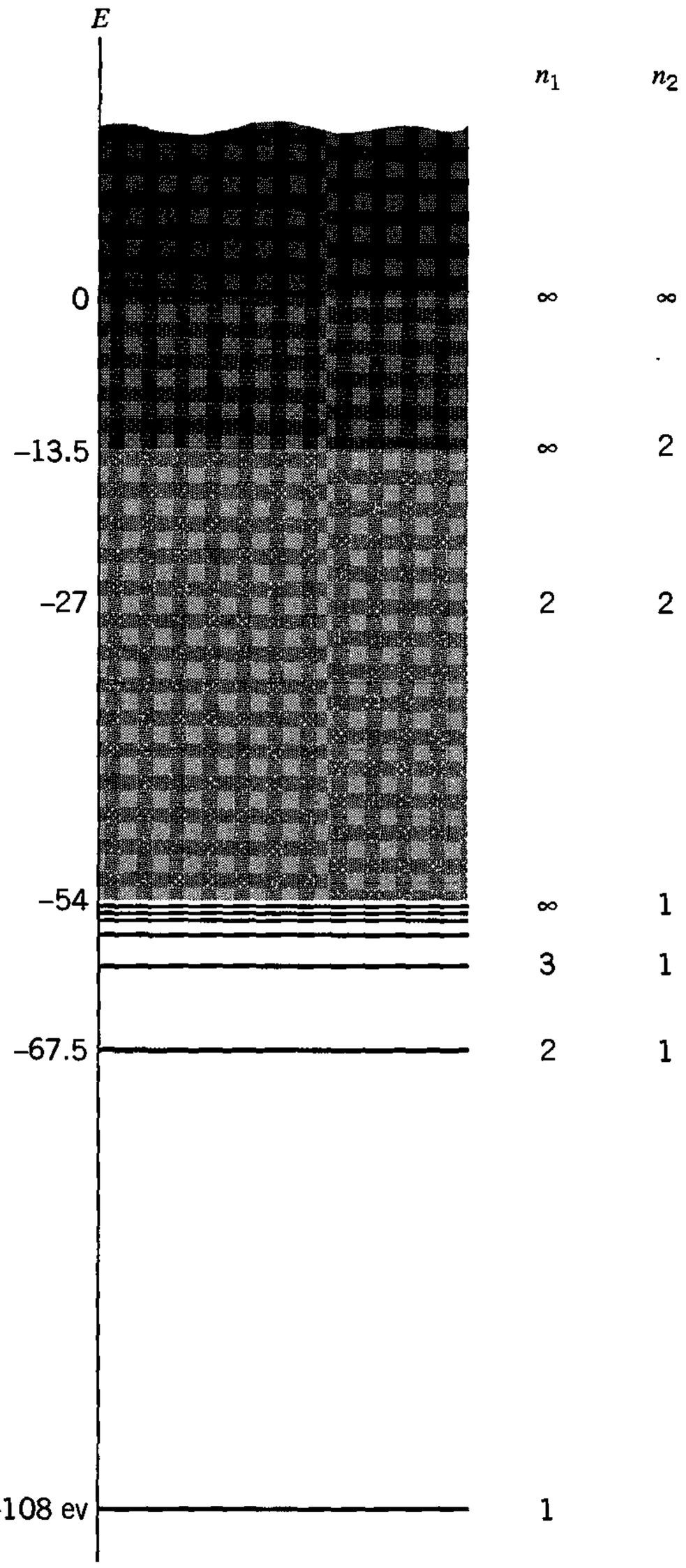


Figure 19.2. Partial energy-level diagram for the unperturbed states of two electrons in the Coulomb field of a helium nucleus. The principal quantum numbers for the two electrons are n_1 and n_2. The continuous portions of the spectrum are shaded. The doubly excited discrete state $n_1 = n_2 = 2$ is embedded in the continuum of states, as are all higher excited levels. After a radiationless autoionizing transition to the continuum, electron 1 is emitted with positive energy and an effective continuous n_1 value as defined in Eq. (13.113), while electron 2 drops to the $n_2 = 1$ level.

and has a pronounced peak at $E_k = E_s$. This behavior, shown in Figure 19.3, confirms that transitions which tend to conserve the unperturbed energy are dominant. The interval ΔE is generally comparable in magnitude to E_s itself. If this is so, $\hbar/\Delta E$ is a very short time, and there is a considerable range of t such that (19.96) is fulfilled and yet the initial state s is not appreciably depleted. During such times, our per-

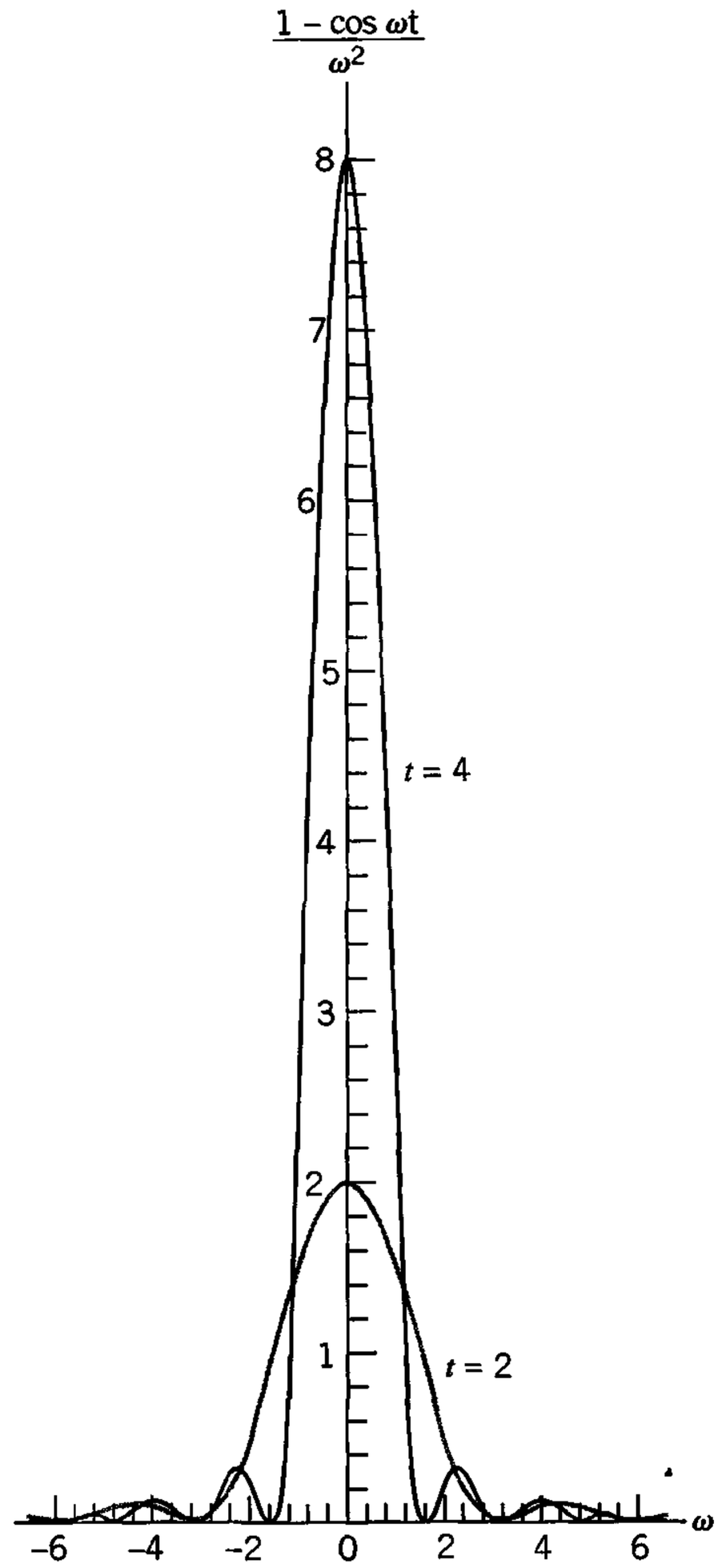


Figure 19.3. Plot of $(1 - \cos \omega t)/\omega^2$ versus ω for different values of the time, t ($t = 2$ and 4). The width of the peak is inversely proportional to t, and its height increases as t^2, so that the area under the curve is proportional to t.

turbation method is expected to be valid, and (19.95) can be simplified to good approximation as

$$\sum_{k \in f} P_{k \leftarrow s}(t) = \frac{2}{\hbar} |\langle k|V|s\rangle|^2 \rho_f(E_s) \int_{-\infty}^{+\infty} \frac{1 - \cos \omega_{ks} t}{\omega_{ks}^2} \, d\omega_{ks} \tag{19.97}$$

where condition (19.96) has been invoked in extending the limits of the integral to $\pm\infty$. Under the conditions stated, we can evaluate the integral for $t > 0$ as

$$\int_{-\infty}^{+\infty} \frac{1 - \cos \omega_{ks} t}{\omega_{ks}^2} \, d\omega_{ks} = t \int_{-\infty}^{+\infty} \frac{\sin x}{x} \, dx = \pi t \tag{19.98}$$

Hence, when summed over the set of similar final continuum states, the total transition probability increases *linearly* with time, and we can speak of a constant transition probability per unit time or *transition rate w*,

$$\boxed{w = \frac{d}{dt}\sum_{k\in f} P_{k\leftarrow s}(t) = \frac{2\pi}{\hbar}\,|\langle k|V|s\rangle|^2\rho_f(E_s)} \tag{19.99}$$

This remarkable result comes about because we have summed over transitions that conserve the unperturbed energy strictly ($E_k = E_s$) and transitions that violate this energy conservation ($E_k \neq E_s$). According to (19.93), the former increase quadratically with time, but the probability of the latter oscillates periodically. The sum is a compromise between these two tendencies: the transition rate is constant. Fermi has called formula (19.99) the *Golden Rule of time-dependent perturbation theory* because it plays a fundamental part in many applications.

In view of its universal relevance, it is useful to give an alternate heuristic derivation of the Golden Rule from (19.88). Without further discussion and detailed justification of each step, we write a chain of equalities:

$$\begin{aligned}
w &= \frac{dP_{k\leftarrow s}(t)}{dt} = \frac{d}{dt}\sum_{k\in f}|\langle k|\tilde{T}(t,0)|s\rangle|^2 \\
&= \frac{d}{dt}\sum_{k\in f}\frac{1}{\hbar^2}\int_0^t \langle k|V(t')|s\rangle e^{i\omega_{ks}t'}\,dt' \int_0^t \langle k|V(t'')|s\rangle^* e^{-i\omega_{ks}t''}\,dt'' \\
&\approx \frac{d}{dt}\frac{1}{\hbar}\int_{-\infty}^{+\infty} d\omega_{ks}\int_0^t\int_0^t dt'\,dt''\,\langle k|V(t')|s\rangle\langle k|V(t'')|s\rangle^*\;\; e^{i\omega_{ks}(t'-t'')}\rho_f(E_k) \\
&= \frac{d}{dt}\frac{2\pi}{\hbar}\,\rho_f\,(E_s)\int_0^t\int_0^t dt'\,|\langle k|V(t')|s\rangle|^2\delta(t'-t'')\,dt'' = \frac{2\pi}{\hbar}\,|\langle k|V(t)|s\rangle|^2\rho_f(E_s)
\end{aligned} \tag{19.100}$$

If V is constant in time, the Golden Rule results.

We illustrate the Golden Rule by applying it to the problem of radiationless atomic transitions, caused by the interaction (19.94), from an initial $(2s)^2$ two-electron state (see Figure 19.2). The Golden Rule requires that we evaluate the matrix element

$$\langle f|V|i\rangle = \iint \psi^*_{1s}(\mathbf{r}_2)\psi^*_{ks}(\mathbf{r}_1)\,\frac{e^2}{|\mathbf{r}_1-\mathbf{r}_2|}\,\psi_{2s}(\mathbf{r}_1)\psi_{2s}(\mathbf{r}_2)\,d^3r_1\,d^3r_2 \tag{19.101}$$

Here we have labeled the initial and final states by i and f, respectively. In the final state, one electron is in the lowest energy state ($1s$) and the other is in the continuum, with wave number k. When the standard multipole expansion

$$\begin{aligned}
\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|} &= \sum_{\ell=0}^{\infty}\frac{r_<^\ell}{r_>^{\ell+1}}P_\ell(\hat{\mathbf{r}}\cdot\hat{\mathbf{r}}') \\
&= 4\pi\sum_{\ell=0}^{\infty}\frac{r_<^\ell}{r_>^{\ell+1}}\frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell}Y_\ell^{m*}(\hat{\mathbf{r}}_1)Y_\ell^m(\hat{\mathbf{r}}_2)
\end{aligned} \tag{19.102}$$

is substituted in (19.101), it is evident that the matrix element vanishes, unless the continuum electron also is in an s state ($\ell = 0$), as anticipated in the notation for $\psi_{ks}(\mathbf{r}_1)$. If the continuum state is approximated by a plane wave corresponding to

energy $E_k = \hbar^2 k^2/2m = 2E_{2s} - E_{1s}$, the autoionization or Auger transition rate becomes, with this selection rule, for an atom orion with nuclear charge ze,

$$w = \frac{2\pi e^4}{\hbar}\left|(4\pi)^2 N \int_0^\infty\int_0^\infty e^{-Zr_2/a}\,\frac{\sin kr_1}{kr_1}\,\frac{1}{r_{1,2>}}\times e^{-Zr_1/2a}\left(1-\frac{Zr_1}{2a}\right)e^{-Zr_2/2a}\left(1-\frac{Zr_2}{2a}\right)r_1^2\,dr_1\,r_2^2\,dr_2\right|^2 \rho_f(E_k)$$

Here N denotes the product of the four normalization constants for the wave functions in the matrix element. The product of N^2 and the free particle energy density of final states is

$$N^2\,\rho_f(E_k) = \frac{mk}{2^7\pi^5\hbar^2}\left(\frac{Z}{a}\right)^9$$

In terms of the Bohr frequency, $me^4/\hbar^3$ ($\cong 4\times 10^{16}$ sec^{-1}), the calculation of the autoionization/Auger rate gives the simple result

$$w = C\,\frac{me^4}{\hbar^3} \tag{19.103}$$

where C is a numerical constant or order of 10^{-3}.

Exercise 19.10. Deduce the selection rule for autoionizing transitions from a $2s\,np$ initial to a $1s\,k\ell$ final state. Do the same for transitions from $(2p)^2$ to $1s\,k\ell$. Identify the terms in the multipole expansion of the electron-electron Coulomb interaction that are responsible for the transitions.

Exercise 19.11. Show that the approximate autoionization/Auger rate for the hydrogenic $(2s)^2$ state is independent of Z. Give a physical explanation. Is this conclusion changed, if Coulomb wave functions (Section 13.8) are used to describe the continuum state, instead of plane waves?

Even if the perturbation is weak, there may be good reasons for proceeding beyond the first perturbation term in evaluating the transition amplitude. If V is constant in time and we retain the second term in (19.20), we obtain for $t_0 = 0$ and $k \neq s$,

$$\begin{aligned}\langle k|\tilde{T}(t,0)|s\rangle = {} & \frac{\langle k|V|s\rangle}{E_k - E_s}\,(1 - e^{i\omega_{ks}t}) \\ & + \sum_m \langle k|V|m\rangle\,\frac{\langle m|V|s\rangle}{E_s - E_m}\left(\frac{1-e^{i\omega_{ks}t}}{E_k - E_s} - \frac{1-e^{i\omega_{km}t}}{E_k - E_m}\right)\end{aligned} \tag{19.104}$$

The application of this formula requires care. Nonreversing transitions at constant rates to states k which lie in a continuum with $E_k \approx E_s$ may still be possible in second order if the matrix elements $\langle m|V|s\rangle = 0$ for *all* states that conserve the unperturbed energy, $E_m \approx E_s$. Under these conditions, the second term inside the bracket in (19.104) is negligible because it oscillates in time with small amplitude. The Golden Rule still holds provided that the first-order transition matrix element of the perturbation V is replaced by the expression

$$\langle k|V|s\rangle \to \sum_m \frac{\langle k|V|m\rangle\langle m|V|s\rangle}{E_s - E_m} \tag{19.105}$$

This effective transition matrix element is sometimes interpreted by describing the transition as a two-step process, in which a *virtual* transition from the initial state s to an *intermediate state* m is followed by a second virtual transition from the state m to the final state k. The (unperturbed) energy of the intermediate state can be entirely different from the energy of the initial and final states, but in the overall transition from the initial to the final states energy is conserved.

In the general case, the second term in the bracket in (19.104) cannot be neglected. The energy denominators may vanish if there are energy-conserving intermediate states, and, as for the degenerate stationary state perturbation theory discussed in the last chapter, a more accurate treatment is required.[5]

A constant transition rate can also occur if the perturbation, instead of being constant in time, varies harmonically with frequency ω:

$$V(t) = V_0 e^{-i\omega t} + V_0^\dagger e^{i\omega t} \tag{19.106}$$

The interaction (19.84) of a charged particle with an electromagnetic wave is an example. Such a harmonic perturbation induces enduring transitions between an initial unperturbed state s and a final state k, if frequency matching $\omega_{ks} \approx \omega$ or $\omega_{sk} \approx \omega$ is achieved.

If $E_k > E_s$, we have the case $\omega_{ks} \approx \omega$. Substitution of (19.106) into (19.88) shows that (19.90) generalizes to

$$P_{k \leftarrow s}(t) = |\langle k|\tilde{T}(t, 0)|s\rangle|^2 = \frac{2}{\hbar^2} |\langle k|V|s\rangle|^2 \frac{1 - \cos(\omega_{ks} - \omega)t}{(\omega_{ks} - \omega)^2} \tag{19.107}$$

In arriving at this result, rapidly oscillating terms containing $e^{i(\omega_{ks}+\omega)t}$ are neglected because they are not effective in causing lasting transitions.

The same arguments that were employed in deriving the Golden Rule (19.99) can be invoked here to calculate the transition rate by evaluating the integral

$$\int_{-\infty}^{+\infty} \frac{1 - \cos(\omega_{ks} - \omega)t}{(\omega_{ks} - \omega)^2} d(\omega_{ks} - \omega) = \pi t \tag{19.108}$$

This procedure then leads to a generalized Golden Rule,

$$\boxed{w = \frac{2\pi}{\hbar} |\langle k|V_0|s\rangle|^2 \rho_f(E_s + \hbar\omega)} \tag{19.109}$$

if the perturbation is monochromatic, and the final states form a quasi-continuum with a final density of states ρ_f. The transition described here corresponds to absorption of energy induced by the harmonic perturbation (19.106).

As a specific example, we consider the absorption of electromagnetic radiation by an atom. The incident field is represented by a plane wave vector potential:

$$V(t) = \frac{e}{mc} \mathbf{A} \cdot \mathbf{p} = \frac{e}{mc} [A_0 e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)} + A_0^\dagger e^{-i(\mathbf{k}\cdot\mathbf{r} - \omega t)}] \hat{\mathbf{e}} \cdot \mathbf{p} \tag{19.110}$$

which is a limiting case of (19.44) for monochromatic radiation, and thus an infinitely extended plane wave packet. For such a wave, the time-average incident en-

[5]For an example, see the discussion of resonance fluorescence in Sakurai (1967), Section 2–6.

ergy per unit area and unit time, i.e., the *intensity*, is found from the Poynting vector (19.54) to be

$$I_0\hbar\omega = \frac{\omega^2}{2\pi c}|A_0|^2 \tag{19.111}$$

where I_0 is the incident flux of quanta (photons) with energy $\hbar\omega$. Since in the transition an energy $\hbar\omega = E_k^{(0)} - E_s^{(0)}$ is absorbed, the rate of energy absorption is

$$\hbar\omega\, w = \frac{2\pi e^2\omega}{m^2c^2}|A_0|^2|\langle k|e^{i\omega\hat{\mathbf{n}}\cdot\mathbf{r}/c}\,\hat{\mathbf{e}}\cdot\mathbf{p}|s\rangle|^2\rho_f(E_s + \hbar\omega) \tag{19.112}$$

If the final energy level is narrow, we can write

$$\rho_f(E_s + \hbar\omega) = \delta(E_k - E_s + \hbar\omega) = \hbar^{-1}\,\delta(\omega_{ks} - \omega) \tag{19.113}$$

The absorption cross section $\sigma(\omega)$ was defined in Section 19.4, where the incident radiation was realistically represented by a finite wave packet. Here we use the idealization of infinite plane waves. The cross section is then simply calculated from the energy absorption rate,

$$I_0\sigma(\omega) = w \tag{19.114}$$

Combining Eqs. (19.111)–(19.114), we obtain

$$\sigma(\omega) = \frac{4\pi^2e^2}{\omega m^2c\hbar}|\langle k|e^{i\omega\hat{\mathbf{n}}\cdot\mathbf{r}/c}\,\mathbf{p}\cdot\hat{\mathbf{e}}|s\rangle|^2\,\delta(\omega_{ks} - \omega) \tag{19.115}$$

which is the same result as (19.83).

Exercise 19.12. Derive the Golden Rule for stimulated emission induced by a harmonic perturbation (19.106). This case occurs when the final unperturbed energy E_k lies *below* the initial energy E_s.

8. ***Exponential Decay and Zeno's Paradox.*** Suppose that a system, which is perturbed by a constant V as described in the last section, is known to be in the unperturbed energy eigenstate state s at time t. The probability that the system will make a nonreversing transition in the ensuing time interval between t and $t + dt$ is equal to $w\,dt$, if the conditions under which the Golden Rule (19.99) was derived obtain and k symbolizes the totality of available final states. *Stochastic processes* with constant w are familiar in probability theory. The probability $P_s(t + dt)$ of finding the system in state s at time $t + dt$ can easily be derived if we argue as follows: The system will be in state s at time $t + dt$ only if (a) it was in s at time t, and (b) it has not decayed from this state during the interval dt. Since the probability for not decaying is $(1 - w\,dt)$, we have

$$P_s(t + dt) = P_s(t)(1 - w\,dt) \tag{19.116}$$

with the initial condition $P(0) = 1$. Solving (19.116), we infer the probability of finding the system at any time t still undecayed in the initial state:

$$\boxed{P_s(t) = e^{-wt}} \tag{19.117}$$

This is the famous *exponential decay law*. The property (19.116) is implied by the more general relation

$$P_s(t_1 + t_2) = P_s(t_1)P_s(t_2) \tag{19.118}$$

which is characteristic of the exponential function.

Caution is required in making the preceding argument because it assumes that $P_s(t)$ changes only by virtue of transitions *out of* state s *into* other states but disregards the possibility that state s may be replenished by transitions from other states. In particular, it assumes that there are no reverse transitions from the final states back into the initial state. Even if these assumptions are valid, the probability $P_s(t + dt)$ can be equated to the product of $P_s(t)$ and $(1 - w\,dt)$ only if the actual determination of whether or not at time t the system is in state s does not influence its future development. In general, this condition for the validity of (19.116) is emphatically not satisfied in quantum mechanics. Usually, such a measurement interferes with the course of events and alters the chances of finding the system in s at time $t + dt$ from what it would have been had we refrained from attempting the measurement. Obviously, an explanation is required.

If the measurement is performed within a very short time interval, in violation of condition (19.96) and before the system has begun to populate the final states with a probability increasing linearly in time, the effects on the initial state can be dramatic, as we will see later. First, however, we imagine that the time interval dt, while very short compared with the depletion time of the initial state, is much longer than $\hbar/\Delta E$, as demanded by (19.96).

Starting with

$$P_s(0) = |\langle s|\tilde{T}(0, 0)|s\rangle|^2 = 1$$

quantum mechanics requires that the probability at all times is to be calculated from the *amplitudes*, which are the primary concepts, hence

$$P_s(t) = |\langle s|\tilde{T}(t, 0)|s\rangle|^2 \quad \text{and} \quad P_s(t + dt) = |\langle s|\tilde{T}(t + dt, 0)|s\rangle|^2$$

There is in general no reason why these expressions should be related as

$$|\langle s|\tilde{T}(t + dt, 0)|s\rangle|^2 = |\langle s|\tilde{T}(t, 0)|s\rangle|^2(1 - w\,dt) \tag{19.119}$$

as was assumed in writing (19.116). However, it can be shown that the very conditions which ensure a constant transition rate also establish the validity of Eqs. (19.116) and (19.119).

We can gain a qualitative understanding of the irreversible decay of a discrete initial state s embedded in a continuum of final states k with similar unperturbed energies, if we recall that the time evolution of the transition amplitudes is governed by the coupled linear integral equations (19.15), or their equivalent differential form, adapted to the choice $t_0 = 0$:

$$i\hbar \frac{d}{dt}\langle k|\tilde{T}(t, 0)|s\rangle = \sum_n e^{i\omega_{kn}t}\langle k|V|n\rangle\langle n|\tilde{T}(t, 0)|s\rangle \tag{19.120}$$

According to these equations, a state n feeds a state k if $\langle k|V|n\rangle \neq 0$. Thus, transitions from the initial state s to the various available final states k occur, and at the same time these final states contribute to the amplitude of the initial state. As the amplitudes of the final states k increase from their initial value, zero, they must grow at the expense of the initial state, since probability is conserved and the time development operator $\tilde{T}(t, 0)$ is unitary. We might expect that as the amplitudes of the states k increase, they would begin to feed back into state s. Indeed, this is what happens, but because of the different frequencies ω_{ks} with which this feedback occurs, the contributions from the many transition amplitudes $\langle k|\tilde{T}(t, 0)|s\rangle$ to the equation of motion for $\langle s|\tilde{T}(t, 0)|s\rangle$ are all of different phases. Hence, if there are many states k, forming essentially a continuum, these contributions tend to cancel. It is

this destructive interference which causes the gradual and irreversible exponential depletion of the initial state without corresponding regeneration.

To make these notions more precise and derive the exponential decay law, we must solve the equations of motion under the same assumptions as before (i.e., constant perturbation V, transitions from a discrete initial state s to a quasi-continuum of final states), but we must remove the uncomfortable restriction to times that are short compared with the lifetime of the initial state. In effect, this means that it is no longer legitimate to replace the transition amplitude $\langle s|\tilde{T}(t, 0)|s\rangle$ on the right-hand side of (19.120) by its initial value, $\langle s|\tilde{T}(t, 0)|s\rangle \approx 1$. However—and this is the fundamental assumption here—we continue to neglect all other contributions to the change in $\langle k|\tilde{T}(t, 0)|s\rangle$ and for $t \geq 0$ improve on the first-order approximation (19.88) by using the equations

$$i\hbar \frac{d}{dt} \langle k|\tilde{T}(t, 0)|s\rangle = e^{i\omega_{ks}t}\langle k|V|s\rangle\langle s|\tilde{T}(t, 0)|s\rangle \quad (k \neq s) \tag{19.121}$$

The justification for this assumption is essentially physical and a posteriori. To a certain extent, it is based on our previous experience with the short-term approximation. If the perturbation is weak, $\langle k|\tilde{T}(t, 0)|s\rangle$ will remain small for those transitions for which ω_{ks} is appreciably different from zero. Only those transition amplitudes $\langle k|\tilde{T}(t, 0)|s\rangle$ are likely to be important which conserve unperturbed energy, such that $\omega_{ks} \approx 0$. On the other hand, the matrix elements $\langle k|V|n\rangle$ that connect two possible final unperturbed states k, $n \neq s$ for which $E_k \approx E_n$, are usually small and will be neglected. (Such transitions are, however, basic in scattering processes; see Chapter 20.)

The integral form of Eq. (19.121) is

$$\langle k|\tilde{T}(t, 0)|s\rangle = -\frac{i}{\hbar} \langle k|V|s\rangle \int_0^t e^{i\omega_{ks}t'}\langle s|\tilde{T}(t', 0)|s\rangle \, dt' \quad (k \neq s) \tag{19.122}$$

The equation of motion for $\langle s|\tilde{T}(t, 0)|s\rangle$ is, rigorously,

$$i\hbar \frac{d}{dt} \langle s|\tilde{T}(t, 0)|s\rangle = \sum_{k \neq s} e^{i\omega_{ks}t}\langle s|V|k\rangle\langle k|\tilde{T}(t, 0)|s\rangle + \langle s|V|s\rangle\langle s|\tilde{T}(t, 0)|s\rangle \tag{19.123}$$

The term $k = s$ has been omitted from the sum and appears separately on the right-hand side. (In decay problems, we frequently have $\langle s|V|s\rangle = 0$, but in any case this term produces only a shift in the unperturbed energy levels s, as is already known from the Rayleigh-Schrödinger perturbation theory.)

If (19.122) is substituted into (19.123), we obtain the differential-integral equation for the probability amplitude that the system will at time $t \geq 0$ still dwell in the initial state s:

$$\begin{aligned} \frac{d}{dt} \langle s|\tilde{T}(t, 0)|s\rangle = &-\frac{1}{\hbar^2} \sum_{k \neq s} |\langle k|V|s\rangle|^2 \int_0^t \langle s|\tilde{T}(t', 0)|s\rangle e^{i\omega_{ks}(t'-t)} \, dt' \\ &- \frac{i}{\hbar} \langle s|V|s\rangle\langle s|\tilde{T}(t, 0)|s\rangle \end{aligned} \tag{19.124}$$

The solution of this equation demands care. We are interested in times t which imply rapid oscillations of the factor $e^{i\omega_{ks}(t'-t)}$ in the integrand as a function of the final state energy, E_k. The slowly varying amplitude $\langle s|\tilde{T}(t, 0)|s\rangle$ can therefore be re-

moved from the t' integrand, and the remaining integral can be evaluated by using the formula (A.19) from the Appendix:

$$\int_0^t e^{-i(\omega - i\varepsilon)\tau}\, d\tau = \frac{1}{i\omega + \varepsilon} = \pi\delta(\omega) - i\mathsf{P}\frac{1}{\omega} \quad (\omega t \gg 1) \tag{19.125}$$

Here P denotes the Cauchy principal value. The resulting differential equation is solved for $t \geq 0$ by

$$\boxed{\langle s|\tilde{T}(t, 0)|s\rangle = \exp\left[-\frac{\pi t}{\hbar^2}\sum_{k\neq s}|\langle k|V|s\rangle|^2\,\delta(\omega_{ks}) - \frac{it}{\hbar}\Delta E_s\right]} \tag{19.126}$$

where we have denoted the perturbative energy shift of the level s, up to second order in V, by

$$\Delta E_s = \langle s|V|s\rangle + \sum_{k\neq s}\frac{|\langle k|V|s\rangle|^2}{E_s - E_k} \tag{19.127}$$

Equation (19.126) is the anticipated result, since under the assumptions made in deriving the Golden Rule,

$$w = \frac{2\pi}{\hbar^2}\sum_{k\neq s}|\langle k|V|s\rangle|^2\delta(\omega_{ks}) = \frac{2\pi}{\hbar}|\langle k|V|s\rangle|^2\rho_f(E_s) \tag{19.128}$$

Hence, we see that

$$\boxed{\langle s|\tilde{T}(t, 0)|s\rangle = \exp\left(-\frac{w}{2}t - \frac{i}{\hbar}t\Delta E_s\right)} \tag{19.129}$$

describing the exponential decay of the unstable state.

To obtain nonreversing transitions and a progressive depletion of the initial state, it is essential that the discrete initial state be coupled to a large number of final states with similar energies. However, the fact remains that the exponential decay law, for which we have so much empirical support in radioactive decay processes, is not a rigorous consequence of quantum mechanics but the result of somewhat delicate approximations.

If (19.129) is substituted back into (19.122), the integration can be carried out and we obtain for $t \geq 0$,

$$\langle k|\tilde{T}(t, 0)|s\rangle = \langle k|V|s\rangle\frac{1 - \exp\left(-\frac{w}{2}t\right)\exp\left[-\frac{i}{\hbar}(E_s + \Delta E_s - E_k)t\right]}{E_k - (E_s + \Delta E_s) + i\hbar\frac{w}{2}} \tag{19.130}$$

Hence, the probability that the system has decayed into state k is

$$\boxed{P_{k\leftarrow s}(t) = |\langle k|V|s\rangle|^2\frac{1 - 2\exp\left(-\frac{\Gamma}{2\hbar}t\right)\cos\left(\frac{E_s + \Delta E_s - E_k}{\hbar}t\right) + \exp\left(-\frac{\Gamma}{\hbar}t\right)}{(E_k - E_s - \Delta E_s)^2 + \frac{\Gamma^2}{4}}} \tag{19.131}$$

where we have set $\Gamma = \hbar w$. After a time that is very long compared with the lifetime $\hbar/\Gamma$, we obtain the distribution

$$\boxed{P_{k\leftarrow s}(+\infty) = \frac{|\langle k|V|s\rangle|^2}{(E_k - E_s - \Delta E_s)^2 + \dfrac{\Gamma^2}{4}}} \tag{19.132}$$

exhibiting the typical bell-shaped resonance behavior with a peak at $E_k = E_s + \Delta E_s$ and a width equal to Γ. The transition probability to a selected final state k oscillates as a function of time before it reaches its asymptotic value (19.132). In spite of these oscillations, the total probability of finding the system in any final state increases monotonically with time.

Exercise 19.13. Prove from (19.131) that the sum (integral) $\sum_{k\neq s} |\langle k|\tilde{T}(t,0)|s\rangle|^2$ over the final states is equal to $1 - \exp(-\Gamma t/\hbar)$ as required by the conservation of probability.

In a somewhat imprecise manner, the results of this section can be interpreted as implying that the interaction causes the state s to change from a strictly stationary state of H_0 with energy E_s into a decaying state with (normalized) probability $p(E)\,dE$ for having an energy between E and $E + dE$:

$$\boxed{p(E)\,dE = \frac{1}{2\pi}\frac{\Gamma}{(E - E_s)^2 + \dfrac{\Gamma^2}{4}}\,dE} \tag{19.133}$$

A remark about the shape of the absorption cross section calculated in Section 19.5 may now be made. If the mechanism responsible for the depletion of an energy level E_2, after it is excited by absorption of radiation from a stable discrete energy level E_1, broadens the excited state in accord with (19.133), the absorption probability for the frequency $\omega = (E - E_1)/\hbar$ must be weighted by the probability (19.133). The absorption cross section, instead of (19.83), is therefore more accurately:

$$\sigma(\omega) = \frac{4\pi^2\alpha}{m^2\omega_{ks}}\left|\langle k|\exp\left(i\frac{\omega_{ks}}{c}\,\hat{\mathbf{n}}\cdot\mathbf{r}\right)\mathbf{p}\cdot\hat{\mathbf{e}}|s\rangle\right|^2 \frac{1}{2\pi}\frac{\Gamma/\hbar}{(\omega - \omega_{21})^2 + \Gamma^2/4\hbar^2} \tag{19.134}$$

showing the characteristic resonance (or *Lorentz*) profile of the absorption "line."

Exercise 19.14. Show that in the limit $\Gamma \to 0$ the distribution (19.133) becomes a delta function and the cross section (19.134) approaches the form (19.83).

We saw that exponential decay of a discrete initial state embedded in a quasi-continuum of final states presupposes that the system is allowed to evolve without an intervening observation of its state at least during a brief time interval $\hbar/\Delta E$, which is of the order of a typical period associated with the system. On the other hand, if this assumption is violated and the state of the system is observed within a much shorter time interval, we can no longer expect that its time development will be unaffected. Here we consider the extreme possibility that a system which is initially represented by a state vector $\Psi(0)$ and which under the action of a time-

independent Hamiltonian H evolves into $\Psi(t)$, in the Schrödinger picture, is subjected to frequently repeated observations to ascertain whether or not it is still in its initial state. In other words, we contemplate a measurement of the projection operator, or density operator, $\rho_0 = |\Psi(0)\rangle\langle\Psi(0)|$. This operator has eigenvalues 1 and 0, corresponding to "yes" and "no" as the possible answers to the experimental interrogation that is designed to determine if the system has survived in the initial state. The observations are assumed to be *ideal* measurements leaving the system in state $\Psi(0)$ if the answer is "yes," corresponding to eigenvalue 1. The measurement thus resets the clock, and $\Psi(0)$ serves again as the initial state of the subsequent time evolution. If N such repeated measurements are made in the time interval t, the probability that at *every one* of the interrogations the system is found to have survived in its initial state is

$$\begin{aligned} P_N(t) &= |\langle\Psi(0)|\Psi(t/N)\rangle|^{2N} \\ &= |\langle\Psi(0)|T^\dagger(t/N, 0)|\Psi(0)\rangle\langle\Psi(0)|T(t/N, 0)|\Psi(0)\rangle|^N \end{aligned} \tag{19.135}$$

Since we are interested in the limit $N \to \infty$, it is appropriate to expand the time development operator in powers of t/N

$$T\left(\frac{t}{N}, 0\right) = 1 - \frac{i}{\hbar}\frac{t}{N}H - \frac{1}{2\hbar^2}\left(\frac{t}{N}\right)^2 H^2 + \ldots \tag{19.136}$$

Substituting this approximation into (19.135), we obtain

$$\begin{aligned} P_N(t) &= |1 + \langle\Psi(0)|H|\Psi(0)\rangle^2(t/N)^2 - \langle\Psi(0)|H^2|\Psi(0)\rangle(t/N)^2|^N \\ &= [1 - (\Delta H)_0^2(t/N)^2]^N \end{aligned} \tag{19.137}$$

In the limit $N \to \infty$ the system is ceaselessly observed to check if it is still represented by its initial state. The probability for finding it so is

$$\lim_{N\to\infty} P_N(t) = \lim_{N\to\infty}\left[1 - (\Delta H)_0^2\left(\frac{t}{N}\right)^2\right]^N = e^0 = 1 \tag{19.138}$$

and we conclude that under these circumstances the system is never observed to change at all. This peculiar result is known as *Zeno's paradox* in quantum mechanics (also as *Turing's paradox*). Experiments have shown that it is indeed possible to delay the decay of an unstable system by subjecting an excited atomic state to intermittent observations with a high recurrence rate. This result was derived for general dynamical systems, without the use of perturbation theory.

Exercise 19.15. A spin one-half particle with magnetic moment is exposed to a static magnetic field. If the state were continuously observed in an ideal measurement, show that the polarization vector would not precess.

Problems

1. Calculate the cross section for the emission of a photoelectron ejected when linearly polarized monochromatic light of frequency ω is incident on a complex atom. Simulate the initial state of the atomic electron by the ground state wave function of an isotropic three-dimensional harmonic oscillator and the final state by a plane wave. Obtain the angular distribution as a function of the angle of emission and sketch it on a polar graph for suitable assumed values of the parameters.

2. Calculate the total cross section for photoemission from the K shell as a function of the frequency of the incident light and the frequency of the K-shell absorption edge, assuming that $\hbar\omega$ is much larger than the ionization potential but that nevertheless the photon momentum is much less than the momentum of the ejected electron. Use a hydrogenic wave function for the K shell and plane waves for the continuum states.
3. By considering the double commutator

$$[[H, e^{i\mathbf{k}\cdot\mathbf{r}}], e^{-i\mathbf{k}\cdot\mathbf{r}}]$$

obtain as a generalization of the Thomas-Reiche-Kuhn sum rule the formula

$$\sum_n (E_n - E_s)|\langle n|e^{i\mathbf{k}\cdot\mathbf{r}}|s\rangle|^2 = \frac{\hbar^2 k^2}{2m}$$

Specify the conditions on the Hamiltonian H required for the validity of this sum rule.

4. A charged particle moving in a linear harmonic oscillator potential is exposed to electromagnetic radiation. Initially, the particle is in the oscillator ground state. Discuss the conditions under which the electric dipole–no retardation approximation is good. In this approximation, show that the first-order perturbation value of the integrated absorption cross section is equal to the sum of dipole absorption cross sections, calculated exactly.
5. For the system described in Problem 4, derive the selection rules for transitions in the electric quadrupole approximation, which correspond to retaining the second term in the expansion (19.64). Calculate the absorption rate for quadrupole transitions and compare with the rate for dipole transitions.

CHAPTER 20

The Formal Theory of Scattering

It is natural to look at a scattering process as a transition from one unperturbed state to another. In the formal theory of scattering, infinite plane waves serve as idealizations of very broad and long wave packets, replacing the formulation of Chapter 13 in terms of finite wave packets. Although the formal theory is patterned after the description of simple elastic deflection of particles by a fixed potential, it is capable of enormous generalization and applicable to collisions between complex systems, inelastic collisions, nuclear reactions, processes involving photons (such as the Compton effect or pair production), collisions between pions and nucleons, and so forth. Almost every laboratory experiment of atomic, nuclear, and particle physics can be described as a generalized scattering process with an initial incident state, an interaction between the components of the system, and a final scattered state.[1] Formal scattering theory enables us to predict the general form of observable quantities, such as transition probabilities or cross sections, quickly and directly from the symmetry properties of a system, and it is readily adapted to a relativistic formulation. The scattering matrix, already introduced in Chapters 6 and 16, contains all relevant dynamical information. This chapter centers on a generalization of this concept, the scattering operator.

1. ***The Equations of Motion, the Transition Matrix, the S Matrix, and the Cross Section.*** A collision can be thought of as a transition between two unperturbed states. If the scattering region is of finite extent, the initial and final states are simply plane wave eigenstates of definite momentum of the unperturbed Hamiltonian, $H_0 = \mathbf{p}^2/2m$, and the scattering potential causes transitions from an initial state with propagation vector $\mathbf{k}$ to the final states characterized by propagation vectors $\mathbf{k}'$.

At first sight, it may seem strange that an incident wave can be represented by an infinite plane wave that is equally as intense in front of the scatterer as behind it. In Chapter 13, the unphysical appearance of incident waves behind the scatterer was avoided by superposing waves of different $\mathbf{k}$ and canceling the unwanted portion of the wave by destructive interference. The resulting theory, involving wave packets and Fourier integrals at every step, was correct but clumsy. In the present chapter, precisely the same results will be achieved in a more elegant fashion by the use of suitable mathematical limiting procedures. To avoid mistakes, however, it is advisable always to keep in mind the physical picture of the scattering of particles, which first impinge upon and subsequently move away from the scatterer.

In scattering problems, we are interested in calculating transition amplitudes

[1]A comprehensive treatise on scattering processes is Goldberger and Watson (1964). See also Newton (1982). A fine textbook is Taylor (1972).

between states that evolve in time under the action of an unperturbed Hamiltonian H_0 whose eigenstates are defined by

$$H_0|n\rangle = E_n|n\rangle \tag{20.1}$$

As in Chapter 13, nonrelativistic elastic scattering of a particle without spin from a fixed potential will be considered, since this process, to which we refer as *simple scattering*, is the prototype of all more complex processes. The Hamiltonian is

$$H = H_0 + V = \frac{\mathbf{p}^2}{2m} + V \tag{20.2}$$

Although we assume that H_0 is simply the kinetic energy operator, in more sophisticated applications H_0 may include part of the interaction. The interaction operator V is assumed to be time-independent.

In the interaction picture, the solution of the equation of motion may be written in terms of the eigenvectors of H_0 as

$$|\tilde{\Psi}(t)\rangle = \sum_{k,s} |k\rangle\langle k|\tilde{T}(t, t_0)|s\rangle\langle s|\tilde{\Psi}(t_0)\rangle \tag{20.3}$$

According to (19.15), the equation of motion for the transition amplitudes is expressible as

$$\langle k|\tilde{T}(t, t_0)|s\rangle = \delta_{ks} - \frac{i}{\hbar}\sum_n \langle k|V|n\rangle \int_{t_0}^{t} e^{i\omega_{kn}t'}\langle n|\tilde{T}(t', t_0)|s\rangle\, dt' \tag{20.4}$$

where

$$\hbar\omega_{kn} = E_k - E_n \tag{20.5}$$

and where it has been assumed that the unperturbed eigenvectors are normalized to unity:

$$\langle k|n\rangle = \delta_{kn} \tag{20.6}$$

This means that, if the unperturbed states are plane waves, periodic boundary conditions must be imposed in a large box of volume L^3. Eventually, the limit $L \to \infty$ may be taken. Typically, in simple scattering of spinless particles, the unperturbed states of interest are plane wave states $e^{i\mathbf{k}\cdot\mathbf{r}}/L^{3/2}$ in the coordinate representation, where $\mathbf{k}$ stands for the incident or scattered propagation vector, quantized as in Section 4.4 to satisfy the periodic boundary conditions.

To describe scattering properly, we should choose the initial state $|\tilde{\Psi}(t_0)\rangle$ in (20.3) to represent a wave packet moving freely toward the interaction region. Here we employ a mental shortcut and idealize the wave packet as an incident plane wave, making the unrealistic assumption that somehow at time t_0 the scattering region has been embedded in a perfect plane wave that is now being released. By pushing the initial time into the distant past and letting $t_0 \to -\infty$, we avoid the unphysical consequences of this assumption, which brings confusing transients into the calculation. Similarly, $t \to +\infty$ signals that the scattering process is complete. We thus demand that the transition matrix element between an incident, or "*in*," state s and a scattered, or "*out*," state k, $\langle k|\tilde{T}(t, t_0)|s\rangle$, converges to a well-defined limit as $t_0 \to -\infty$ and $t \to +\infty$. Scattering theory requires us to solve the coupled equations (20.4) subject to the condition of the existence of the limit

$$\boxed{S_{ks} = \langle k|\tilde{T}(+\infty, -\infty)|s\rangle} \tag{20.7}$$

for the matrix elements of the time development operator between unperturbed asymptotic in and out states. This matrix is called the *scattering matrix*, or simply the *S matrix*; it plays a central role in this chapter. We encountered it in special cases earlier, in Chapters 6, 13, and 16.

To solve the equations for the transition matrix elements, we recall that in first-order perturbation theory we would write

$$\langle k|\tilde{T}(t, t_0)|s\rangle = \delta_{ks} - \frac{i}{\hbar}\langle k|V|s\rangle \int_{t_0}^{t} e^{i\omega_{ks}t'}\, dt' \tag{20.8}$$

Using this expression as a clue, we devise the *Ansatz*:

$$\langle k|\tilde{T}(t, t_0)|s\rangle = \delta_{ks} - \frac{i}{\hbar} T_{ks} \int_{t_0}^{t} e^{i\omega_{ks}t' + \alpha t'}\, dt' \tag{20.9}$$

In generalizing (20.8) to (20.9), the known matrix $\langle k|V|s\rangle$ has been replaced by an unknown matrix T_{ks} in the expectation that the perturbation approximation might be avoided. To make the integral meaningful as $t_0 \to -\infty$, a factor $e^{\alpha t}$ has been inserted in the integrand with the understanding that α is positive and that the limit $\alpha \to 0$ must be taken after the limit $t_0 \to -\infty$. Equation (20.9) will be assumed to give $\langle k|\tilde{T}(t, t_0)|s\rangle$ correctly only for times t that satisfy the relation

$$|t| \ll (1/\alpha) \tag{20.10}$$

It is essential to keep these restrictions in mind. If they are disregarded, the equations of the formal theory may lead to painful contradictions. Such contradictions easily arise because the formal theory is designed to be a shorthand notation in which conditions like (20.10) are implied but never spelled out. The formal theory thus operates with a set of conventions from which it derives its conciseness and flexibility. Those who consider the absence of explicit mathematical instructions at every step too high a price to pay can always return to the wave packet form of the theory. The connection between the two points of view is never lost if it is noted that $1/\alpha$ measures crudely the length of time during which the wave packet strikes, envelops, and passes the scattering region. If v is the mean particle velocity, v/α is roughly the length of the wave packet.

Having given some motivation for the form (20.9), we now ask if the T_{ks}, known as the *transition matrix elements*, can be determined so that (20.9) is the solution of (20.4). In the next section, we will show that under the conditions prevailing in scattering problems such a solution does indeed exist. Moreover, it is rigorous and not approximate.

Assuming the existence of this solution, we can draw an important conclusion immediately. Upon integrating (20.9), we obtain

$$\langle k|\tilde{T}(t, -\infty)|s\rangle = \delta_{ks} + \frac{T_{ks}e^{i\omega_{ks}t + \alpha t}}{\hbar(-\omega_{ks} + i\alpha)} \tag{20.11}$$

as $\lim_{\substack{\alpha\to 0 \\ t_0\to -\infty}} e^{\alpha t_0} = 0$. In the limit $t \to +\infty$ and $\alpha \to 0$, but subject to condition (20.10), and using Eq. (A.17) in the Appendix, we obtain:

$$\boxed{S_{ks} = \langle k|\tilde{T}(+\infty, -\infty)|s\rangle = \delta_{ks} - 2\pi i\delta(E_k - E_s)T_{ks}} \tag{20.12}$$

This formula provides an important connection between the S matrix and the transition matrix. Kronecker deltas and delta functions are mixed together in this expression, because at this stage we don't need to commit ourselves to any particular

choice of representation for the states labeled by s and k. If that choice is eventually made in favor of the propagation vectors $\mathbf{k}$ (the momentum of units of $\hbar$) and if $\mathbf{k}$-normalization is used as $L \to \infty$, the δ_{ks} turns into $\delta(\mathbf{k}' - \mathbf{k})$.

For states $k \neq s$, we thus have at finite times, during the scattering process

$$|\langle k|\tilde{T}(t, -\infty)|s\rangle|^2 = \frac{|T_{ks}|^2 e^{2\alpha t}}{\hbar^2(\omega_{ks}^2 + \alpha^2)} \tag{20.13}$$

Hence, for the rate of transition into state k,

$$\frac{d}{dt}|\langle k|\tilde{T}(t, -\infty)|s\rangle|^2 = \frac{2\alpha}{\omega_{ks}^2 + \alpha^2} e^{2\alpha t} \frac{1}{\hbar^2}|T_{ks}|^2 \tag{20.14}$$

In the limit $\alpha \to 0$, which must always be taken but at finite values of t, this becomes

$$\frac{d}{dt}|\langle k|\tilde{T}(t, -\infty|s\rangle|^2 = \frac{2\pi}{\hbar^2}\delta(\omega_{ks})|T_{ks}|^2 = \frac{2\pi}{\hbar}\delta(E_k - E_s)|T_{ks}|^2 \tag{20.15}$$

if $k \neq s$. The solution thus implies a constant transition rate—precisely what we expect to be the effect of the scatterer causing transitions from state s to k. This makes it clear why T_{ks} is called the transition matrix.

Equation (20.15) is meaningful only if there is a quasi-continuum of unperturbed states with energies $E_k \approx E_s$. In Section 20.2, we will demonstrate that the matrix T_{ks} exists and that in scattering problems (20.4) has solutions of the form (20.9).

If the theory is to be useful, we must establish the connection between the transition rate, (20.15), and the scattering cross section. The unperturbed states are now assumed to be normalized momentum eigenstates. Since

$$\delta(E_k - E_{k'}) = \frac{m}{\hbar^2 k}\delta(k - k')$$

we obtain for the total transition rate from an incident momentum state $\mathbf{k}$ into a solid angle $d\Omega$,

$$\begin{aligned} w &= \sum_{\mathbf{k}'} \frac{d}{dt}|\langle \mathbf{k}'|\tilde{T}(t, -\infty)|\mathbf{k}\rangle|^2 = \frac{2\pi}{\hbar} d\Omega \int_0^\infty \frac{m}{\hbar^2 k}\delta(k - k')|T_{\mathbf{k}'\mathbf{k}}|^2 k'^2 \frac{L^3}{(2\pi)^3} dk' \\ &= \frac{mkL^3}{(2\pi)^2\hbar^3}|T_{\mathbf{k}'\mathbf{k}}|^2 \, d\Omega \end{aligned} \tag{20.16}$$

where $\mathbf{k}'$ is the momentum of the scattered particle ($k' = k$). The factor $(L/2\pi)^3$ is the $\mathbf{k}$-space density of free-particle states in the cube of length L, subject to periodic boundary conditions [see Eq. (4.53)]. Hence, if $v = \hbar k/m$ is the velocity of the incident particles, the transition rate reduces to

$$w = \frac{m^2 v L^3}{(2\pi)^2\hbar^4}|T_{\mathbf{k}'\mathbf{k}}|^2 \, d\Omega \tag{20.17}$$

The probability of finding a particle in a unit volume of the incident beam is $1/L^3$. Hence, v/L^3 is the probability that a particle is incident on a unit area perpendicular to the beam per unit time. If this probability current density is multiplied by the differential cross section $d\sigma$, as defined in Section 13.1, the transition rate w is obtained; hence,

$$\boxed{d\sigma = \frac{w}{v/L^3} = \left(\frac{mL^3}{2\pi\hbar^2}\right)^2 |T_{\mathbf{k}'\mathbf{k}}|^2 \, d\Omega} \tag{20.18}$$

If $\langle k|V|s\rangle$ is used as an approximation instead of T_{ks}, (20.15) is equivalent to the Golden Rule (19.99) of time-dependent perturbation theory, and the cross-section formula (20.18) reduces to the Born approximation of Section 13.4.

2. *The Integral Equations of Scattering Theory.* We now substitute $\langle k|\tilde{T}(t, -\infty)|s\rangle$ from (20.11) into (20.4) and immediately set $\alpha t = 0$ in accordance with the restriction (20.10). The matrix T_{ks} must then satisfy the system of simultaneous linear equations,

$$\boxed{T_{ks} = \langle k|V|s\rangle + \frac{1}{\hbar}\sum_n \frac{\langle k|V|n\rangle T_{ns}}{-\omega_{ns} + i\alpha}} \tag{20.19}$$

If the transition matrix elements satisfy this equation, the expression (20.11) is a solution of the equation of motion (20.4) for times $|t| \ll 1/\alpha$.

The summation in (20.19) extends over all unperturbed eigenstates. Hence, it includes states for which $E_n = E_s$. The corresponding denominator, $-\omega_{ns} + i\alpha$, vanishes in the limit $\alpha \to 0$, and we would thus be led to a meaningless equation if the spectrum of H_0 were truly discrete. The unmanageable singularity can be averted, and the transition matrix can be expected to exist if the unperturbed states form a quasi-continuum with energies very close to the energy E_s of the initial state. Precisely this situation prevails in scattering, where the initial and final states have the same, or very nearly the same, energies. The transition amplitudes and the matrix elements T_{ks} are then proportional to $1/L^3$. Hence, the amplitudes of the final states, $k \neq s$, will remain small, at least for $t \ll 1/\alpha$. It follows that the initial state is not appreciably depleted over a period of the order $1/\alpha$, and we can have a *constant rate of transition*. If the initial state, instead of being part of a quasi-continuum of energies were a truly discrete state, the method of solution outlined here would fail. Such a state decays exponentially and, as was seen in Section 19.8, an altogether different solution to the equation of motion is appropriate.

Equation (20.19) can be used to connect the formal theory with the more explicit description of scattering in Chapter 13. For this purpose, it is convenient to define a set of vectors $\Psi_s^{(+)}$ in Hilbert space by the linear equations,

$$T_{ks} = \sum_j (\Psi_k, V\Psi_j)(\Psi_j, \Psi_s^{(+)}) = (\Psi_k, V\Psi_s^{(+)}) \tag{20.20}$$

Substituting this scalar product in (20.19), we obtain an equation that $\Psi_s^{(+)}$ must satisfy:

$$(\Psi_k, V\Psi_s^{(+)}) = \sum_n (\Psi_k, V\Psi_n)\frac{(\Psi_n, V\Psi_s^{(+)})}{E_s - E_n + i\hbar\alpha} + (\Psi_k, V\Psi_s)$$

or, since this must be true for all k

$$\begin{aligned}\Psi_s^{(+)} &= \Psi_s + \sum_n \Psi_n \frac{(\Psi_n, V\Psi_s^{(+)})}{E_s - E_n + i\hbar\alpha} \\ &= \Psi_n + \sum_n \frac{1}{E_s - H_0 + i\hbar\alpha}\Psi_n(\Psi_n, V\Psi_s^{(+)})\end{aligned} \tag{20.21}$$

However, by completeness,

$$\sum_n \Psi_n(\Psi_n, V\Psi_s^{(+)}) = V\Psi_s^{(+)} \tag{20.22}$$

Hence, we obtain as a final result the *implicit equation*

$$\boxed{\Psi_s^{(+)} = \Psi_s + \frac{1}{E_s - H_0 + i\hbar\alpha} V\Psi_s^{(+)}} \tag{20.23}$$

This is a fundamental equation of the formal theory of scattering. The problem of obtaining the transition matrix has thus been reduced to solving (20.23), known as the *Lippmann-Schwinger equation.*

By applying the operator, $E_s - H_0 + i\hbar\alpha$ to (20.23), it is immediately established that in the limit $\alpha \to 0$,

$$(E_s - H_0)\Psi_s^{(+)} = V\Psi_s^{(+)} \tag{20.24}$$

That is, $\Psi_s^{(+)}$ becomes an eigenvector of $H = H_0 + V$, and E_s the corresponding eigenvalue. Hence, the procedure of solving the equation of motion outlined in this chapter can succeed only if, in the limit $\alpha \to 0$, E_s is an eigenvalue of *both* H_0 and H. In simple scattering systems where $V \to 0$ as $r \to \infty$, both Hamiltonian operators have the same continuous spectrum: The energy can have any value between 0 and ∞, and the presence of the potential does not change this fact. Throughout this chapter it will be assumed that the continuous portions of the spectra of H_0 and H coincide and extend from $E = 0$ to ∞, although for more complicated systems this is a severe restriction that cannot and need not be maintained. In addition, H may possess discrete eigenvalues that have no counterpart in the spectrum of H_0, corresponding to bound states.

Actually, the periodic boundary conditions on the cube of length L cause the entire spectra of both H_0 and H to be discrete. Even the quasi-continuous energy levels of H_0 and H will then generally not coincide, because the two operators differ by V. But the shift, ΔE_k, of each level goes to zero as $L \to \infty$. The limit $L \to \infty$ must be taken before the limit $\alpha \to 0$, since v/α signifies approximately the length of the incident wave packet and we must require that $L > v/\alpha$, or else the normalization box could not contain the wave packet.[2] If the limit $L \to \infty$ is accompanied by a change of normalization of positive energy states from unity to $\mathbf{k}$-normalization, as defined in Section 4.4, the formal identification $L = 2\pi$ allows the same equations to be used after the limit has been taken as before.

In retrospect, it is instructive to rederive (20.23) by making the eigenvalue equation (20.24) the starting point, and to ask for a solution of the eigenvalue problem of H, assuming that the fixed eigenvalue E_s lies in the continuous spectrum that is common to both H and H_0. Equation (20.24) may then be regarded as an inhomogeneous equation to be solved. Its homogeneous counterpart,

$$(H_0 - E_s)\Psi_s \equiv (H_0 - E_s)|s\rangle = 0 \tag{20.25}$$

does have solutions (by assumption). According to the theorems of Section 18.2 the inhomogeneous equation has a solution if and only if the inhomogeneous term $V\Psi_s^{(+)}$ is orthogonal to the subspace spanned by the solution of the homogeneous equation (20.25). This condition is generally *not* satisfied. We can nevertheless ignore this apparent obstacle in solving (20.24), because, when the spectrum is quasi-

[2]In the formal scattering theory, the order in which the several limiting processes are executed is very important. No simple mathematical notation has yet been invented for operators that depend on parameters whose limiting values are to be taken after the operation.

continuous, the component of $V\Psi_s^{(+)}$ in the troublesome subspace can be neglected without any appreciable alteration.

According to Section 18.2, the general "solution" of the inhomogeneous equation (20.24) is

$$\Psi_s^{(+)} = \Psi_s + G_+(E_s)V\Psi_s^{(+)} \tag{20.26}$$

where $G_+(E_s)$ is a particular solution of the operator equation

$$(E_s - H_0)G(E_s) = 1 - P_s \tag{20.27}$$

This equation has infinitely many different solutions. In order to obtain the solutions $\Psi_s^{(+)}$ we must obviously make the choice

$$G_+(E_s) = \frac{1}{E_s - H_0 + i\hbar\alpha} \tag{20.28}$$

since then (20.23) and (20.26) become identical. The operator (20.28) is indeed a particular solution of (20.27), since $(E_s - H_0)G_+(E_s)$ gives zero when it acts on an eigenvector of H_0 with eigenvalue E_s and, in the limit $\alpha \to 0$, has the effect of the identity operation when it acts on an eigenvector Ψ_r of H_0 with $E_r \neq E_s$.

The implicit equation (20.23) for $\Psi_s^{(+)}$ looks like the integral equation (13.28), and the two equations are equivalent if the unperturbed eigenstates Ψ_s are momentum eigenstates. To prove this, it is best to go back to (20.21) and write it in the coordinate representation. If we set

$$\langle \mathbf{r} | \Psi_s^{(+)} \rangle = \psi_{\mathbf{k}}^{(+)}(\mathbf{r}) \tag{20.29}$$

and

$$\langle \mathbf{r} | \Psi_s \rangle = \frac{1}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{r}}$$

we obtain for the local potential $\langle \mathbf{r}' | V | \mathbf{r}'' \rangle = V(\mathbf{r})\delta(\mathbf{r}' - \mathbf{r}'')$,

$$\psi_{\mathbf{k}}^{(+)}(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} e^{i\mathbf{k}\cdot\mathbf{r}} - \frac{1}{(2\pi)^3} \iint \frac{e^{i\mathbf{k}'\cdot(\mathbf{r}-\mathbf{r}')}}{k'^2 - (k^2 + i\varepsilon)} d^3k' \frac{2m}{\hbar^2} V(\mathbf{r}')\, \psi_{\mathbf{k}}^{(+)}(\mathbf{r}')\, d^3r' \tag{20.30}$$

By comparison with the equations of Section 13.3, we see that after the $\mathbf{k}'$-integration is performed, this equation becomes identical to the integral equation (13.39).

The preceding discussion shows that the Green's function $G_+(\mathbf{r}, \mathbf{r}')$ is proportional to the matrix element of the operator

$$\boxed{G_+(E) = (E - H_0 + i\hbar\alpha)^{-1}}$$

in the coordinate representation:

$$G_+(\mathbf{r}, \mathbf{r}') = -4\pi \frac{\hbar^2}{2m} \langle \mathbf{r} | G_+(E) | \mathbf{r}' \rangle \tag{20.31}$$

As shown in Chapter 13, this particular Green's function ensures that the solutions $\psi^{(+)}$ of the integral equation (13.28) asymptotically represent *outgoing* spherical waves in addition to incident plane waves.

Exercise 20.1. Work out the Green's function $\langle x|G_+(E)|x'\rangle$ for a free particle in one dimension.

The operator G_+ is only one particular solution of (20.24). The particular solution

$$\boxed{G_-(E_s) = \frac{1}{E_s - H_0 - i\hbar\alpha}} \tag{20.32}$$

of (20.27) leads to the replacement of (20.26) by the equation

$$\Psi_s^{(-)} = \Psi_s + G_-(E_s)V\Psi_s^{(-)} \tag{20.33}$$

If Ψ_s is a momentum eigenvector, the eigenstate $\Psi_s^{(-)}$ of H, when projected into the coordinate representation, describes asymptotically spherically *incoming* waves.

Another solution of (20.27) which is often useful is the operator

$$G_1(E_s) = \frac{1}{2}G_+(E_s) + \frac{1}{2}G_-(E_s) = \frac{1}{2}\left(\frac{1}{E_s - H_0 + i\hbar\alpha} + \frac{1}{E_s - H_0 - i\hbar\alpha}\right)$$

The expression on the right may be simplified and written as [Eq. (A.20) in the Appendix]:

$$\boxed{G_1(E) = \mathsf{P}\left(\frac{1}{E - H_0}\right)} \tag{20.34}$$

With this operator, we can define a set of eigenstates $\Psi^{(1)}$ by the equation

$$\Psi^{(1)} = \Psi_s + \mathsf{P}\left(\frac{1}{E - H_0}\right)V\Psi_s^{(1)} \tag{20.35}$$

In the coordinate representation, the operator $G_1(E)$ is related to the *standing wave* Green's function (13.38).

The solution of equations like (20.26), (20.33), and (20.35) depends on the choice of the free particle states Ψ_s. To guard against confusion with commonly used terminology, we note that, although for simple scattering the states $|\Psi_s^{(+)}\rangle$ in the coordinate representation represent asymptotically *outgoing* spherically scattered waves, they correspond to the "in" states of particles with momentum $\hbar\mathbf{k}$, before they reach the scattering region. Similarly, the states $|\Psi_k^{(-)}\rangle$ represent *incoming* spherically scattered waves, but correspond to the "out" states of particles with momentum $\hbar\mathbf{k}'$, after the scattering has taken place.

If we identify the unperturbed initial and final states as the momentum eigenstates, $|\mathbf{k}\rangle$ and $|\mathbf{k}'\rangle$, and make the proper adjustment for box normalization, the transition matrix can, by (20.20), be expressed as

$$T_{\mathbf{k}'\mathbf{k}} = (\Psi_{\mathbf{k}'}, V\Psi_{\mathbf{k}}^{(+)}) = \frac{(2\pi)^{3/2}}{L^3}\int e^{-i\mathbf{k}'\cdot\mathbf{r}}\, V(\mathbf{r})\psi_{\mathbf{k}}^{(+)}(\mathbf{r})\, d^3r \tag{20.36}$$

Comparing this with (13.43), we find a simple relation between the *transition matrix* and the *scattering amplitude*:

$$\boxed{T_{\mathbf{k}'\mathbf{k}} = -\frac{2\pi\hbar^2}{mL^3}\, f_{\mathbf{k}}(\hat{\mathbf{k}}')} \tag{20.37}$$

In the limit $L \to \infty$, and with **k**-normalization, this equation turns into

$$-(2\pi)^2 \, mT_{\mathbf{k}'\mathbf{k}} = \hbar^2 f_{\mathbf{k}}(\mathbf{k}') \tag{20.38}$$

Substituting $T_{\mathbf{k}'\mathbf{k}}$ from (20.37) into (20.18), we obtain

$$\boxed{d\sigma = |f_{\mathbf{k}}(\hat{\mathbf{k}}')|^2 \, d\Omega} \tag{20.39}$$

This result is identical with (13.23).

3. ***Properties of the Scattering States.*** The fundamental problem of scattering theory is to solve the equation

$$\Psi_s^{(+)} = \Psi_s + \frac{1}{E_s - H_0 + i\hbar\alpha} V\Psi_s^{(+)} \tag{20.23}$$

The solutions can then be used to determine the transition matrix which, according to the last section, is directly related to the cross section.

Formally, we may solve (20.23) by multiplying it by $E_s - H_0 + i\hbar\alpha$, and adding and subtracting $-V\Psi_s$ on the right-hand side of the equation. Thus, we obtain

$$(E_s - H + i\hbar\alpha)\Psi_s^{(+)} = (E_s - H + i\hbar\alpha)\Psi_s + V\Psi_s$$

or

$$\Psi_s^{(+)} = \Psi_s + \frac{1}{E_s - H + i\hbar\alpha} V\Psi_s \tag{20.40}$$

The important distinction between this equation and (20.23) is the appearance of H rather than H_0 in the denominator.

If the solution (20.40) is substituted in (20.20) for the transition matrix, we get

$$T_{ks} = (\Psi_k, V\Psi_s) + \left(\Psi_k, V\frac{1}{E_s - H + i\hbar\alpha} V\Psi_s\right) \tag{20.41}$$

In this way, the cross section for a scattering process can in principle be calculated. However, for practical purposes not much is gained, because the effect of the operator $(E_s - H + i\hbar\alpha)^{-1}$ is not known unless the eigenvectors of H have already been determined. Since this is the problem we want to solve, it is usually necessary to resort to approximation methods to solve (20.23).

Exercise 20.2. Show that, if $E_r = E_s$,

$$T_{rs} = (\Psi_r^{(-)}, V\Psi_s) \tag{20.42}$$

The crudest approximation is obtained if in (20.23) the term proportional to V on the right-hand side of the equation is neglected altogether:

$$\Psi_s^{(+)} \cong \Psi_s \tag{20.43}$$

This is simply the first term in the solution of (20.23) obtained by successive approximation. If we define the *resolvent operator*,

$$\mathcal{G}_+(E) = \frac{1}{E - H + i\hbar\alpha} \tag{20.44}$$

Eq. (20.40) can be written as

$$\Psi_s^{(+)} = \Psi_s + \mathcal{G}_+(E_s)V\Psi_s \tag{20.45}$$

Using the identity,

$$\boxed{\mathcal{G}_+(E) = G_+(E) + G_+(E)V\mathcal{G}_+(E)} \tag{20.46}$$

we find by iteration the operator identity,

$$\mathcal{G}_+ = [\sum_{k=0}^{n} (G_+V)^k]G_+ + (G_+V)^{n+1}\mathcal{G}_+ \tag{20.47}$$

Exercise 20.3. Prove the operator identity

$$\boxed{\frac{1}{A} = \frac{1}{B} + \frac{1}{B}(B - A)\frac{1}{A}} \tag{20.48}$$

and exploit it to verify (20.46).

The expansion (20.47) applied to Eq. (20.40) produces the formal series expansion

$$\boxed{\Psi_s^{(+)} = \Psi_s + G_+(E_s)V\Psi_s + G_+(E_s)VG_+(E_s)V\Psi_s + \cdots} \tag{20.49}$$

The nth *Born approximation* to the scattering state $\Psi_s^{(+)}$ consists of terminating the expansion (20.49) arbitrarily after n terms.

Exercise 20.4. Show that formally the series (20.49) is also arrived at by rewriting (20.26) as

$$\Psi_s^{(+)} = \frac{1}{1 - G_+(E_s)V}\Psi_s \tag{20.50}$$

and expanding $(1 - G_+V)^{-1}$ as a power series.

The convergence of the *Born series* (20.49) is often difficult to ascertain, but it is easy to see that it will certainly *not* converge, if the equation

$$\lambda_0 G_+(E_s)V\Psi = \Psi$$

has an eigenvalue λ_0 whose absolute value is less than 1. The operator $(1 - \lambda G_+V)^{-1}$ has a singularity at $\lambda = \lambda_0$; consequently, the radius of convergence of the series expansion of this operator in powers of λ must be less than $|\lambda_0|$. If $|\lambda_0| < 1$, the Born series, which corresponds to $\lambda = 1$, is divergent. If, as frequently happens in cases of practical interest, the Born series fails to converge or converges too slowly to be useful, more powerful, but also more involved, approximation techniques are available for the determination of $\Psi^{(+)}$.

If the first Born approximation (20.43) is substituted into the transition matrix, we obtain from (20.36) and (20.37):

$$f_{\text{Born}} = -\frac{m}{2\pi\hbar^2}\int e^{-i\mathbf{k}'\cdot\mathbf{r}}\, Ve^{i\mathbf{k}\cdot\mathbf{r}}\, d^3r = -\frac{mL^3}{2\pi\hbar^2}\langle\mathbf{k}'|V|\mathbf{k}\rangle \tag{20.51}$$

in agreement with (13.48). The first Born approximation is the result of a first-order perturbation treatment of scattering, in which the accurate equation (20.9) is replaced by the approximate equation (20.8).

The formal solution (20.40) can be used to demonstrate the orthonormality of the eigenvectors $\Psi^{(+)}$. This is seen by the following simple manipulations:

$$\begin{aligned}(\Psi_k^{(+)}, \Psi_s^{(+)}) &= \left(\Psi_k + \frac{1}{E_k - H + i\hbar\alpha} V\Psi_k, \Psi_s^{(+)}\right) \\ &= \left(\Psi_k, \Psi_s^{(+)} + V\frac{1}{E_k - H - i\hbar\alpha}\Psi_s^{(+)}\right) \\ &= \left(\Psi_k, \Psi_s^{(+)} + V\frac{1}{E_k - E_s - i\hbar\alpha}\Psi_s^{(+)}\right) \\ &= \left(\Psi_k, \Psi_s^{(+)} - \frac{1}{E_s - H_0 + i\hbar\alpha} V\Psi_s^{(+)}\right)\end{aligned}$$

If we finally use (20.23), we get the result

$$(\Psi_k^{(+)}, \Psi_s^{(+)}) = (\Psi_k, \Psi_s) = \delta_{ks} \tag{20.52}$$

This formula is valid only in the limit $L \to \infty$, when E_r becomes an eigenvalue of both H and H_0. Entirely analogous arguments can be made to show that

$$(\Psi_k^{(-)}, \Psi_s^{(-)}) = \delta_{ks} \tag{20.53}$$

Corresponding to an orthonormal set of Ψ_s we thus obtain two sets, $\Psi_s^{(+)}$ and $\Psi_s^{(-)}$, of orthonormal eigenvectors of the total Hamiltonian H. The question that arises is whether these sets are complete. It would appear that each set by itself is a complete set, because the vectors Ψ_s form a complete set, and $\Psi_s^{(+)}$ (or $\Psi_s^{(-)}$) goes over into Ψ_s as $V \to 0$. However, one reservation is called for: The Hamiltonian H may have discrete energy eigenvalues corresponding to bound states produced by the interaction V. These discrete states, which have no counterpart in the spectrum of H_0 and are never found among the solutions of (20.20), are orthogonal to the scattering states and must be added to all the $\Psi_s^{(+)}$ (or $\Psi_s^{(-)}$) to complete the set of eigenvectors.

4. ***Properties of the Scattering Matrix.*** The S matrix was defined in terms of time-dependent transition amplitudes in Section 20.1. It is related to the transition matrix, and therefore to the energy eigenstates of H and H_0, by Eqs. (20.12) and (20.20), leading to:

$$S_{ks} = \delta_{ks} - 2\pi i\delta(E_k - E_s)T_{ks} = \delta_{ks} - 2\pi i\delta(E_k - E_s)(\Psi_k, V\Psi_s^{(+)}) \tag{20.54}$$

The S matrix connects the initial state with the final states of interest. An alternative approach to the S matrix is to think of scattering as described by idealized stationary states. If, owing to its preparation, the system is known to be in the eigenstate $\Psi_s^{(+)}$ of the full Hamiltonian H, the S matrix element S_{ks} is the probability amplitude for detecting the system in state $\Psi_k^{(-)}$. Therefore, when we express $\Psi_s^{(+)}$ as a linear combination of the $\Psi_k^{(-)}$, all belonging to the same energy eigenvalue,

$$\Psi_k^{(+)} = \sum_k \Psi_k^{(-)} S_{ks} \tag{20.55}$$

we expect that the expansion coefficients S_{ks} are the elements of the S matrix. From the orthonormality of the scattering states, we obtain

$$S_{ks} = (\Psi_k^{(-)}, \Psi_s^{(+)}) \tag{20.56}$$

In this section, we prove that the two expressions (20.54) and (20.56) are indeed equal. The definition (20.55) implies that the S matrix is diagonal with respect to the energy. The S-matrix elements between states of different energy are zero, as is seen explicitly in the formula (20.54).

In the case of simple scattering, the incident state is represented asymptotically by the free particle "in" state $\Psi_{\mathbf{k}} = |\mathbf{k}\rangle$ which feeds the outgoing spherical wave associated with the state $\Psi_{\mathbf{k}}^{(+)}$. In a scattering experiment, we are asking for the probability amplitude that the particle is asymptotically found in the free particle "out" state $\Psi_{\mathbf{k}'} = |\mathbf{k}'\rangle$, which is fed by the incoming spherical wave associated with the state $\Psi_{\mathbf{k}'}^{(-)}$. The expansion (20.55) takes the explicit form

$$\Psi_{\mathbf{k}}^{(+)} = \sum_{\mathbf{k}'} \Psi_{\mathbf{k}'}^{(-)} S_{\mathbf{k}'\mathbf{k}} = \sum_{\mathbf{k}'} \Psi_{\mathbf{k}'}^{(-)} \langle \mathbf{k}' | S | \mathbf{k} \rangle \tag{20.57}$$

where we have defined the *scattering operator* S by its matrix elements,

$$\langle \mathbf{k}' | S | \mathbf{k} \rangle = (\Psi_{\mathbf{k}'}^{(-)}, \Psi_{\mathbf{k}}^{(+)}) = S_{\mathbf{k}'\mathbf{k}} \tag{20.58}$$

between the "in" state $|\mathbf{k}\rangle$ and the "out" state $|\mathbf{k}'\rangle$.

To prove the equality of the representations (20.54) and (20.56), we use the formula

$$\Psi_k^{(-)} = \Psi_k + \frac{1}{E_k - H - i\hbar\alpha} V\Psi_k \tag{20.59}$$

which is the analogue of (20.40) for $\Psi_k^{(-)}$ and substitute it in (20.56):

$$\begin{aligned} S_{ks} &= (\Psi_k, \Psi_s^{(+)}) + \left(\frac{1}{E_s - H - i\hbar\alpha} V\Psi_k, \Psi_s^{(+)} \right) \\ &= \left(\Psi_k, \Psi_s + \frac{1}{E_s - H_0 + i\hbar\alpha} V\Psi_s^{(+)} \right) + \left(\frac{1}{E_k - H - i\hbar\alpha} V\Psi_k, \Psi_s^{(+)} \right) \end{aligned}$$

where the Lippmann-Schwinger equation (20.23) has been used in the last equality. Since Ψ_k is an eigenstate of H_0 and $\Psi_s^{(+)}$ is an eigenstate of H, we can reduce this expression to

$$S_{ks} = \delta_{ks} + \left(-\frac{1}{E_k - E_s - i\hbar\alpha} + \frac{1}{E_k - E_s + i\hbar\alpha} \right) (\Psi_k, V\Psi_s^{(+)}) \tag{20.60}$$

Using a standard representation of the delta function, Eq. (A.21) in the Appendix, we recover the formula (20.54) for the S-matrix elements.

If plane waves with periodic boundary conditions in a box of volume L^3 are chosen for the set of eigenvectors Ψ_s, the scattering matrix becomes

$$S_{\mathbf{k}'\mathbf{k}} = \delta_{\mathbf{k}'\mathbf{k}} + \frac{4\pi^2 i}{kL^3} \delta(k - k') f_{\mathbf{k}}(\hat{\mathbf{k}}') \tag{20.61}$$

As we saw in Section 20.2, the differential scattering cross section must be proportional to $|S_{\mathbf{k}'\mathbf{k}} - \delta_{\mathbf{k}'\mathbf{k}}|^2$ or $|f_{\mathbf{k}}(\hat{\mathbf{k}}')|^2$.

From (20.55) and (20.20) it is easy to derive the simple formula

$$T_{js} = \sum_r (\Psi_j, V\Psi_r^{(-)}) S_{rs} \tag{20.62}$$

This also follows directly from the definition of the scattering matrix. If $E_j = E_s$, on account of (20.42) it may also be written as

$$T_{js} = \sum_r T^*_{rj} S_{rs} \tag{20.63}$$

The scattering matrix owes its central importance to the fact that it is *unitary*. To prove the unitary property in the time-independent formulation, we must show that

$$\sum_n S^*_{n\ell} S_{nj} = \delta_{\ell j} \tag{20.64}$$

and

$$\sum_n S_{\ell n} S^*_{jn} = \delta_{\ell j} \tag{20.65}$$

The first of these equations follows immediately from the definition (20.55) and the orthonormality of the $\Psi^{(+)}$ and $\Psi^{(-)}$. The second equation is proved by using (20.56) to construct

$$\sum_n S_{\ell n} S^*_{jn} = \sum_n (\Psi^{(-)}_\ell, \Psi^{(+)}_n)(\Psi^{(+)}_n, \Psi^{(-)}_j)$$

If $\Psi^{(b)}_i$ denotes the bound states, we have as a result of completeness the closure relation

$$\sum (\Phi_a, \Psi^{(+)}_n)(\Psi^{(+)}_n, \Phi_b) + \sum_i (\Phi_a, \Psi^{(b)}_i)(\Psi^{(b)}_i, \Phi_b) = (\Phi_a, \Phi_b)$$

for any two states Φ_a and Φ_b. Applying this to the previous equation, we get

$$\sum_n S_{\ell n} S^*_{jn} = (\Psi^{(-)}_l, \Psi^{(-)}_j) - \sum_n (\Psi^{(-)}_l, \Psi^{(b)}_i)(\Psi^{(b)}_i, \Psi^{(-)}_j) = (\Psi^{(-)}_l, \Psi^{(-)}_j) = \delta_{\ell j}$$

because the bound states are orthogonal to the scattering states.

In the time-dependent formulation, the S-matrix element S_{ks} is by its definition (20.7) equal to the matrix element of $\tilde{T}(+\infty, -\infty)$ between the initial unperturbed state Ψ_s and the final unperturbed state Ψ_k. Hence, formally,

$$\boxed{S = \tilde{T}(+\infty, -\infty)} \tag{20.66}$$

The time development operator is unitary, and thus the operator S is unitary, as expected. The S-matrix element S_{ks} is the probability amplitude for finding the system at $t = +\infty$ in state Ψ_k if it was known to have been in state Ψ_s at $t = -\infty$.

The connection between the two definitions of the S matrix is further clarified by the two equivalent expressions for the stationary scattering states:

$$\Psi^{(+)}_s = \Psi_s + G_+(E_s) V \Psi^{(+)}_s \tag{20.67}$$

which is the Lippmann-Schwinger equation, and

$$\Psi^{(+)}_s = \sum_k [\Psi_k S_{ks} + G_-(E_k) V \Psi^{(-)}_k S_{ks}] \tag{20.68}$$

obtained by substituting (20.33) into (20.55). If we now imagine that wave packets are constructed from these expressions by superposition, we may relate them to the time-dependent description of the scattering process. At $t = -\infty$, only the first term

in (20.67) contributes, and at $t = +\infty$, only the first terms in (20.68) contribute, since the retarded (G_+) wave vanishes before the scattering and the advanced (G_-) wave vanishes after the scattering. Hence, the matrix S_{ks} connects the free initial state s with the free final states k as described by

$$\boxed{S_{ks} = \langle k|S|s\rangle = \langle k|\tilde{T}(+\infty, -\infty)|s\rangle} \tag{20.69}$$

The scattering operator S is useful because it depends only on the nature of the system and the partition of the Hamiltonian H into a "free" Hamiltonian H_0 and an interaction V, but not on the particular incident state. A simple application to scattering from a central-force field in the next section will illustrate the advantages of working with the scattering operator.

For formal manipulations in scattering theory, it is sometimes convenient to use the operators T, $\Omega^{(+)}$, and $\Omega^{(-)}$ defined by the equations

$$T_{ks} = (\Psi_k, T\Psi_s) = \langle k|T|s\rangle = (\Psi_k, V\Psi_s^{(+)}) \tag{20.70}$$

and

$$\Omega^{(+)}\Psi_r = \Psi_r^{(+)}, \qquad \Omega^{(-)}\Psi_r = \Psi_r^{(-)} \quad \text{for all } r \tag{20.71}$$

From the last definition it is evident that $\Omega^{(+)}$ and $\Omega^{(-)}$ preserve the norm of vectors on which they act. Nevertheless, in general they are not unitary (only *isometric*), because the full Hamiltonian $H = H_0 + V$ may support discrete bound eigenstates, which can be expanded in terms of the unperturbed states Ψ_r, but are orthogonal to the states $\Psi_r^{(+)}$ and $\Psi_r^{(-)}$. Therefore, $\Omega^{(+)\dagger}\Omega^{(+)} = \Omega^{(-)\dagger}\Omega^{(-)} = I$, but $\Omega^{(+)}\Omega^{(+)\dagger} = \Omega^{(-)}\Omega^{(-)\dagger} = I$ is not necessarily valid.[3]

Exercise 20.5. Prove that

$$T = V\Omega^{(+)} \tag{20.72}$$

and

$$S = [\Omega^{(-)}]^\dagger\Omega^{(+)} \tag{20.73}$$

Exercise 20.6. Show that

$$H\Omega^{(\pm)} = \Omega^{(\pm)} H_0 \tag{20.74}$$

and from this relation deduce that

$$[S, H_0] = 0 \tag{20.75}$$

and verify again that the S matrix is diagonal with respect to the energy.

5. ***Rotational Invariance, Time Reversal Symmetry, and the S Matrix.*** If a spatial rotation is applied to all the states of a physical system in which scattering occurs, the initial and final momentum states are rotated rigidly. According to Section 17.2, this is accomplished by applying a unitary operator U_R. If the forces are central, $V = V(r)$, the Hamiltonians H and H_0 are both invariant under rotations, and the scattering matrix will be the same before and after the rotation. In this case we have

$$\langle U_R\mathbf{k}'|S|U_R\mathbf{k}\rangle = \langle \mathbf{k}'|S|\mathbf{k}\rangle$$

[3]For a more detailed exposition, see Taylor (1972).

Hence, the scattering matrix cannot depend on the absolute orientation of the vectors $\mathbf{k}$ and $\mathbf{k}'$ in space. It can only be a function of the energy and of the angle between the initial and final momenta. If the particles have no spin, the completeness of the Legendre polynomials allows us to write the scattering matrix in the form

$$\langle \mathbf{k}' | S | \mathbf{k} \rangle = \delta(k - k') \sum_{\ell=0}^{\infty} F_\ell(k) P_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}}') \tag{20.76}$$

with undetermined coefficients $F_\ell(k)$. The delta function has been included as a separate factor, because we already know that the S matrix has nonvanishing elements only "*on the energy shell*," i.e., between two states of the same energy.

The coefficients $F_\ell(k)$ can be determined to within a phase factor by invoking the unitarity of the scattering matrix:

$$\int \langle \mathbf{k}' | S | \mathbf{k}'' \rangle \langle \mathbf{k} | S | \mathbf{k}'' \rangle^* \, d^3k'' = \delta(\mathbf{k} - \mathbf{k}')$$

Substituting (20.76) here and carrying out the integration in $\mathbf{k}''$-space by recourse to the addition theorem of spherical harmonics, Eq. (11.100), we obtain

$$k^2 \delta(k - k') \sum_{\ell=0}^{\infty} \frac{4\pi}{2\ell + 1} |F_\ell(k)|^2 P_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}}') = \delta(\mathbf{k} - \mathbf{k}')$$

We now use the identity (11.105) to write:

$$\delta(\mathbf{k} - \mathbf{k}') = \frac{\delta(k - k')}{k^2} \sum_{\ell=0}^{\infty} \frac{2\ell + 1}{4\pi} P_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}}') \tag{20.77}$$

From the last two equations, we immediately find that the coefficients $F_\ell(k)$ must be of the form

$$F_\ell(k) = \frac{2\ell + 1}{4\pi k^2} e^{2i\delta_\ell(k)} \tag{20.78}$$

where the $\delta_\ell(k)$ are real functions of the momentum (or energy). Inserting this result in (20.76), we conclude that the scattering matrix is expressible as

$$\langle \mathbf{k}' | S | \mathbf{k} \rangle = \delta(k - k') \sum_{\ell=0}^{\infty} \frac{2\ell + 1}{4\pi k^2} e^{2i\delta_\ell(k)} P_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}}') \tag{20.79}$$

On the other hand, according to (20.61) and (20.77), this matrix element can also be written in terms of the scattering amplitude as

$$\langle \mathbf{k}' | S | \mathbf{k} \rangle = \delta(k - k') \left[\sum_{\ell=0}^{\infty} \frac{2\ell + 1}{4\pi k^2} P_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}}') + \frac{i}{2\pi k} f_k(\hat{\mathbf{k}}') \right] \tag{20.80}$$

Here we have chosen the dimension of the normalization cube $L = 2\pi$, so that the unperturbed eigenstates are normalized as $\langle \mathbf{k}' | \mathbf{k} \rangle = \delta(\mathbf{k} - \mathbf{k}')$ throughout. Comparing (20.79) and (20.80), we get

$$f_k(\hat{\mathbf{k}}') = \frac{1}{k} \sum_{\ell=0}^{\infty} (2\ell + 1) e^{2i\delta_\ell(k)} \sin \delta_\ell(k) P_\ell(\hat{\mathbf{k}} \cdot \hat{\mathbf{k}}') \tag{20.81}$$

We have thus rederived the main result of the partial wave analysis (13.70) directly from the rotational invariance of the scattering operator.

Exercise 20.7. Transform the matrix element (20.79) into the orbital angular momentum representation (see Exercise 17.40), and show that

$$\langle \alpha \ell' m' | S | \alpha \ell m \rangle = e^{2i\delta_\ell(k)} \delta_{m'm} \delta_{\ell'\ell} \tag{20.82}$$

verifying that $e^{2i\delta_\ell(k)}$ are the eigenvalues of the S matrix for a rotationally invariant interaction, in agreement with Eq. (13.75).[4]

In order to analyze the symmetry properties of the scattering states under the time reversal operation, we write the fundamental integral equations using momentum eigenstates as the unperturbed states:

$$\Psi_{\mathbf{k}}^{(+)} = \Psi_{\mathbf{k}} + \frac{1}{E_k - H_0 + i\hbar\alpha} V\Psi_{\mathbf{k}}^{(+)} \tag{20.83}$$

and

$$\Psi_{\mathbf{k}}^{(-)} = \Psi_{\mathbf{k}} + \frac{1}{E_k - H_0 - i\hbar\alpha} V\Psi_{\mathbf{k}}^{(-)} \tag{20.84}$$

If we now apply the antiunitary operator Θ defined in Section 17.9 and choose the phases of the momentum eigenstates such that $\Theta\Psi_{\mathbf{k}} = \Psi_{-\mathbf{k}}$, we obtain

$$\Theta\Psi_{\mathbf{k}}^{(+)} = \Psi_{-\mathbf{k}} + \frac{1}{E_k - H_0 - i\hbar\alpha} \Theta V \Theta^{-1} \Theta \Psi_{\mathbf{k}}^{(+)} \tag{20.85}$$

where use has been made of the invariance of H_0 under time reversal. Comparing this relation with (20.84), we observe that $\Psi_{\mathbf{k}}^{(+)}$ and $\Psi_{-\mathbf{k}}^{(-)}$ are mutually time-reversed states,

$$\Theta\Psi_{\mathbf{k}}^{(+)} = \Psi_{-\mathbf{k}}^{(-)} \tag{20.86}$$

if the interaction V is invariant under time reversal:

$$\Theta V \Theta^{-1} = V \tag{20.87}$$

In this case, the S matrix satisfies the condition

$$\begin{aligned} \langle \mathbf{k}' | S | \mathbf{k} \rangle &= (\Psi_{\mathbf{k}'}^{(-)}, \Psi_{\mathbf{k}}^{(+)}) = (\Theta\Psi_{\mathbf{k}'}^{(-)}, \Theta\Psi_{\mathbf{k}}^{(+)})^* \\ &= (\Psi_{-\mathbf{k}'}^{(-)}, \Psi_{-\mathbf{k}}^{(+)}) = \langle -\mathbf{k} | S | -\mathbf{k}' \rangle \end{aligned} \tag{20.88}$$

owing to the antiunitary property of Θ. For the scattering amplitude, this implies by (20.54) the relation

$$f_{\mathbf{k}}(\mathbf{k}') = f_{-\mathbf{k}'}(-\mathbf{k}) \tag{20.89}$$

This equation, derived from very general symmetry properties, expresses the equality of two scattering processes obtained by reversing the path of the particle and is known as the *reciprocity relation.*

6. ***The Optical Theorem.*** From the unitary property of the scattering matrix, we can derive an important theorem for the scattering amplitudes. If we substitute the

[4]The same symbol (δ) is used for delta functions and phase shifts in this section, but the context always determines the meaning unambiguously.

expression (20.54) for the S matrix in the unitarity condition (20.64) and work out the result, we get

$$\boxed{2\pi \sum_n \delta(E_n - E_r) T^*_{nr} T_{nj} = i(T_{rj} - T^*_{jr})} \tag{20.90}$$

By (20.37), this formula can also be written in terms of the scattering amplitudes. Replacing the summation by an integration, we obtain by use of the appropriate density-of-states factor

$$2\pi \frac{2\pi\hbar^2}{mL^3} \frac{L^3}{(2\pi)^3} \int f^*_{\mathbf{k}}(\hat{\mathbf{k}}'') f_{\mathbf{k}'}(\hat{\mathbf{k}}'') \frac{m}{\hbar^2 k''} \delta(k'' - k)\, k''^2\, dk''\, d\Omega'' = -i[f_{\mathbf{k}'}(\hat{\mathbf{k}}) - f^*_{\mathbf{k}}(\hat{\mathbf{k}}')]$$

or

$$\boxed{\int f^*_{\mathbf{k}}(\hat{\mathbf{k}}'') f_{\mathbf{k}'}(\hat{\mathbf{k}}'')\, d\Omega'' = \frac{4\pi}{k} \frac{f_{\mathbf{k}'}(\hat{\mathbf{k}}) - f^*_{\mathbf{k}}(\hat{\mathbf{k}}')}{2i}} \tag{20.91}$$

As a special case of this relation, we may identify $\mathbf{k}'$ and $\mathbf{k}$ and then obtain by comparison with (20.39):

$$\boxed{\sigma = \int d\sigma = \int |f_{\mathbf{k}}(\hat{\mathbf{k}}'')|^2\, d\Omega'' = \frac{4\pi}{k} \operatorname{Im} f_{\mathbf{k}}(\hat{\mathbf{k}})} \tag{20.92}$$

This formula shows that the imaginary part of the *forward scattering amplitude* $f_{\mathbf{k}}(\hat{\mathbf{k}})$ measures the loss of intensity that the incident beam suffers because of the scattering. It therefore expresses the conservation of probability, which is a consequence of the Hermitian property of the Hamiltonian. The unitarity of S is directly linked to the Hermitian property of H, since according to (20.66) S is the limit of the time development operator (in the interaction picture).

Equations (20.90), (20.91), and especially (20.92) are generically known as expressions of the *optical theorem*, because of the analogy with light that passes through a medium. In optics, the imaginary part of the complex index of refraction is related to the total absorption cross section. Application of the optical theorem to scattering from a central-force potential was the subject of Exercise 13.14.

Exercise 20.8. Derive the optical theorem (20.90) directly from (20.11), using conservation of probability.

Exercise 20.9. Show that the first Born approximation violates the optical theorem. Explain this failure and show how it can be remedied by including the second Born approximation for the forward scattering amplitude.

Problems

1. Obtain the "scattering states" (energy eigenstates with $E \geq 0$) for a one-dimensional delta-function potential, $g\delta(x)$. Calculate the matrix elements $\langle k' | S | k \rangle$ and verify the unitarity of the S matrix. Obtain the transmission coefficient, and compare with Eq. (6.19) and Exercise 6.13. Perform the calculations in both the coordinate and momentum representations.

2. Use the Born approximation to calculate the differential and total cross sections for the elastic scattering of electrons by a hydrogen atom that is in its ground state. Approximate the interaction between the continuum electron and the atom by the static field of the atom and neglect exchange phenomena.
3. The cross section for two-quantum annihilation of positrons of velocity v with an electron at rest has, for $v \ll c$, the "classical" value

$$\sigma = \pi\left(\frac{e^2}{m_e c^2}\right)\frac{c}{v}$$

Use this information to derive the annihilation probability per unit time from a plane wave state and (assuming that annihilation occurs only if the two particles are at the same place) estimate the decay probability in the singlet ground state of positronium.

4. Using the Born approximation, and neglecting relativistic effects, express the differential cross section for scattering of an electron from a spherically symmetric charge distribution $\rho(r)$ as the product of the Rutherford scattering cross section for a point charge and the square of a *form factor* F. Obtain an expression for the form factor and evaluate it as a function of the momentum transfer for (a) a uniform charge distribution of radius R, and (b) a Gaussian charge distribution with the same root-mean-square radius.
5. If the nonlocal *separable* scattering potential

$$\langle \mathbf{r}' | V | \mathbf{r}'' \rangle = \lambda u(r')u(r'')$$

is given, work out explicitly and solve the integral equation for $\Psi^{(+)}$. Obtain the scattering amplitude, and discuss the Born series for this exponential.

CHAPTER 21

Identical Particles

When we endeavor to describe the quantum behavior of systems of several identical particles, altogether new, strange, and unclassical features are encountered because there is no way of keeping track of each particle separately when the wave functions of two indistinguishable particles overlap. The indistinguishability of identical particles has important, experimentally verifiable consequences. For instance, the spectra of many-electron atoms would be altogether different if the atomic electrons were distinguishable. The Pauli exclusion principle, which explains the periodic table of elements and the stability of complex atoms, presupposes the indistinguishability of electrons. The quantum theory for identical particles is most efficiently formulated in terms of operators that represent the occupation number, or density of occupation, of one-particle states and the related creation and annihilation operators. Since it affords a direct route to quantum field theory, and the physics of particles and quasiparticles which can be created and destroyed, we will work in this so-called second quantization mode, but the connection to wave mechanics in configuration space will also be displayed. By extending the principles of one-particle nonrelativistic quantum mechanics to the description of identical particles, we will be led to bosons (symmetric under permutations and obeying Bose-Einstein statistics) and fermions (antisymmetric, with Fermi-Dirac statistics) as the only types of particles consistent with perfect indistinguishability.

1. ***The Indistinguishability of and the State Vector Space for Identical Particles.*** The indistinguishability of two particles makes it impossible to follow them individually in the region where both may be found simultaneously. Similar complications do not arise in classical mechanics from the identity of two bodies because their wave packets do not overlap, and the particles move in separate, distinguishable, continuous orbits. Even if they are not followed as they move along their orbits, two objects that are identical can, classical mechanics assumes, always be made distinguishable by marking them in such a way that there is no measurable influence on the physical process under consideration. For instance, we suppose that coloring the balls in billiards differently has no influence on their motion; yet it serves to distinguish them individually and to identify them after any number of collisions.

If the particles are distinguishable, we saw in Section 15.4 that the appropriate quantum mechanical description is a state vector space constructed as a direct product space from the one-particle theory. The same strategy can be employed for n indistinguishable particles, but only at the expense of labeling or numbering the identical particles. Since the order in which the particles are labeled has, by the definition of particle identity, no physical significance, state vectors (or wave functions) that differ only in the permutation of the labels must define the same state. Thus, there is in this mode of description a symmetry under the $n!$ permutations of the particle labels, violating the principle that (except for a phase factor) there is exactly one state vector for every physical state of a system. The redundancy in the

description of identical particles must then be removed by limiting the state vector space to a subspace that is invariant under the permutations of labels.

We prefer to build up the quantum mechanics of n identical particles without introducing redundant states that must be subsequently excised. Let us assume the existence of a complete set of dynamical variables K, appropriate for a single particle. As in Section 9.1, we characterize the corresponding set of eigenvalues collectively by K_i, without attaching primes. A particular state of the system is then specified by stating that of the n particles n_1 have the value K_1, n_2 have the value K_2, etc. More general states are obtained by superposition of such particular n-particle states. It is impossible to tell which particle has the value K_1, which one the value K_2, etc. For example, the coordinates x, y, z and the spin component S_z constitute a complete set of dynamical variables for electrons. By measurement, it can be determined that there is one electron at position $\mathbf{r}_1$ with spin component $(\hbar/2)\sigma_1$ (where $\sigma_1 = +1$ or -1), another electron at position $\mathbf{r}_2$ with spin component $(\hbar/2)\sigma_2$, etc., but it is not possible to make any further identification of the electrons.

A basic arbitrariness is inherent in the definition of the term *particle* as it is used here, and the identity of two particles is to a certain extent a matter of convention. For example, protons and neutrons may be regarded as two distinct species, distinguishable by differences in mass, charge, magnetic moment, and decay properties. Yet, since interactions can transform protons into neutrons and vice versa, it is frequently convenient to describe them as two different states of the same species, the *nucleon*. All nucleons are then considered as identical, and proton and neutron states are characterized by different values of a new dynamical variable, the *isobaric* "*spin*," or *isospin*. In high-energy processes, which are dominated by the strong interactions, strangeness, hypercharge, and other quantum numbers associated with the symmetry group $SU(3)$ place the neutron and proton in a particular multiplet (octet) which has six other, "strange" spin one-half baryonic members. At a deeper level, these assignments of quantum numbers are understood in terms of the constituent quark structure of these composite "particles." The quarks themselves can be treated as species with distinguishable flavors, that is, up, down, and strange, and so forth, and distinctive colors (red, green, blue), or they may be thought of as one kind of particle that can be in states with different flavor and color "quantum numbers."

In certain problems, on the other hand, it may be convenient to regard entire nuclei, atoms, or molecules as particles. Their composite nature implies the existence of some internal degrees of freedom, which must be included among the dynamical variables used to describe the states of the system.

The latitude in the definition of what is meant by a particle is no obstacle in the development of the quantum mechanics for identical particles. On the contrary, the general principles of the theory are applicable to any species of particle, characterized by a set of dynamical variables that can be regarded as complete in the context of the problems under study.

In developing a mathematical framework for these concepts, we define the quantum state vector space for a system of n identical particles on the basic assumption that

> *any complete set of operators K, which describes the behavior of a single particle, can also be employed for n particles of the same kind.*

This minimal assumption, postulated to be true even in the presence of interactions between the particles, implies that the composite system retains the properties of the individual constituent particles to a considerable extent.

The physical context defines the meaning of this assumption. For example, in the theory of atomic and molecular structure and spectra, and in solid-state physics, it makes sense to consider all electrons as replicas of a single electron, with its position and momentum coordinates and its spin (one-half), and *no additional attributes*. In relativistic quantum electrodynamics, however, it is necessary to include a further degree of freedom, corresponding to the positron which, as the antiparticle of the electron, represents another state that can be excited by interaction with the electromagnetic field.

As another example, helium in its gaseous or condensed (liquid or solid) form is properly regarded as an assemblage of identical inert helium atoms in their ground state, provided that the temperatures are not excessively high. At stellar temperatures, this picture breaks down, and the internal degrees of freedom of the helium atom become active. At even higher, thermonuclear, temperatures, the degrees of freedom relevant to the structure of the nucleus must be taken into account.

As long as we assume that the system composed of identical particles is merely a collection of the constituent particles, we can postulate that to each eigenvalue K_i of K corresponds an *occupation number operator* N_i whose eigenvectors characterize the states in which a definite number, n_i, of the particles has the value K_i. The eigenvalues of N_i are the *occupation numbers* n_i. As a fundamental postulate, we assume that the totality of the operators N_i forms a complete set of commuting Hermitian operators for the system of identical particles. (If the system is composed of several different species of identical particles, occupation number operators for each of these are introduced.) In the next two sections, we will see that this postulate is responsible primarily for the unique roles played by Bose-Einstein and Fermi-Dirac statistics. However, we must first construct the state vector space of the many-body system by a suitable generalization of one-particle quantum mechanics, paying due attention to the indistinguishability of the particles.

The fundamental postulate implies that in the state vector space of the many-body system, also known as *Fock space*, the basis vectors

$$|n_1, n_2, n_3, \ldots\rangle \tag{21.1}$$

which allot the eigenvalue K_1 to n_1 particles, the eigenvalue K_2 to n_2 particles, etc., constitute a complete set of orthonormal basis vectors for the system of identical particles. The most general state of the system is a linear combination of the kets (21.1).

In particular, we have the ''no-particle'' (or vacuum) state

$$\Psi^{(0)} \equiv |\mathbf{0}\rangle = |0, 0, 0, \ldots\rangle \tag{21.2}$$

Next there are the one-particle states

$$\Psi_i^{(1)} \equiv |K_i\rangle = |0, 0, \ldots, n_i = 1, 0, 0 \ldots\rangle \tag{21.3}$$

spanning the one-particle subspace of the much larger vector space of the many-body system, which is composed of states with zero, one, two . . . particles and their linear combinations. The principles of quantum mechanics in the one-particle subspace, which have been the subject of most of this book so far, will now be generalized to systems containing an arbitrary, and variable, number of identical particles.

A remark on notation: Equation (21.3) puts us on guard and warns us that in this and the next chapter the bra-ket notation will be used in both its narrow single-particle form, $|K_i\rangle$, and its Fock space version, sometimes even in the same equation.

Does the postulate adopted in this section provide the necessary ingredients for a quantum theory of interacting particles? To answer this question, it must be shown

that a consistent theory can be constructed in this framework and that the theoretical calculations are in satisfactory agreement with experimental observations. The assumption that the state of a system of interacting particles can be expanded in terms of the states of noninteracting particles finds broad support in many successful calculations. Especially when the relatively weak electromagnetic interactions predominate, as in atoms, molecules, and solids, perturbation theory affords an account of almost unlimited precision, provided that the theory is "renormalized." Renormalization involves rules designed to make consistent and accurate predictions possible, in spite of the appearance of divergent quantities in the calculations. Quantum electrodynamics and its offspring, electroweak gauge theory, are the most mature versions of the theory of interacting fields and particles. Even the theory of the fundamentally nonperturbative strong interactions, quantum chromodynamics, owes its success to the existence of a regime ("asymptotic freedom") in which the concepts of perturbation theory are applicable and the constituents may be regarded as nearly noninteracting.

2. ***Creation and Annihilation Operators.*** For a given species of particles, it is useful to define *creation operators* $a_i^\dagger$ with the property

$$a_i^\dagger |n_1, n_2, \ldots n_{i-1}, n_i, n_{i+1}, \ldots\rangle \propto |n_1, n_2, \ldots n_{i-1}, n_i + 1, n_{i+1}, \ldots\rangle \quad (21.4)$$

adding to the basis state one particle with quantum number K_i. From the definition of a Hermitian adjoint operator, it follows conversely that an *annihilation operator* a_i has the property

$$a_i |n_1, n_2, \ldots n_{i-1}, n_i, n_{i+1}, \ldots\rangle \propto |n_1, n_2, \ldots n_{i-1}, n_i - 1, n_{i+1}, \ldots\rangle \quad (21.5)$$

and thus in effect removes one particle with quantum number K_i. The constants of proportionality in these defining relations are yet to be determined. In the interest of simplicity, we begin this determination by requiring

$$a_i^\dagger \Psi^{(0)} = a_i^\dagger |\mathbf{0}\rangle = \Psi_i^{(1)} = |K_i\rangle = |0, 0, \ldots n_i = 1, 0 \ldots\rangle \quad (21.6)$$

$$a_i \Psi_j^{(1)} = a_i |K_j\rangle = a_i |0, 0, \ldots n_j = 1, 0, \ldots\rangle = \delta_{ij} \Psi^{(0)} = \delta_{ij} |\mathbf{0}\rangle \quad (21.7)$$

Since the vacuum contains no particle to be destroyed, we also demand that

$$a_i \Psi^{(0)} = a_i |\mathbf{0}\rangle = 0 \quad (21.8)$$

Exercise 21.1. If $\Psi^{(1)}$ is an arbitrary one-particle state, prove that the probability, $|\langle 0|a_i|\Psi^{(1)}\rangle|^2 = |\langle K_i|\Psi^{(1)}\rangle|^2$, of finding the particle to have quantum number K_i is equal to the expectation value $\langle \Psi^{(1)}|a_i^\dagger a_i|\Psi^{(1)}\rangle$.

In order to establish how a unitary transformation from one basis to another in the one-particle theory can be expressed in this new notation, as in Section 9.5, we introduce a second complete set of one-particle observables L with eigenvalues L_q. The corresponding occupation number operators will be denoted by $\tilde{N}_q$ with eigenvalues $\tilde{n}_q$. The transformation between the two representations is effected by

$$|K_i\rangle = \sum_q |L_q\rangle\langle L_q|K_i\rangle \quad (21.9)$$

The complex transformation coefficients $\langle L_q|K_i\rangle$ form a unitary matrix.

The vacuum or zero-particle state $\Phi^{(0)}$ is the same in the new as in the old basis:

$$\Phi^{(0)} = \Psi^{(0)} = |\mathbf{0}\rangle \quad (21.10)$$

The one-particle states in the new basis are

$$\Phi_q^{(1)} = |0, 0, \ldots, \tilde{n}_q = 1, 0, 0, \ldots\rangle = |L_q\rangle$$

We now introduce creation operators $b_q^\dagger$ and annihilation operators b_q in complete analogy with (21.4) and (21.5):

$$b_q^\dagger|\tilde{n}_1, \tilde{n}_2, \ldots \tilde{n}_{q-1}, \tilde{n}_q, \tilde{n}_{q+1}, \ldots\rangle \propto |\tilde{n}_1, \tilde{n}_2, \ldots \tilde{n}_{q-1}, \tilde{n}_q + 1, \tilde{n}_{q+1}, \ldots\rangle \quad (21.11)$$

$$b_q|\tilde{n}_1, \tilde{n}_2, \ldots \tilde{n}_{q-1}, \tilde{n}_q, \tilde{n}_{q+1}, \ldots\rangle \propto |\tilde{n}_1, \tilde{n}_2, \ldots \tilde{n}_{q-1}, \tilde{n}_q - 1, \tilde{n}_{q+1}, \ldots\rangle \quad (21.12)$$

Evidently, for the one-particle states, the two bases are connected by the equation,

$$a_i^\dagger|\mathbf{0}\rangle = \Psi_i^{(1)} = |K_i\rangle = \sum_q |L_q\rangle\langle L_q|K_i\rangle = \sum_q b_q^\dagger\langle L_q|K_i\rangle|\mathbf{0}\rangle$$

This equation is satisfied by the transformation condition

$$\boxed{a_i^\dagger = \sum_q b_q^\dagger\langle L_q|K_i\rangle} \quad (21.13)$$

The Hermitian adjoint of this equation,

$$\boxed{a_i = \sum_q \langle K_i|L_q\rangle b_q} \quad (21.14)$$

is trivially satisfied when acting on the vacuum state. Applying a_i to a one-particle state, we have, by the use of Eqs. (21.7) and (21.8) and their analogues,

$$a_i\Psi_j^{(1)} = \delta_{ij}\Psi^{(0)} = \sum_q \langle K_i|L_q\rangle\langle L_q|K_j\rangle\Psi^{(0)} = \sum_{q,r} \langle K_i|L_q\rangle\langle L_r|K_j\rangle\delta_{qr}\Psi^{(0)}$$
$$= \sum_q b_q\langle K_i|L_q\rangle \sum_r \Phi_r^{(1)}\langle L_r|K_j\rangle = \sum_q b_q\langle K_i|L_q\rangle\Psi_j^{(1)}$$

This relation is again satisfied by Eq. (21.14). Hence we may adopt the operator Eqs. (21.13) and (21.14) in the subspace of no particles and one particle.

It is a basic notion of quantum mechanics that the creation of a particle with quantum number K_i is equivalent to the creation of a particle with any one of the quantum numbers L_q, each contributing an amplitude $\langle L_q|K_i\rangle$ in a linear superposition. The operator Eq. (21.13) is consistent with this general requirement. The change to yet another, third one-particle basis M_s is effected by the relation

$$a_i^\dagger = \sum_q b_q^\dagger\langle L_q|K_i\rangle = \sum_s \left(\sum_q b_q^\dagger\langle L_q|M_s\rangle\right)\langle M_s|K_i\rangle$$

explicitly exhibiting the equivalence of all one-particle bases or representations for constructing the many-body theory, if (21.13) is valid.

We therefore make the fundamental assumption that the transformation equations (21.13) and (21.14) will hold beyond the no-particle and one-particle subspaces as operator equations in the entire state vector space of the system of any number of identical particles. We will see that this assumption determines the structure of the commutation relations that the creation and annihilation operators must satisfy and narrows the freedom we have in deriving the constants of proportionality which enter (21.4), (21.5), (21.11), and (21.12). Natural, and almost innocuous, as this assumption, or *principle of unitary symmetry*, may appear, it is ultimately responsible for the restriction to Bose-Einstein and Fermi-Dirac statistics in the quantum theory of many-particle systems. Hypothetical quanta or particles called *anyons*,

which violate the fundamental assumption made here, are useful constructs in certain approximate models of (especially, two-dimensional) many-body systems, but they and other unconventional versions of quantum statistics (so-called parastatistics) are outside the scope of this book.

3. ***The Algebra of Creation and Annihilation Operators.*** The first step in deriving the algebraic properties of creation and annihilation operators is to note that two creation operators $a_i^\dagger$ and $a_j^\dagger$, when applied successively to a state, produce the same physical state, although the normalization may depend on the order in which the two particles are created. Hence, for any state Ψ,

$$a_i^\dagger a_j^\dagger \Psi = \lambda a_j^\dagger a_i^\dagger \Psi \tag{21.15}$$

To show that the constant λ is independent of the state Ψ and of the subscripts i and j—and indeed of the particular one-particle representation chosen—it is advantageous to consider what happens in changing to another one-particle representation. Substituting $a_i^\dagger$ from Eq. (21.13) into Eq. (21.15), we get

$$(a_i^\dagger a_j^\dagger - \lambda a_j^\dagger a_i^\dagger)\Psi = \sum_{k,\ell} \langle L_k | K_i \rangle \langle L_\ell | K_j \rangle (b_k^\dagger b_\ell^\dagger - \lambda b_\ell^\dagger b_k^\dagger)\Psi = 0 \tag{21.16}$$

The transformation coefficients are arbitrary complex numbers, except for the constraint imposed by unitarity:

$$\sum_k \langle K_i | L_k \rangle \langle L_k | K_j \rangle = \delta_{ij} \tag{21.17}$$

If the theory is to have the same form in any representation (unitary symmetry), Eq. (21.16) can be satisfied for all states Ψ only if, for every value of k and ℓ,

$$b_k^\dagger b_\ell^\dagger - \lambda b_\ell^\dagger b_k^\dagger = 0$$

Hence, also,

$$b_\ell^\dagger b_k^\dagger - \lambda b_k^\dagger b_\ell^\dagger = 0$$

and consequently

$$\lambda^2 = 1 \quad \text{or} \quad \lambda = \pm 1$$

According to the assumption spelled out in Section 21.2, these algebraic relations for creation operators must hold for any of the unitarily equivalent one-particle representations, including the original basis created from the vacuum by the operators $a_i^\dagger$.

It follows that there are two and only two classes of algebraic relations for creation operators of various particle species. They must satisfy either the *commutation relations*

$$\boxed{a_i^\dagger a_j^\dagger - a_j^\dagger a_i^\dagger = 0} \tag{21.18}$$

or the *anticommutation relations*

$$\boxed{a_i^\dagger a_j^\dagger + a_j^\dagger a_i^\dagger = 0} \tag{21.19}$$

The possibility that for a specific kind of particle (21.18) holds for some creation operators and (21.19) for others need not be contemplated if there are no super-

selection rules operative that would restrict superpositions and the unitary transformations $\langle L_k | K_i \rangle$ in the one-particle theory to certain subspaces.

By taking the Hermitian adjoint of Eqs. (21.18) and (21.19), we obtain, respectively,

$$\boxed{a_i a_j - a_j a_i = 0} \tag{21.20}$$

and

$$\boxed{a_i a_j + a_j a_i = 0} \tag{21.21}$$

A similar argument can now be made for two operators a_i and $a_j^\dagger$, with $i \neq j$. Consider

$$a_i a_j^\dagger \Psi = \mu a_j^\dagger a_i \Psi$$

The unitary transformations (21.13) and (21.14) give for $i \neq j$,

$$(a_i a_j^\dagger - \mu a_j^\dagger a_i)\Psi = \sum_{k,\ell} \langle K_i | L_k \rangle \langle L_\ell | K_j \rangle (b_k b_\ell^\dagger - \mu b_\ell^\dagger b_k)\Psi = 0 \tag{21.22}$$

and this must be valid for any state. In satisfying this condition, we must again take the unitary constraint (20.17) into account, leading to the inference that for all values of k and ℓ, but subject to the restriction $k \neq \ell$, we must have

$$b_k b_\ell^\dagger - \mu b_\ell^\dagger b_k = 0 \tag{21.23}$$

and thus also

$$a_i a_j^\dagger - \mu a_j^\dagger a_i = 0 \quad \text{for} \quad i \neq j \tag{21.24}$$

If we substitute this back into Eq. (21.22), we get (for $i \neq j$)

$$\sum_k \langle K_i | L_k \rangle \langle K_k | L_i \rangle (b_k b_k^\dagger - \mu b_k^\dagger b_k)\Psi = 0$$

This equation is compatible with the unitarity condition (21.17) only if

$$b_k b_k^\dagger - \mu b_k^\dagger b_k = A \tag{21.25}$$

where the operator A is independent of the subscript k. If we now expand the expression $a_i a_i^\dagger - \mu a_i^\dagger a_i$ by use of the transformation relation (21.13) and (21.14) and apply the conditions (21.23) and (21.17), we conclude that

$$a_i a_i^\dagger - \mu a_i^\dagger a_i = A \tag{21.26}$$

Hence, the Hermitian operator A is the same for all pairs of creation and annihilation operators in any one-particle basis. By letting (21.26) act on the vacuum, we see from Eqs. (21.6) and (21.8) that $A|\mathbf{0}\rangle = |\mathbf{0}\rangle$. Hence, we may choose A to be the identity operator, $A = 1$:

$$a_i a_i^\dagger - \mu a_i^\dagger a_i = 1 \tag{21.27}$$

To complete the theoretical framework, we define the Hermitian operators N_i whose eigenvalues are the occupation numbers, n_i, of the particles for which K has the value K_i. The sum of these,

$$N = \sum_i N_i \tag{21.28}$$

must be the additive operator measuring the total number of particles, and this must be invariant under a change of the one-particle basis:

$$N = \sum_i N_i = \sum_k \tilde{N}_k$$

Other than the trivial identity operator, essentially only one additive operator is invariant under the transformations (21.13) and (21.14). It is

$$\sum_i a_i^\dagger a_i = \sum_k b_k^\dagger b_k \tag{21.29}$$

which is invariant by virtue of the unitarity of the transformation $\langle K_i | L_k \rangle$. The operator

$$\sum_i a_i a_i^\dagger = \sum_k b_k b_k^\dagger$$

is also invariant, but because of the commutation relation (21.27), it can be expressed in terms of (21.29) and the identity.

Exercise 21.2. Verify the invariance of the additive operator (21.29) under the unitary transformation (21.13), (21.14) (principle of unitary symmetry).

Thus, we expect N_i to have the form

$$N_i = x\, a_i^\dagger a_i + y\, 1$$

where x and y are real constants. We must require that $N_i|\mathbf{0}\rangle = 0$, hence, $y = 0$. Furthermore, for a one-particle state,

$$N_i \Psi_i^{(1)} = \Psi_i^{(1)}$$

From (21.7) and (21.6) it follows that $x = 1$. Thus, the occupation number operators are simply

$$\boxed{N_i = a_i^\dagger a_i} \tag{21.30}$$

These operators are positive definite and have nonnegative expectation values, as expected.

If the basis vectors (21.1) are to be eigenstates of N_i, it is easy to see from the defining relations (21.4) and (21.5) for the creation and annihilation operators that the following conditions must be met:

$$N_i a_k - a_k N_i = N_i a_k^\dagger - a_k^\dagger N_i = 0 \quad \text{for} \quad i \neq k \tag{21.31}$$

as well as

$$N_i a_i - a_i N_i = -a_i \tag{21.32}$$

and

$$N_i a_i^\dagger - a_i^\dagger N_i = a_i^\dagger \tag{21.33}$$

If the commutation relations that have so far been proved, (21.18)–(21.21), and (21.24), are employed to bring the operator a_k to the right in Eq. (21.31), we deduce:

$$(1 \mp \mu) N_i a_k = 0$$

for all $i \neq k$. The upper sign, or $\mu = +1$, corresponds to (21.18), and (21.20) and the lower sign, or $\mu = -1$, to (21.19) and (21.21).

We have thus arrived at the conclusion that, consistent with our fundamental assumptions, there are two and only two forms of quantum mechanics for identical particles:

Bose-Einstein Case

$$\boxed{\begin{aligned} a_k^\dagger a_\ell^\dagger - a_\ell^\dagger a_k^\dagger &= 0 \\ a_k a_\ell - a_\ell a_k &= 0 \\ a_k a_\ell^\dagger - a_\ell^\dagger a_k &= \delta_{k\ell}\, 1 \end{aligned}} \tag{21.34}$$

Fermi-Dirac Case

$$\boxed{\begin{aligned} a_k^\dagger a_\ell^\dagger + a_\ell^\dagger a_k^\dagger &= 0 \\ a_k a_\ell + a_\ell a_k &= 0 \\ a_k a_\ell^\dagger + a_\ell^\dagger a_k &= \delta_{k\ell}\, 1 \end{aligned}} \tag{21.35}$$

(In the literature, the brace $\{A, B\} \equiv AB + BA$ is often used to denote an anticommutator, similar to the bracket $[A, B] \equiv AB - BA$ for a commutator. For our purposes, little is to be gained by using the $\{\ \}$ notation, and so we prefer to write out anticommutators explicitly.)

The designations Bose-Einstein (B.E.) and Fermi-Dirac (F.D.) associated with the two classes of commutation relations will be seen to be justified when the statistical behavior of identical particles is examined. Characteristically, each particle species belongs to one or the other of these two kinds of *statistics*. We call the former species *bosons* since they obey Bose-Einstein statistics, and the latter *fermions*, since they obey Fermi-Dirac statistics. A connection between the statistics of a particle and the transformation of its states under spatial rotation is found in quantum field theory, leading to the conclusion that all particles with integral spin are bosons and all particles with half-integral spin are fermions.[1]

One immediate consequence of Eq. (21.35) is that, for fermions, $a_k^\dagger a_k^\dagger = 0$; i.e., there are no fermion states in which two or more particles share the same quantum number. This is the familiar expression of the *Pauli exclusion principle*.

Exercise 21.3. Prove that the occupation number operators N_i for different i commute for either Bose-Einstein or Fermi-Dirac statistics.

Exercise 21.4. From the commutation relations for a_i and $a_i^\dagger$, deduce that the operator $N_i = a_i^\dagger a_i$ has as eigenvalues all nonnegative integers in the case of Bose-Einstein statistics, but only 0 and 1 in the case of Fermi-Dirac statistics.

The operators $N_i = a_i^\dagger a_i$ determine the constants of proportionality in Eqs. (21.4) and (21.5) except for a phase factor. Indeed

$$\langle n_1, n_2, \ldots n_i \ldots | N_i | n_1, n_2, \ldots n_i \ldots \rangle = n_i = \langle n_1, n_2, \ldots n_i \ldots | a_i^\dagger a_i | n_1, n_2, \ldots n_i \ldots \rangle$$

Hence,

$$a_i | n_1, n_2, \ldots n_i \ldots \rangle = e^{i\alpha}\sqrt{n_i}\, | n_1, n_2, \ldots n_i - 1 \ldots \rangle \tag{21.36}$$

[1]Gross (1993), Chapter 7.

For Bose-Einstein statistics, we may set $\alpha = 0$. For Fermi-Dirac statistics, it is consistent with the anticommutation relations to choose $e^{i\alpha}$ to be equal to 1 if the number of occupied one-particle states with index less than i is even, and equal to -1 if this number is odd. Other phase assignments are possible, but they must always be consistent with the commutation and anticommutation relations.

Exercise 21.5. Prove the feasibility of the suggested phase assignments in Eq. (21.36).

To summarize the effect of the creation and annihilation operators, we omit all identifying subscripts and focus on the action of the typical operators a and $a^\dagger$ on the eigenstates $|n\rangle$ of the occupation number operator $N = a^\dagger a$, such that

$$a^\dagger a|n\rangle = n|n\rangle \tag{21.37}$$

We have

Bose-Einstein Case

$$a|n\rangle = \sqrt{n}|n\rangle \qquad a^\dagger|n\rangle = \sqrt{n+1}|n\rangle \tag{21.38}$$

Fermi-Dirac Case

$$\begin{aligned} a|0\rangle &= 0 & a^\dagger|0\rangle &= e^{-i\alpha}|1\rangle \\ a|1\rangle &= e^{i\alpha}|0\rangle & a^\dagger|1\rangle &= 0 \end{aligned} \tag{21.39}$$

For bosons (B.E. case) the structure of the theory is isomorphic to that of the harmonic oscillator, since the energy levels of the oscillator are spaced equally and can be interpreted as representing n quanta (or particles) of energy $\hbar\omega$. The oscillator Hamiltonian (10.70) is just

$$H = \hbar\omega\,(N + 1/2) \tag{21.40}$$

It links the second quantization for bosons to the operator ålgebra treatment of the oscillator in Section 10.6.

4. ***Dynamical Variables.*** Useful dynamical variables for a system of an arbitrary number of identical particles can now be constructed. The operator $\mathcal{K}$ that measures the total value of an *additive one-particle quantity* K, like the kinetic energy, is

$$\mathcal{K} = \sum_i K_i N_i = \sum_i K_i a_i^\dagger a_i \tag{21.41}$$

where a_i and $a_i^\dagger$ destroy and create particles with quantum number K_i. It is desirable to express the operator $\mathcal{K}$ in terms of an arbitrary set of creation and annihilation operators, b_i and $b_i^\dagger$, which are related to a_i and $a_i^\dagger$ by (21.13) and (21.14). Applying this transformation to (21.41) yields

$$\boxed{\mathcal{K} = \sum_{qr} b_q^\dagger b_r \langle L_q|K|L_r\rangle} \tag{21.42}$$

since

$$\sum_i \langle L_q|K_i\rangle K_i \langle K_i|L_r\rangle = \langle L_q|K|L_r\rangle \tag{21.43}$$

The operator $\mathcal{H}$ is the most general Hermitian operator that is bilinear in annihilation and creation operators.

An *additive two-particle operator*, like the mutual potential energy, can be used to define

$$\mathcal{V} = \frac{1}{2}\sum_{i\neq j} N_i N_j V_{ij} + \frac{1}{2}\sum_i N_i(N_i - 1)V_{ii} = \frac{1}{2}\sum_{ij}(N_i N_j - N_i\delta_{ij})V_{ij} \quad (21.44)$$

where we may assume, without loss of generality, that the numbers V_{ij} form a real symmetric matrix. It is easy to see that the *pair distribution operator* $P_{ij} = N_i N_j - \delta_{ij}N_i$, occurring in (21.44), may be written as

$$P_{ij} = N_i N_j - \delta_{ij}N_i = a_i^\dagger a_j^\dagger a_j a_i \quad (21.45)$$

for either the Bose-Einstein or Fermi-Dirac case. Hence, Eq. (21.44) may be written more concisely as

$$\mathcal{V} = \frac{1}{2}\sum_{ij} a_i^\dagger a_j^\dagger a_j a_i \, V_{ij} \quad (21.46)$$

After the transformation (21.13), (21.14) is applied, this operator acquires its generalized form

$$\boxed{\mathcal{V} = \frac{1}{2}\sum_{qrst} b_q^\dagger b_r^\dagger b_s b_t \langle qr|V|ts\rangle} \quad (21.47)$$

where

$$\sum_{ij}\langle L_q|K_i\rangle\langle K_i|L_t\rangle\langle L_r|K_j\rangle\langle K_j|L_s\rangle V_{ij} = \langle qr|V|ts\rangle \quad (21.48)$$

is the general two-particle matrix element. The order of quantum numbers in (21.48) implies the convention that q and t belong to one particle, and r and s to the other. From the reality and symmetry properties of the V_{ij} we deduce that the two-particle operator matrix elements have the property

$$\boxed{\langle qr|V|ts\rangle = \langle rq|V|st\rangle} \quad (21.49)$$

Figure 21.1 shows a diagram representation of the two-particle operator $b_q^\dagger b_r^\dagger b_s b_t\langle qr|V|ts\rangle$, which is a typical term in the interaction operator $\mathcal{V}$, (21.47).

Exercise 21.6. Verify Eq. (21.49).

Exercise 21.7. Show that the square of an additive one-particle operator is generally expressible as the sum of an additive two-particle operator and an additive one-particle operator.

Exercise 21.8. Prove from the commutation relations that

$$\langle \mathbf{0}|a_i a_j a_k^\dagger a_\ell^\dagger|\mathbf{0}\rangle = \delta_{jk}\delta_{i\ell} \pm \delta_{ik}\delta_{j\ell} \quad (21.50)$$

the sign depending on the statistics, B.E. (+) or F.D. (−). Also calculate the vacuum expectation value $\langle \mathbf{0}|a_h a_i a_j a_k^\dagger a_\ell^\dagger a_m^\dagger|\mathbf{0}\rangle$.

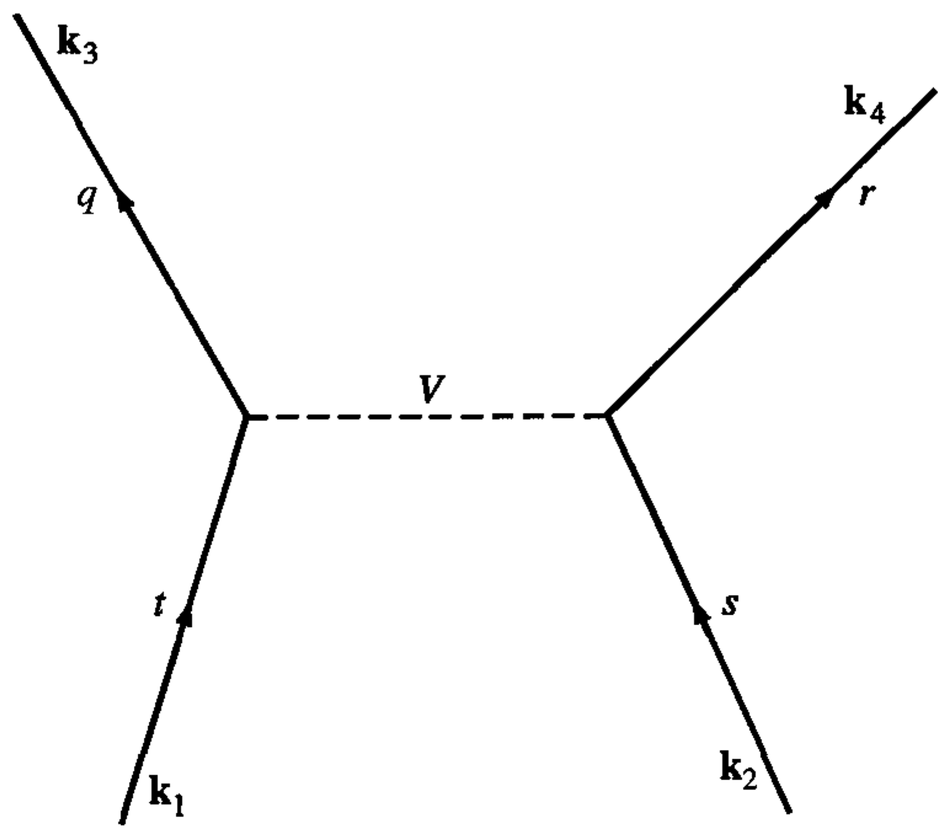


Figure 21.1. Diagram representing the two-particle matrix element $\langle qr|V|ts\rangle$. The vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ (in units of $\hbar$) are "in" momenta, and $\mathbf{k}_3$ and $\mathbf{k}_4$ are "out" momenta. (See Problems 1 and 2 in Chapter 21.)

5. ***The Continuous One-Particle Spectrum and Quantum Field Operators.*** So far in this chapter we have implicitly assumed that the eigenvectors of the complete set of commuting one-particle operators, from which the n-particle basis is constructed, are normalized to unity, corresponding to a discrete set of eigenvalues. Now we must consider a one-particle basis belonging to the continuous spectrum of a complete set of commuting operators symbolized by x, such that the one-particle eigenvectors are normalized as

$$\langle x', \sigma'|x'', \sigma''\rangle = \delta(x' - x'')\delta_{\sigma',\sigma''}$$

The continuous eigenvalues x' of x and the discrete eigenvalues σ' of σ label the basis vectors. The most common example of a continuous basis occurs if x is the ordinary position coordinate of a particle. The discrete variable σ may represent the z-component of the spin of the particle, as explained in Section 16.3.

The formalism of the preceding sections can be modified and transcribed without difficulty to be applicable to this case, but the occupation number representation is no longer useful, unless by a discretization procedure the coordinate space is artificially partitioned into a large number of small but finite cells.

It is customary to write $\psi_{\sigma'}(\mathbf{r}')$ for the annihilation operator, which destroys a particle with quantum number σ' at position $\mathbf{r}'$, instead of using the notation a_i which is employed in the discrete case. An operator like $\psi_{\sigma'}(\mathbf{r}')$, which depends on the position coordinates, is generally referred to as a *quantum field operator* or simply a *field*.

With this new notation, we now generalize the commutation relations (21.34) and (21.35) and postulate:

Bose-Einstein Case

$$\begin{aligned}
\psi_{\sigma'}(\mathbf{r}')\psi_{\sigma''}(\mathbf{r}'') - \psi_{\sigma''}(\mathbf{r}'')\psi_{\sigma'}(\mathbf{r}') &= 0 \\
\psi^\dagger_{\sigma'}(\mathbf{r}')\psi^\dagger_{\sigma''}(\mathbf{r}'') - \psi^\dagger_{\sigma''}(\mathbf{r}'')\psi^\dagger_{\sigma'}(\mathbf{r}') &= 0 \\
\psi_{\sigma'}(\mathbf{r}')\psi^\dagger_{\sigma''}(\mathbf{r}'') - \psi^\dagger_{\sigma''}(\mathbf{r}'')\psi_{\sigma'}(\mathbf{r}') &= \delta(\mathbf{r}' - \mathbf{r}'')\delta_{\sigma'\sigma''}
\end{aligned} \tag{21.51}$$

Fermi-Dirac Case

$$\begin{aligned}
\psi_{\sigma'}(\mathbf{r}')\psi_{\sigma''}(\mathbf{r}'') + \psi_{\sigma''}(\mathbf{r}'')\psi_{\sigma'}(\mathbf{r}') &= 0 \\
\psi^{\dagger}_{\sigma'}(\mathbf{r}')\psi^{\dagger}_{\sigma''}(\mathbf{r}'') + \psi^{\dagger}_{\sigma''}(\mathbf{r}'')\psi^{\dagger}_{\sigma'}(\mathbf{r}') &= 0 \\
\psi_{\sigma'}(\mathbf{r}')\psi^{\dagger}_{\sigma''}(\mathbf{r}'') + \psi^{\dagger}_{\sigma''}(\mathbf{r}'')\psi_{\sigma'}(\mathbf{r}') &= \delta(\mathbf{r}' - \mathbf{r}'')\delta_{\sigma'\sigma''}
\end{aligned} \tag{21.52}$$

The operator $\psi^{\dagger}_{\sigma'}(\mathbf{r}')\psi_{\sigma'}(\mathbf{r}')$ represents the density at the position $\mathbf{r}'$ of particles for which the discrete variable has the value σ', and

$$N = \sum_{\sigma'} \int \psi^{\dagger}_{\sigma'}(\mathbf{r}')\psi_{\sigma'}(\mathbf{r}')\, d^3r' \tag{21.53}$$

is the operator whose eigenvalues are the total number of particles.

For an additive one-particle operator, we have

$$\mathcal{K} = \sum_{\sigma'\sigma''} \int\int \psi^{\dagger}_{\sigma'}(\mathbf{r}')\langle \mathbf{r}'\sigma' | K | \mathbf{r}''\sigma''\rangle \psi_{\sigma''}(\mathbf{r}'')\, d^3r'\ d^3r'' \tag{21.54}$$

and for the common kind of additive two-particle operator,

$$\begin{aligned}
\mathcal{V} = \frac{1}{2} \sum_{\sigma'\sigma'\sigma''\sigma'''} \int\int\int \psi^{\dagger}_{\sigma'}(\mathbf{r}')\psi^{\dagger}_{\sigma''}(\mathbf{r}'')\psi_{\sigma'''}(\mathbf{r}''')\psi_{\sigma''''}(\mathbf{r}'''') \\
\times \langle \mathbf{r}'\sigma', \mathbf{r}''\sigma'' | V | \mathbf{r}''''\sigma'''', \mathbf{r}'''\sigma'''\rangle\, d^3r'\ d^3r''\ d^3r'''\ d^3r''''
\end{aligned} \tag{21.55}$$

An arbitrary n-particle state $\Psi^{(n)}$ can be conveniently represented by the function

$$\boxed{\psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \ldots, \mathbf{r}_n\sigma_n) = \frac{1}{\sqrt{n!}} \langle \mathbf{0} | \psi_{\sigma_1}(\mathbf{r}_1)\psi_{\sigma_2}(\mathbf{r}_2) \ldots \psi_{\sigma_n}(\mathbf{r}_n) | \Psi^{(n)}\rangle} \tag{21.56}$$

Evidently, with respect to permutations of the subscripts $1, 2, \ldots n$, ψ is symmetric in the Bose-Einstein case and antisymmetric in the Fermi-Dirac case. In the trivial case of a one-particle state, $n = 1$, Eq. (21.56) reduces to

$$\psi(\mathbf{r}\sigma) = \langle \mathbf{0} | \psi_{\sigma}(\mathbf{r}) | \Psi^{(1)}\rangle \equiv \langle \mathbf{r}\sigma | \Psi^{(1)}\rangle \tag{21.57}$$

which is simply the ordinary one-particle Schrödinger wave function with spin, representing the state $\Psi^{(1)}$.

The representation (21.56) of an n-particle state is convenient for many purposes. The function $\psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \ldots, \mathbf{r}_n\sigma_n)$, which is symmetric (B.E. case) or antisymmetric (F.D. case) under exchange of any two-particle labels, is the wave function of the n-particle system in the $3n$-dimensional Cartesian configuration space of the coordinates $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n$. Such a representation of a state of n identical particles implies a numerical labeling of the identical particles, but configuration space wave functions, which, under particle label exchange, are neither fully symmetric nor antisymmetric, are ruled out as not representing a physical state of n particles. The Fock representation, with its emphasis on occupation numbers of intrinsically indistinguishable one-particle properties, has no room for such physically inadmissible states and, therefore, provides a more economical description.[2]

The usefulness of the n-particle wave function (21.56) in configuration space is attributable to the circumstance that, if the operators $\mathbf{r}\sigma$ constitute a complete set

[2]For a more complete account, see Schweber (1961), Chapter 6.

of observables for a single particle, the function ψ as given by formula (21.56) determines the state $\Psi^{(n)}$ of n identical particles. The formula that inverts (20.56) is

$$\boxed{|\Psi^{(n)}\rangle = \sum_{\sigma_1\sigma_2\ldots} \int \frac{1}{\sqrt{n!}} \boldsymbol{\psi}^\dagger_{\sigma_n}(\mathbf{r}_n) \ldots \boldsymbol{\psi}^\dagger_{\sigma_2}(\mathbf{r}_2)\boldsymbol{\psi}^\dagger_{\sigma_1}(\mathbf{r}_1)|\mathbf{0}\rangle\psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \ldots, \mathbf{r}_n\sigma_n)\, d^3r_1 \ldots d^3r_n} \tag{21.58}$$

For its proof, the following lemma is required:

$$\begin{aligned} F_n &\equiv \langle \mathbf{0}|\boldsymbol{\psi}_{\sigma_1}(\mathbf{r}_1)\boldsymbol{\psi}_{\sigma_2}(\mathbf{r}_2) \ldots \boldsymbol{\psi}_{\sigma_n}(\mathbf{r}_n)\boldsymbol{\psi}^\dagger_{\sigma'_n}(\mathbf{r}'_n) \ldots \boldsymbol{\psi}^\dagger_{\sigma'_1}(\mathbf{r}'_1)|\mathbf{0}\rangle \\ &= \sum_{P_n} \operatorname{sgn}(P_n)\delta(\mathbf{r}_1 - \mathbf{r}'_1)\delta(\mathbf{r}_2 - \mathbf{r}'_2) \ldots \delta(\mathbf{r}_n - \mathbf{r}'_n)\delta_{\sigma_1\sigma'_1}\,\delta_{\sigma_2\sigma'_2} \ldots \delta_{\sigma_n\sigma'_n} \end{aligned} \tag{21.59}$$

where the sum is to be taken over all permutations P_n of the coordinates $\mathbf{r}_1\sigma_1$, $\mathbf{r}_2\sigma_2, \ldots \mathbf{r}_n\sigma_n$ (but not the primed variables) and $\operatorname{sgn}(P_n)$ is the sign of the permutation. For the Bose-Einstein case, $\operatorname{sgn}(P_n) = +1$; for the Fermi-Dirac case, $\operatorname{sgn}(P_n) = +1$ if the permutation P_n is even and $\operatorname{sgn}(P_n) = -1$ if it is odd.

Formula (21.59) is proved by induction. It is obviously valid for $n = 1$. Use of the commutation relations (20.51) or (20.52) gives

$$\begin{aligned} F_n = {} & \langle \mathbf{0}|\boldsymbol{\psi}_{\sigma_1}(\mathbf{r}_1) \ldots \boldsymbol{\psi}_{\sigma_{n-1}}(\mathbf{r}_{n-1})\boldsymbol{\psi}^\dagger_{\sigma'_{n-1}}(\mathbf{r}'_{n-1}) \ldots \boldsymbol{\psi}^\dagger_{\sigma'_1}(\mathbf{r}'_1)|\mathbf{0}\rangle\delta(\mathbf{r}_n - \mathbf{r}'_n)\,\delta_{\sigma_n\sigma'_n} \\ & \pm \langle \mathbf{0}|\boldsymbol{\psi}_{\sigma_1}(\mathbf{r}_1) \ldots \boldsymbol{\psi}_{\sigma_{n-1}}(\mathbf{r}_{n-1})\boldsymbol{\psi}^\dagger_{\sigma'_n}(\mathbf{r}'_n)\boldsymbol{\psi}_{\sigma_n}(\mathbf{r}_n)\boldsymbol{\psi}^\dagger_{\sigma'_{n-1}}(\mathbf{r}'_{n-1}) \ldots \boldsymbol{\psi}^\dagger_{\sigma'_1}(\mathbf{r}'_1)|\mathbf{0}\rangle \end{aligned}$$

By continuing to move the operator $\boldsymbol{\psi}_{\sigma_n}(\mathbf{r}_n)$ to the right, step by step, and by applying Eq. (21.59) for $n - 1$, we eventually obtain

$$\begin{aligned} F_n = {} & \sum_{P_{n-1}} \operatorname{sgn}(P_{n-1})\delta(\mathbf{r}_1 - \mathbf{r}'_1) \ldots \delta(\mathbf{r}_{n-1} - \mathbf{r}'_{n-1})\delta(\mathbf{r}_n - \mathbf{r}'_n) \\ & \pm \sum_{P_{n-1}} \operatorname{sgn}(P_{n-1})\delta(\mathbf{r}_1 - \mathbf{r}'_1) \ldots \delta(\mathbf{r}_{n-2} - \mathbf{r}'_{n-2})\delta(\mathbf{r}_{n-1} - \mathbf{r}'_n)\delta(\mathbf{r}_n - \mathbf{r}'_{n-1}) \\ & + \sum_{P_{n-1}} \operatorname{sgn}(P_{n-1})\delta(\mathbf{r}_1 - \mathbf{r}'_1) \ldots \delta(\mathbf{r}_{n-2} - \mathbf{r}'_{n-1})\delta(\mathbf{r}_{n-1} - \mathbf{r}'_n)\delta(\mathbf{r}_n - \mathbf{r}'_{n-2}) \\ & \pm \ldots \end{aligned}$$

which can be contracted into the right-hand side of Eq. (21.59). (The Kronecker deltas carrying the spin variable have been suppressed for brevity in the last equation.)

Armed with the generalized orthogonality relation (21.59), we can now derive Eq. (21.58) from (21.56) if we assume that a general n-particle state can be represented by

$$|\Psi^{(n)}\rangle = \frac{1}{\sqrt{n!}} \sum_{\sigma_1\sigma_2\ldots} \int d^3r_1 \ldots d^3r_n \boldsymbol{\psi}^\dagger_{\sigma_n}(\mathbf{r}_n) \ldots \boldsymbol{\psi}^\dagger_{\sigma_2}(\mathbf{r}_2)\boldsymbol{\psi}^\dagger_{\sigma_1}(\mathbf{r}_1)|\mathbf{0}\rangle f(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \ldots, \mathbf{r}_n\sigma_n) \tag{21.60}$$

If this expression is substituted for Ψ in Eq. (21.56) and the integrations are carried out, using lemma (21.59),

$$\psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \ldots, \mathbf{r}_n\sigma_n) = \frac{1}{n!} \sum_{P_n} \operatorname{sgn}(P_n) f(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \ldots, \mathbf{r}_n\sigma_n)$$

results. On the other hand, owing to the commutation relations for the creation operators, (21.60) can be written as

$$|\Psi^{(n)}\rangle = \frac{1}{\sqrt{n!}} \sum_{\sigma_1\sigma_2\ldots} \int d^3r_1 \ldots d^3r_n \psi^\dagger_{\sigma_n}(\mathbf{r}_n) \ldots \psi^\dagger_{\sigma_2}(\mathbf{r}_2)\psi^\dagger_{\sigma_1}(\mathbf{r}_1)|\mathbf{0}\rangle$$
$$\times \frac{1}{n!} \sum_{P_n} \operatorname{sgn}(P_n) f(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \ldots, \mathbf{r}_n\sigma_n)$$

Hence, Eq. (21.58) follows.

By taking the inner product of $\langle\Psi^{(n)}|$ with (21.58) and using the complex conjugate of Eq. (21.56), we get

$$\langle\Psi^{(n)}|\Psi^{(n)}\rangle = \sum_{\sigma_1\sigma_2\ldots} \int |\psi(\mathbf{r}_1\sigma_1, \ldots, \mathbf{r}_n\sigma_n)|^2 \, d^3r_1 \ldots d^3r_n \tag{21.61}$$

This equation establishes the normalization of the wave function ψ, which can thus be made the basis of a straightforward probability interpretation in configuration space: The particles are labeled numerically, as if they were distinguishable, 1 through n, and $|\psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \ldots, \mathbf{r}_n\sigma_n)|^2$ is understood as the probability density of finding particle 1 at position $\mathbf{r}_1$ with spin σ_1, particle 2 at position $\mathbf{r}_2$ with spin σ_2, and so forth. This description has been the traditional approach to the quantum mechanics of n particles. For the special case of $n = 2$, we already encountered it in Sections 15.4 and 18.8 (helium atom).

Expectation values (and off-diagonal matrix elements) of operators can be expressed either in the language of creation and annihilation operators or in terms of wave functions in multidimensional space. Thus, if $K(\mathbf{r}, \mathbf{p})$ is a one-particle operator, which is a function of the coordinate $\mathbf{r}$ and the conjugate momentum $\mathbf{p}$, the expectation value of the sum operator K in a normalized n-particle state $\Psi^{(n)}$ is

$$\begin{aligned}(\Psi^{(n)}, \mathcal{K}\Psi^{(n)}) &= \langle\Psi^{(n)}|\mathcal{K}|\Psi^{(n)}\rangle \\ &= \left(\Psi^{(n)}, \sum_{\sigma'} \int \psi^\dagger_{\sigma'}(\mathbf{r}')K\left(\mathbf{r}', \frac{\hbar}{i}\nabla'\right)\psi_{\sigma'}(\mathbf{r}') \, d^3r' \, \Psi^{(n)}\right) \\ &= \sum_{j=1}^{n} \sum_{\sigma_1\sigma_2\ldots} \int \psi^*(\mathbf{r}_1\sigma_1, \ldots \mathbf{r}_n\sigma_n)K\left(\mathbf{r}_j, \frac{\hbar}{i}\nabla_j\right) \psi(\mathbf{r}_1\sigma_1, \ldots \mathbf{r}_n\sigma_n) \, d^3r_1 \ldots d^3r_n\end{aligned} \tag{21.62}$$

Similarly, if the two-particle interaction operator V is *local* in coordinate space and spin-independent, such that

$$\langle\mathbf{r}'\sigma', \mathbf{r}''\sigma''|V|\mathbf{r}''''\sigma'''', \mathbf{r}'''\sigma'''\rangle = V(\mathbf{r}', \mathbf{r}'')\delta(\mathbf{r}' - \mathbf{r}'''')\delta(\mathbf{r}'' - \mathbf{r}''')\delta_{\sigma'\sigma''''}\delta_{\sigma''\sigma'''} \tag{21.63}$$

we may write for an n-particle state,

$$\begin{aligned}(\Psi^{(n)}, \mathcal{V}\Psi^{(n)}) &= \langle\Psi^{(n)}|\mathcal{V}|\Psi^{(n)}\rangle \\ &= \sum_{i>j=1}^{n} \sum_{\sigma_1\sigma_2\ldots} |\psi(\mathbf{r}_1\sigma_1, \ldots, \mathbf{r}_n\sigma_n)|^2 V(\mathbf{r}_i, \mathbf{r}_j) \, d^3r_1 \ldots d^3r_n\end{aligned} \tag{21.64}$$

Exercise 21.9. Prove Eqs. (21.62) and (21.64).

6. ***Quantum Dynamics of Identical Particles.*** Up to this point in this chapter, nothing has been said about changes in time. Following the program of Chapter 14, we now assume that the time development is governed by a Hermitian operator $\mathcal{H}$,

the *Hamiltonian* of the system. For simplicity, we also assume that $\mathcal{H}$ is not explicitly dependent on the time t.

We must decide which picture to use, such as Schrödinger, Heisenberg, or interaction. In this section we will first use the *Heisenberg picture*, because it is instructive to see the similarity of the equation of motion for the field operator $\psi(\mathbf{r}, t)$ with the wave equation in the one-particle theory. Eventually, we return to the Schrödinger picture to derive the equation of motion for the wave function in n-particle configuration space.

All physical observables can be constructed from creation and annihilation operators. It is, therefore, of interest to determine the temporal behavior of the fundamental building blocks and to derive the equation of motion for some arbitrary annihilation operator $a_j(t)$. Since, in what follows, all operators will be evaluated at time t, we will, for brevity, omit reference to the time dependence in some of the ancillary equations. The commutation relations are valid at all times, provided that the two factors are taken *at the same time*.

In order to obtain an equation for the time development of a_j, it is helpful to make explicit assumptions about the structure of the Hamiltonian operator. We cover many important applications by making the physical assumption that $\mathcal{H}$ preserves the total number of particles and that it is composed only of additive one- and two-particle operators. The most general form of $\mathcal{H}$ is thus assumed to be

$$\mathcal{H} = \sum_{k\ell} a_k^\dagger \langle k|H_0|\ell\rangle a_\ell + \frac{1}{2}\sum_{qrst} a_q^\dagger a_r^\dagger \langle qr|V|ts\rangle a_s a_t \tag{21.65}$$

where H_0 is a one-particle operator, such as the kinetic energy plus the potential energy in an external field, and V is a two-particle interaction operator.

To derive the equation of motion for a_j, we must evaluate the commutators

$$[a_j, a_k^\dagger a_\ell] \quad \text{and} \quad [a_j, a_q^\dagger a_r^\dagger a_s a_t]$$

This is easily done by using the commutation relations for creation and annihilation operators. The results, after some straightforward algebra, are

$$[a_j, a_k^\dagger a_\ell] = a_\ell \delta_{kj} \tag{21.66}$$

and

$$[a_j, a_q^\dagger a_r^\dagger a_s a_t] = a_r^\dagger a_s a_t \delta_{qj} + a_q^\dagger a_t a_s \delta_{rj} \tag{21.67}$$

It is remarkable that these relations hold for both Bose-Einstein and Fermi-Dirac statistics.

Exercise 21.10. Prove Eqs. (21.66) and (21.67) for both kinds of statistics.

Using these identities, we obtain in the Heisenberg picture:

$$i\hbar \frac{da_j}{dt} = [a_j, \mathcal{H}] = \sum_\ell \langle j|H_0|\ell\rangle a_\ell + \frac{1}{2}\sum_{rst} a_r^\dagger a_s a_t \langle jr|V|ts\rangle + \frac{1}{2}\sum_{qst} a_q^\dagger a_t a_s \langle qj|V|ts\rangle$$

$$= \sum_\ell \langle j|H_0|\ell\rangle a_\ell + \frac{1}{2}\sum_{rst} a_r^\dagger a_s a_t (\langle jr|V|ts\rangle + \langle rj|V|st\rangle)$$

Finally, the symmetry relation (21.49) gives, for *both statistics*, the desired equation of motion:

$$\boxed{i\hbar \frac{da_j(t)}{dt} = \sum_\ell \langle j|H_0|\ell\rangle a_\ell(t) + \sum_{qrs} a_q^\dagger(t) a_r(t) a_s(t) \langle jq|V|sr\rangle} \tag{21.68}$$

A case of special interest arises if we work in a representation in which the two-particle interaction matrix element is diagonal:

$$\langle jq|V|sr\rangle = V_{jq}\delta_{js}\delta_{qr} \qquad (V_{jq} = V_{qj}) \tag{21.69}$$

If (21.69) holds, the equation of motion for the annihilation operator is

$$\boxed{i\hbar \frac{da_j(t)}{dt} = \sum_{\ell} \langle j|H_0|\ell\rangle a_\ell(t) + \sum_{q} a_q^\dagger(t)a_q(t)a_j(t)V_{jq}} \tag{21.70}$$

To examine the form that this equation takes if the one-particle basis is the coordinate representation, supplemented by the discrete variable σ, we assume that the interaction energy is local and spin-independent, as in (21.63). Equation (21.70) can then be transcribed by using σ and $\mathbf{r}$ to denote the one-particle quantum numbers as labels on the operators $\boldsymbol{\psi}_\sigma(\mathbf{r})$ and $\boldsymbol{\psi}_\sigma^\dagger(\mathbf{r})$, replacing a_j and $a_j^\dagger$ labeled by their subscripts.

This transcription results in

$$\begin{aligned} i\hbar \frac{\partial \boldsymbol{\psi}_\sigma(\mathbf{r}, t)}{\partial t} = &\sum_{\sigma'} \int \langle \mathbf{r}\sigma|H_0|\mathbf{r}'\sigma'\rangle \boldsymbol{\psi}_{\sigma'}(\mathbf{r}', t)\, d^3r' \\ &+ \sum_{\sigma'} \int V(\mathbf{r}, \mathbf{r}')\boldsymbol{\psi}_{\sigma'}^\dagger(\mathbf{r}', t)\, \psi_{\sigma'}\,(\mathbf{r}', t)\, d^3r'\, \boldsymbol{\psi}_\sigma(\mathbf{r}, t) \end{aligned} \tag{21.71}$$

where $V(\mathbf{r}, \mathbf{r}') = V(\mathbf{r}', \mathbf{r})$. If, as is often the case, the one-particle operator H_0 is a function only of coordinates and conjugate momenta so that

$$\langle \mathbf{r}\sigma|H_0|\mathbf{r}'\sigma'\rangle = H_0\left(\mathbf{r}, \frac{\hbar}{i}\nabla\right)\delta(\mathbf{r} - \mathbf{r}')\delta_{\sigma\sigma'} \tag{21.72}$$

Eq. (21.71) can be rewritten as an integral-differential equation in ordinary space:

$$\boxed{\begin{aligned} i\hbar \frac{\partial \boldsymbol{\psi}_\sigma(\mathbf{r}, t)}{\partial t} = &H_0\left(\mathbf{r}, \frac{\hbar}{i}\nabla\right)\boldsymbol{\psi}_\sigma(\mathbf{r}, t) \\ &+ \sum_{\sigma'} \int V(\mathbf{r}, \mathbf{r}')\boldsymbol{\psi}_{\sigma'}^\dagger(\mathbf{r}', t)\boldsymbol{\psi}_{\sigma'}(\mathbf{r}', t)\, d^3r'\, \boldsymbol{\psi}_\sigma(\mathbf{r}, t) \end{aligned}} \tag{21.73}$$

Again, the universality of this equation, applicable to both bosons and fermions, deserves to be emphasized.

If there is no interaction between the particles, (21.73) reduces to

$$i\hbar \frac{\partial \boldsymbol{\psi}_\sigma(\mathbf{r}, t)}{\partial t} = H_0\left(\mathbf{r}, \frac{\hbar}{i}\nabla\right)\boldsymbol{\psi}_\sigma(\mathbf{r}, t) \tag{21.74}$$

which, in form, is identical with the one-particle wave equation—the time-dependent Schrödinger equation. However, it must be remembered that here $\boldsymbol{\psi}_\sigma(\mathbf{r}, t)$ is a *field operator* and not an ordinary *wave function*. Equation (21.74) can be regarded as the *quantized* version of the wave equation (3.1) when the latter is interpreted as a *classical* field equation. The annihilation operator corresponding to the coordinate $\mathbf{r}$ was denoted by $\boldsymbol{\psi}(\mathbf{r})$, rather than some other symbol, in the foreknowledge of its ultimate identification with the quantized Schrödinger field. This is the point of contact between *quantum field theory* and the many-body formulation developed in this chapter, traditionally misnamed *second quantization*.

The interaction term in Eq. (21.73) causes the equation of motion in the Hei-

senberg picture to be nonlinear. Since it involves an integral over all space, it can be interpreted somewhat casually as describing the effect of all the particles on a single one. The effective potential

$$\sum_{\sigma'} \int V(\mathbf{r}, \mathbf{r}')\psi^{\dagger}_{\sigma'}(\mathbf{r}', t)\psi_{\sigma'}(\mathbf{r}', t)\, d^3r' = \sum_{\sigma'} \int V(\mathbf{r}, \mathbf{r}')\rho_{\sigma'}(\mathbf{r}', t)\, d^3r'$$

can be calculated only if the solution of the equation of motion is already known, thus suggesting an iteration procedure to generate a self-consistent solution. Such techniques for solving the many-body problem are, indeed, frequently applied (see Chapter 22).

From the equation of motion (21.73) for the field operators, we can now derive the wave equation in configuration space for an n-particle system. The time-dependent wave function is defined by introducing the time into (21.56) and writing

$$\psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \ldots, \mathbf{r}_n\sigma_n; t) = \frac{1}{\sqrt{n!}} \langle \mathbf{0} | \psi_{\sigma_1}(\mathbf{r}_1, t)\psi_{\sigma_2}(\mathbf{r}_2, t) \ldots \psi_{\sigma_n}(\mathbf{r}_n, t) | \Psi^{(n)} \rangle \tag{21.75}$$

We apply $i\hbar \dfrac{\partial}{\partial t}$ to both sides of this equation and replace the time derivatives of the field operators by the expression on the right-hand side of (21.73). The noninteracting part of the Hamiltonian is easily seen to lead to a sum of n separate terms in the wave equation, one for each particle. The interaction term is reduced by use of the commutation or anticommutation relations for the field operators, moving $\psi^{\dagger}_{\sigma}(\mathbf{r})$ to the left in successive steps and recognizing that $\langle \mathbf{0} | \psi^{\dagger}_{\sigma}(\mathbf{r}) = 0$. When the permutation properties of the wave function are taken into account, the wave equation in configuration space is obtained as

$$i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \ldots \mathbf{r}_n\sigma_n; t) = \left[\sum_{j=1}^{n} H_0\left(\mathbf{r}_j, \frac{\hbar}{i} \nabla_j\right) + \frac{1}{2} \sum_{\substack{j,\ell=1 \\ (j \neq \ell)}}^{n} V(\mathbf{r}_j, \mathbf{r}_\ell) \right] \psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \ldots \mathbf{r}_n\sigma_n; t) \tag{21.76}$$

This equation has the form expected for any n-particle configuration-space wave equation. The permutation symmetry of $\psi(\mathbf{r}_1\sigma_1, \mathbf{r}_2\sigma_2, \ldots, \mathbf{r}_n\sigma_n)$ is conserved as a function of time.

Exercise 21.11. Complete the steps in the derivation of (21.76) from (21.75).

Problems

1. (a) Show that if $V(r)$ is a two-particle interaction that depends only on the distance r between the particles, the matrix element of the interaction in the **k**-representation may be reduced to

$$\langle \mathbf{k}_3\mathbf{k}_4 | V | \mathbf{k}_1\mathbf{k}_2 \rangle = \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}_3 - \mathbf{k}_4) \frac{1}{(2\pi)^3} \int V(r)\, e^{-i\mathbf{q}\cdot\mathbf{r}}\, d^3r$$

where $\hbar\mathbf{q}$ is the momentum transfer $\hbar(\mathbf{k}_3 - \mathbf{k}_1)$.

(b) For this interaction, show that the mutual potential energy operator is

$$\mathcal{V} = \frac{1}{2}\iiint d^3k_1\, d^3k_2\, d^3q\, \boldsymbol{\phi}^\dagger(\mathbf{k}_1 + \mathbf{q})\boldsymbol{\phi}^\dagger(\mathbf{k}_2 - \mathbf{q})\boldsymbol{\phi}(\mathbf{k}_2)\boldsymbol{\phi}(\mathbf{k}_1)F(\mathbf{q})$$

where $F(\mathbf{q})$ is the Fourier transform of the displacement-invariant interaction.

2. Show that the diagonal part of the interaction operator $\mathcal{V}$, found in Problem 1 in the **k**-representation, arises from momentum transfers $\mathbf{q} = 0$ and $\mathbf{q} = \mathbf{k}_2 - \mathbf{k}_1$, respectively. Write down the two interaction terms and identify them as *direct* ($\mathbf{q} = 0$) and *exchange* ($\mathbf{q} = \mathbf{k}_2 - \mathbf{k}_1$) interactions. Draw the corresponding diagrams (Figure 21.1).

3. In the **k**-representation, calculate the matrix element of the interaction in Problem 1 for the screened Coulomb potential $V_0 e^{-\alpha r}/\alpha r$ and plot it as a function of q. For bosons and fermions, construct the corresponding two-particle interaction operator $\mathcal{V}$ for identical particles in terms of the creation and annihilation operators in **k**-space.

4. Defining the momentum space annihilation operator[3]

$$\boldsymbol{\phi}(\mathbf{p}) = \int \langle \mathbf{p}|\mathbf{r}\rangle \boldsymbol{\psi}(\mathbf{r})\, d^3r$$

derive the commutation (or anticommutation) relations for $\boldsymbol{\phi}(\mathbf{p})$ and $\boldsymbol{\phi}^\dagger(\mathbf{p})$. For the Bose-Einstein case, show that the mixed commutator of field operators in coordinate and momentum space is

$$[\boldsymbol{\phi}(\mathbf{p}), \boldsymbol{\psi}^\dagger(\mathbf{r})] = \langle \mathbf{p}|\mathbf{r}\rangle$$

5. In the second-quantization formalism, define the additive position and total momentum operators

$$\mathbf{r} = \int \boldsymbol{\psi}^\dagger(\mathbf{r})\mathbf{r}\boldsymbol{\psi}(\mathbf{r})\, d^3r \quad \text{and} \quad \mathbf{p} = \int \boldsymbol{\phi}^\dagger(\mathbf{p})\mathbf{p}\boldsymbol{\phi}(\mathbf{p})\, d^3p$$

and prove that for bosons their commutator is

$$[\mathbf{r}, \mathbf{p}] = i\hbar N\, \mathbf{1}$$

where N is the operator representing the total number of particles. Derive the Heisenberg uncertainty relation for position and momentum of a system of bosons, and interpret the result.

6. Local particle and current density operators at position **r** are defined in the second-quantization formalism as

$$\rho(\mathbf{r}) = \int \boldsymbol{\psi}^\dagger(\mathbf{r}')\delta(\mathbf{r}' - \mathbf{r})\boldsymbol{\psi}(\mathbf{r}')\, d^3r'$$

and

$$j(\mathbf{r}) = \frac{\hbar}{2mi}\int d^3r'\, \boldsymbol{\psi}^\dagger(\mathbf{r}')[\nabla'\delta(\mathbf{r}' - \mathbf{r}) + \delta(\mathbf{r}' - \mathbf{r})\nabla']\boldsymbol{\psi}(\mathbf{r}')$$

(a) Show that the expectation values of these operators for one-particle states are the usual expressions.

(b) Derive the formulas for the operators $\rho(\mathbf{r})$ and $j(\mathbf{r})$ in the momentum representation.

7. Two identical bosons or fermions in a state

$$\Psi^{(2)} = A\sum_{ij} c_i d_j a_j^\dagger a_i^\dagger \Psi^{(0)} = A\sum_j d_j a_j^\dagger \sum_i c_i a_i^\dagger |\mathbf{0}\rangle$$

[3]For simplicity in Problems 4–6 we suppress any spin reference to spin variables.

are said to be *uncorrelated* (except for the effect of statistics). If $\sum |c_i|^2 = \sum |d_i|^2 = 1$, determine the normalization constant A in terms of the sum $S = \sum c_i^* d_i$.

(a) In this state, work out the expectation value of an additive one-particle operator in terms of the one-particle amplitudes c_i and d_i and the matrix elements $\langle i|K|j\rangle$.

(b) Show that if $S = 0$, the expectation value is the same as if the two particles with amplitudes c_i and d_i were distinguishable.

(c) Work out the expectation value of a diagonal interaction operator in terms of c_i, d_i, and the matrix elements $\langle ij|K|k\ell\rangle = V_{ij}\delta_{ik}\delta_{j\ell}$. Show that the result is the same as for distinguishable particles if the states of the two particles do not overlap, i.e., if $c_i d_i = 0$ for all i.

8. A state of n identical particles (bosons or fermions) is denoted by $|\Psi^{(n)}\rangle$.

For $n = 1$, the probability of finding the particle in the one-particle basis state i is the expectation value $\langle\Psi^{(1)}|N_i|\Psi^{(1)}\rangle$. (See Exercise 21.1.)

(a) For $n = 2$, prove that the probability of finding both particles in the one-particle basis state i is the expectation value of $N_i(N_i - 1)/2$.

(b) For $n = 3$, obtain the function of N_i whose expectation value is the probability of finding all three particles in the same basis state i.

(c) For $n = 2$, show that the expectation value of $N_i N_j$ is the probability of finding the two particles in two different basis states, $i \neq j$. Prove that the probability of finding one particle in basis state i and the other particle *not* in basis state i is the expectation value of $N_i(2 - N_i)$.

CHAPTER 22

Applications to Many-Body Systems

In this chapter, "second quantization" as a unifying concept of many-particle physics will be illustrated by several applications. We return to the coupling of angular momenta and present the Hartree-Fock self-consistent field method, leaving detailed discussion of many-body problems in atomic, condensed-matter, and nuclear physics to the specialized literature.[1] The thermal distribution functions for the ideal Bose-Einstein and Fermi-Dirac gas will be derived directly from the commutation and anticommutation relations for the creation and annihilation operators for the two species.

1. ***Angular Momentum in a System of Identical Particles.*** An important example of an observable in a system of identical particles is the angular momentum operator, which according to (21.41) is the additive one-particle operator

$$\mathcal{J} = \sum_{\alpha} \sum_{j} \sum_{mm'} a^{\dagger}_{jm'\alpha} a_{jm\alpha} \langle jm' | \mathbf{J} | jm \rangle \tag{22.1}$$

The one-particle basis is characterized by the angular momentum quantum numbers j and m, as defined in Section 17.4, and α stands for all remaining quantum numbers needed to specify the basis. The total angular momentum operator (22.1) owes its simple structure to the absence of off-diagonal matrix elements of $\mathbf{J}$ with respect to j and α.

The operator $\mathcal{J}^2 = \mathcal{J} \cdot \mathcal{J}$ is not just the sum of the $\mathbf{J}^2$ for the individual particles but contains terms that couple two particles. Thus, it serves as an example of an additive *two-particle* operator. Since it conserves the number of particles, annihilating one and creating one, the operator $\mathcal{J}$ commutes with the total number-of-particles operator, N.

Exercise 22.1. Exhibit the two-particle matrix elements of the square of the total angular momentum explicitly. (See also Exercise 21.7.)

The *two-particle* states in which $\mathcal{J}_z$ and $\mathcal{J}^2$ have the sharp values $M\hbar$ and $J(J+1)\hbar^2$ are readily constructed by the use of the Clebsch-Gordan coefficients defined in (17.52):

$$\Psi^{(2)}_{JM} = C \sum_{m_1 m_2} a^{\dagger}_{j_2 m_2 \alpha_2} a^{\dagger}_{j_1 m_1 \alpha_1} \langle j_1 j_1 m_1 m_2 | j_1 j_2 JM \rangle \Psi^{(0)} \tag{22.2}$$

where $\Psi^{(0)} = |\mathbf{0}\rangle$ is the vacuum state. Since the one-particle state $a^{\dagger}_{jm\alpha}|\mathbf{0}\rangle$ is normalized to unity, (22.2) completely parallels expression (17.52). The normalization constant $C = 1$, unless $\alpha_1 = \alpha_2$ and $j_1 = j_2$.

[1]For further study see Koltun and Eisenberg (1988) and, at a more advanced level, Fetter and Walecka (1971). Thouless (1961) emphasizes models of many-body systems that are exactly soluble.

The expression (22.2) remains an eigenvector of $\mathcal{J}_z$ and $\mathcal{J}^2$ even if $\alpha_1 = \alpha_2$ and $j_1 = j_2 = j$, but the normalization is altered. The symmetry relation (17.61) permits us to rewrite (22.2) in the form

$$\Psi^{(2)}_{JM} = \frac{1}{2}\,[1 \pm (-1)^{J-2j}]C \sum_{m_1 m_2} a^\dagger_{jm_2\alpha_2} a^\dagger_{jm_1\alpha_1} \langle jjm_1m_2 | jjJM\rangle \Psi^{(0)} \tag{22.3}$$

with the upper sign applicable to bosons and the lower sign to fermions. Hence, the angular momenta of two identical bosons (fermions), which share all one-particle quantum numbers except m, cannot couple to a state for which $J - 2j$ is an odd (even) number. If the usual connection between spin and statistics is assumed and bosons (fermions) have integral (half-integral) spin, odd J values of the total angular momentum cannot occur for two alike bosons or fermions with the same α and j.

The value of the normalization constant C in (22.3) may be determined by requiring $(\Psi^{(2)}_{JM}, \Psi^{(2)}_{JM}) = 1$. The unitarity condition (17.59) readily yields the value $C = 1/\sqrt{2}$, so that if $\alpha_1 = \alpha_2$ and $j_1 = j_2 = j$,

$$\boxed{\Psi^{(2)}_{JM} = \frac{1}{\sqrt{2}} \sum_{m_1 m_2} a^\dagger_{jm_2\alpha} a^\dagger_{jm_1\alpha} \langle jjm_1m_2 | jjJM\rangle \Psi^{(0)}} \tag{22.4}$$

Exercise 22.2. Verify the normalization (22.4).

Exercise 22.3. Construct explicitly in terms of states of the form $a^\dagger_{jm_2\alpha} a^\dagger_{jm_1\alpha}|\mathbf{0}\rangle$ the total angular momentum eigenstates for two neutrons in the configurations $(p_{1/2})^2$ and $(p_{3/2})^2$. How would the angular momentum eigenstates look if the two particles were a neutron and a proton but otherwise had the same quantum numbers as before?

Exercise 22.4. Show that if two identical particles with the same quantum numbers α and with angular momentum j couple to zero total angular momentum, the resulting *pair state* is, in an obviously simplified notation,

$$\Psi^{(2)}_{00} = [2(2j+1]^{-1/2} \sum_{m=-j}^{j} (-1)^m a^\dagger_m a^\dagger_{-m} |\mathbf{0}\rangle \tag{22.5}$$

2. ***Angular Momentum and Spin One-Half Boson Operators.*** If we postulate a fictitious *boson* with spin one-half, and with no other dynamical properties, the total angular momentum operator (22.1) for a system of identical particles of this kind takes the form

$$\mathcal{J} = \frac{\hbar}{2}\,(a^\dagger_{1/2} \quad a^\dagger_{-1/2})\boldsymbol{\sigma}\begin{pmatrix} a_{1/2} \\ a_{-1/2}\end{pmatrix} \tag{22.6}$$

Here the creation operators for the two spin states, $m = +1/2$ and $-1/2$ are denoted simply by $a^\dagger_{1/2}$ and $a^\dagger_{-1/2}$. Equation (22.6) may be decomposed into

$$\mathcal{J}_+ = \hbar a^\dagger_{1/2} a_{-1/2} \qquad \mathcal{J}_- = \hbar a^\dagger_{-1/2} a_{1/2} \qquad \mathcal{J}_z = \frac{\hbar}{2}(a^\dagger_{1/2}a_{1/2} - a^\dagger_{-1/2}a_{-1/2}) \tag{22.7}$$

in agreement with our expectations for $\mathcal{J}_\pm$ as the raising (lowering) operator that changes the state $|JM\rangle$ into $|JM\pm1\rangle$. Using the boson commutation relations, we derive from (22.7) the relation

$$\begin{aligned}\mathcal{J}^2 &= \frac{\hbar^2}{4}(a^\dagger_{1/2}a_{1/2} + a^\dagger_{-1/2}a_{-1/2})^2 + \frac{\hbar^2}{2}(a^\dagger_{1/2}a_{1/2} + a^\dagger_{-1/2}a_{-1/2}) \\ &= \frac{\hbar^2}{4}N^2 + \frac{\hbar^2}{2}N = \hbar^2\frac{N}{2}\left(\frac{N}{2}+1\right)\end{aligned} \tag{22.8}$$

Hence, the state with a total number $n = 2J$ of identical spin one-half bosons is an eigenstate of $\mathcal{J}^2$ with eigenvalue $J(J+1)\hbar^2$, where J is either integral or half-integral. The simultaneous eigenstates of the occupation number operators $N_+ = a^\dagger_{1/2}a_{1/2}$ (number of ''spin up'' bosons) and $N_- = a^\dagger_{-1/2}a_{-1/2}$ (number of ''spin down'' bosons) are also simultaneous eigenstates of $\mathcal{J}^2$ and $\mathcal{J}_z = \hbar(N_+ - N_-)/2$. The eigenvalues, $n_\pm$, of the occupation number operators are determined by the relations

$$n = n_+ + n_- = 2J \qquad n_+ - n_- = 2M \tag{22.9}$$

or

$$n_+ = J + M \qquad n_- = J - M \tag{22.10}$$

Hence, by (21.36), the eigenstates, normalized to unity, are

$$|JM\rangle = \frac{(a^\dagger_{1/2})^{J+M}(a^\dagger_{-1/2})^{J-M}}{\sqrt{(J+M)!(J-M)!}}|\mathbf{0}\rangle \tag{22.11}$$

The vacuum state, corresponding to zero boson occupation, represents zero angular momentum, or $|0, 0\rangle = |\mathbf{0}\rangle$.

In terms of the vector model of angular momentum, the representation (22.11) of the state $|JM\rangle$ may be recognized as the projection $M\hbar$ of the resultant of $2J$ spin one-half vectors combined to produce the ''stretched'' vector polygon with all spin one-half vectors ''parallel.'' The requirements of Bose-Einstein statistics for the spins that make up this resultant cause this state to be uniquely defined. At the level of the representations of the group $SU(2)$, discussed in Chapter 17, the connection between spin one-half bosons and the generators of the rotation group can be understood as an extension of the Clebsch-Gordan formula (17.71) for a direct product of n two-dimensional representations of $SU(2)$:

$$\mathbf{2}\otimes\mathbf{2}\otimes\mathbf{2}\otimes\ldots = (\mathbf{n}+\mathbf{1})\oplus(\mathbf{n}-\mathbf{1})^{\lambda_{n-1}}\oplus\ldots\oplus\left(\frac{\mathbf{3}-(-\mathbf{1})^n}{\mathbf{2}}\right)^{\lambda_1} \tag{22.12}$$

where the superscript λ_k ($k = 1, \ldots n-1$) denotes the multiplicity for each irreducible representation contained in the direct product. The $n+1$-dimensional representation $\mathbf{n+1}$ is uniquely contained in (22.12). It corresponds to the linear transformations among the $n+1$ totally symmetric stretched-configuration basis states for the system of n spin one-half bosons.

Exercise 22.5. Work out the decomposition of the direct product (22.12) for $n = 1$ to 6.

The violation of the connection between spin and statistics implied by the use of spin one-half bosons in this section does not vitiate the mathematical procedure that we have outlined. The "spin up" and "spin down" bosons defined here are not particles in the usual sense, since they have no momentum or energy. Rather, they are abstract carriers of spin, allowing an elegant description of angular momentum states. As auxiliary entities, these bosons may be used for a relatively easy evaluation of the Clebsch-Gordan coefficients and of the more complicated structures that arise in the coupling of more than two angular momenta.[2]

3. ***First-Order Perturbation Theory in Many-Body Systems.*** A simple and important illustration of the use of two-particle operators is afforded by a first-order perturbation calculation of the energy eigenvalues of a Hamiltonian which describes a system of interacting identical particles:

$$\mathcal{H} = \sum_i \varepsilon_i a_i^\dagger a_i + \frac{1}{2} \sum_{qrst} a_q^\dagger a_r^\dagger a_s a_t \langle qr | V | ts \rangle \tag{22.13}$$

It is assumed that the eigenstates of the unperturbed Hamiltonian of noninteracting particles

$$\mathcal{H}_0 = \sum_i \varepsilon_i a_i^\dagger a_i \tag{22.14}$$

are known and characterized as $|n_1, n_2, \ldots n_i \ldots\rangle$ by the eigenvalues n_i of the occupation number operators $a_i^\dagger a_i$. If the eigenvalues of $\mathcal{H}_0$ are nondegenerate, first-order perturbation theory gives for the energies the approximate values

$$E_{n_1 n_2 \ldots} = \sum_i n_i \varepsilon_i + \frac{1}{2} \sum_{qrst} \langle n_1 n_2 \ldots | a_q^\dagger a_r^\dagger a_s a_t | n_1 n_2 \ldots \rangle \langle qr | V | ts \rangle \tag{22.15}$$

In evaluating the matrix element of the operator $a_q^\dagger a_r^\dagger a_s a_t$, it is helpful to recognize that, owing to the orthogonality of the unperturbed eigenstates, nonvanishing contributions to the interaction energy are obtained only if $q \neq r$ and either $s = r$ and $t = q$ or $s = q$ and $t = r$, or if $q = r = s = t$. Equation (22.15) is therefore reducible to

$$\begin{aligned} E_{n_1 n_2 \ldots} = \sum_i n_i \varepsilon_i &+ \frac{1}{2} \sum_{q \neq r} n_q n_r [\langle qr | V | qr \rangle \pm \langle qr | V | rq \rangle] \\ &+ \frac{1}{2} \sum_q n_q (n_q - 1) \langle qq | V | qq \rangle \end{aligned} \tag{22.16}$$

The + sign holds for Bose-Einstein statistics and the − sign for Fermi-Dirac statistics. The two matrix elements $\langle qr | V | qr \rangle$ and $\langle qr | V | rq \rangle$, connecting the two one-particle states q and r, are said to have *direct* and *exchange* character, respectively. The last term in (22.16), which accounts for the interaction of particles occupying the same one-particle state, vanishes for fermions, since in that case $n_q = 0$ or 1 (Pauli exclusion principle).

The evaluation of a matrix element of the product of several creation and annihilation operators carried out here is typical of most calculations in many-body theories. The labor involved in such computations is significantly reduced if the operators in a product are arranged in *normal ordering*, i.e., with all annihilation operators standing to the right of all creation operators. The operators in the Ham-

[2]For a full treatment, see J. Schwinger, *On Angular Momentum* in Biedenharn and Van Dam (1965), p. 229.

iltonian (22.13) are already normally ordered. If a product is not yet normally ordered, it may, by repeated application of the commutation relations, be transformed into a sum of normally ordered products. A set of simple manipulative rules may be formulated[3] which permit the expansion of an operator of arbitrary complexity into terms with normal order.

As an example, we choose the fundamental problem of *atomic spectroscopy*, the determination of energy eigenvalues and eigenstates of an atom with n electron.[4] If all spin-dependent interactions are neglected, only electrostatic potentials are effective. In this approximation, both the total orbital and the total spin angular momentum commute with the Hamiltonian. As was suggested in Section 18.6, it is practical to require the eigenvectors of $\mathcal{H}_0$, on which the perturbation theory is based, to be also eigenvectors of the total orbital and the total spin angular momentum. A level with quantum numbers L and S is split by the spin-orbit interaction into a multiplet of eigenstates with definite J values ranging from $|L - S|$ to $L + S$. This scheme of building approximate energy eigenstates for an atom is known as L-S (or *Russell-Saunders*) *coupling*.

If $\mathcal{H}_0$ is a central-force Hamiltonian for noninteracting particles, the unperturbed eigenstates are characterized by the set of occupation numbers for the one-particle states, or *orbitals*, with radial and orbital quantum numbers n_i, ℓ_i. Each pair of quantum numbers n_i, ℓ_i defines an ith atomic (sub)shell. A set of occupation numbers for the atomic orbitals is said to define a *configuration*. A particular configuration usually contains many distinct states of the product form

$$\prod a^{\dagger}_{n_i\ell_i m_i m_s}|\mathbf{0}\rangle$$

Eigenstates of $\mathcal{H}_0$ that are represented by a product of n creation operators are called *independent particle states*.

Although, generally, knowledge of the atomic configuration and the quantum numbers L, S, M_L and M_S is not sufficient to specify the state of an atom unambiguously, in simple cases, such as near closed shells, these specifications may determine the state uniquely. The states of the two-electron atom (e.g., neutral helium) may be fully classified in this way, and we will discuss these in some detail.

If the two electrons are in different shells, the states of any two-electron configuration $(n_1\ell_1)(n_2\ell_2)$ which are simultaneously eigenstates of the total orbital and the total spin angular momentum are, according to Eq. (22.2),

$$|\Psi^{(2)}_{n_1\ell_1 n_2\ell_2}(LSM_LM_S)\rangle$$
$$= \sum_{m_1m_2} \langle \ell_1\ell_2 m_1 m_2 | \ell_1\ell_2 L M_L\rangle \sum_{m'_1m'_2} \langle \tfrac{1}{2}\,\tfrac{1}{2}\, m'_1 m'_2 \,|\tfrac{1}{2}\,\tfrac{1}{2}\, S M_S\rangle a^{\dagger}_{n_2\ell_2 m_2 m'_2} a^{\dagger}_{n_1\ell_1 m_1 m'_1}|\mathbf{0}\rangle \tag{22.17}$$

If the two electrons are in the same shell and the configuration is $(nl)^2$, it is legitimate to set $n_1 = n_2 = n$ and $\ell_1 = \ell_2 = \ell$ in (22.17), provided that a normalization factor of $1/\sqrt{2}$ is furnished.

Exercise 22.6. Use the symmetry relations for Clebsch-Gordan coefficients to show that a configuration $(n\ell)^2$ can only give rise to spin-orbit coupled two-electron states for which $L + S$ is even, i.e., states ^1S, ^3P, ^1D, and so on.

[3]Koltun and Eisenberg (1988), Chapter 8.

[4]A useful introduction to atomic, molecular, and solid state applications of quantum mechanics is Tinkham (1964).

The ground state of the neutral helium atom is described by the configuration $(1s)^2$ and has the spectroscopic character $^1S_0$. In our notation, this state may be expressed as

$$|\Psi^{(2)}_{1010}(0000)\rangle = a^\dagger_{100,-1/2}a^\dagger_{100,1/2}|\mathbf{0}\rangle \tag{22.18}$$

The configuration of the simplest excited states is $(1s)(n\ell)$ with $n > 1$. Since the two spins may couple to 0 or 1, the excited states ($L = \ell$) are classified as singlet ($S = 0$) and triplet ($S = 1$) states. With the appropriate values for the Clebsch-Gordan coefficients substituted in (22.17), we obtain for the triplet states:

$$\begin{aligned}
|\Psi^{(2)}_{10n\ell}(\ell 1m,\ 1)\rangle &= a^\dagger_{n\ell m,1/2}a^\dagger_{100,1/2}|\mathbf{0}\rangle \\
|\Psi^{(2)}_{10n\ell}(\ell 1m,\ 0)\rangle &= \frac{1}{\sqrt{2}}\,(a^\dagger_{n\ell m,-1/2}a^\dagger_{100,1/2} + a^\dagger_{n\ell m,1/2}a^\dagger_{100,-1/2})|\mathbf{0}\rangle \\
|\Psi^{(2)}_{10n\ell}(\ell 1m,\ -1)\rangle &= a^\dagger_{n\ell m,-1/2}a^\dagger_{100,-1/2}|\mathbf{0}\rangle
\end{aligned} \tag{22.19}$$

and for the singlet states:

$$|\Psi^{(2)}_{10n\ell}(\ell 0m0)\rangle = \frac{1}{\sqrt{2}}\,(a^\dagger_{n\ell m,-1/2}a^\dagger_{100,1/2} - a^\dagger_{n\ell m,1/2}a^\dagger_{100,-1/2})|\mathbf{0}\rangle \tag{22.20}$$

Owing to the anticommutation properties of the creation operators, the triplet states are symmetric under exchange of the spin quantum numbers of the two particles and antisymmetric under exchange of the set of spatial (radial and orbital) quantum numbers. The situation is reversed for the singlet states.

The perturbation interaction, arising from the Coulomb repulsion of the electrons, is diagonal with respect to all the unperturbed states that we have constructed, and the first-order corrections to the energy are the expectation values of the interaction in these states. These energies were already worked out in terms of direct and exchange integrals in Section 18.8. We now see that the identity of the electrons, manifested in their statistics, results in a definite correlation between the spatial, orbital, symmetry and the total spin S of the system. The states of parahelium are singlet states, and the states of orthohelium are triplet states. In complex atoms, the connection between S and the spatial symmetry of the state is less simple and not necessarily unique, but S remains instrumental in classifying the orbital symmetry of the states and thus serves as a quantum number on which the energy levels depend, even though spin-dependent interactions are neglected and the interaction depends only on the position coordinates of the electrons.[5]

4. ***The Hartree-Fock Method.*** One of the most useful methods for approximating the ground state of a system of n interacting *fermions* is based on the variational property of the Hamiltonian

$$\mathcal{H} = \sum_{\alpha\alpha'} b^\dagger_\alpha\langle\alpha|H_0|\alpha'\rangle b_{\alpha'} + \frac{1}{2}\sum_{\alpha\beta\alpha'\beta'} b^\dagger_\alpha b^\dagger_\beta\langle\alpha\beta|V|\alpha'\beta'\rangle b_{\beta'}b_{\alpha'} \tag{22.21}$$

The essence of the *Hartree-Fock method* is to seek a new one-particle basis with creation operators $a^\dagger_k$ such that the independent-particle state

$$|\Psi_\nu\rangle = a^\dagger_n a^\dagger_{n-1}\cdots a^\dagger_2 a^\dagger_1|\mathbf{0}\rangle \tag{22.22}$$

[5]For a compact treatment of the theory of atomic spectra in terms of the second quantization formalism, see Judd (1967).

renders the expectation value of $\mathcal{H}$ stationary. In this new basis, the Hamiltonian appears as

$$\mathcal{H} = \sum_{i,\ell} a_i^\dagger \langle i|H_0|\ell\rangle a_\ell + \frac{1}{2}\sum_{qrst} a_q^\dagger a_r^\dagger \langle qr|V|ts\rangle a_s a_t \tag{22.23}$$

The exact eigenstates of the Hamiltonian are usually not as simple as $|\Psi_\nu\rangle$ but can be thought of as linear combinations of independent-particle states, with the expression (22.22) as the leading term. Although the variational method per se does not single out the ground state from all energy eigenstates, the ground state is of paramount interest, and the knowledge that it *minimizes* the expectation value of $\mathcal{H}$ greatly aids its determination. We will use $|\Psi_\nu\rangle$ as written in the form (22.22) to denote the ground state. Excited states of the n-particle system will then be expressed in terms of $|\Psi_\nu\rangle$. For example, the state $a_j^\dagger a_k|\Psi_\nu\rangle$ (with k and j labeling occupied and unoccupied one-particle states, respectively) is an independent-particle state similar in structure to (22.22) but orthogonal to $|\Psi_\nu\rangle$ and may be regarded as an approximation to an excited state of the system.

The variation to be considered is a basis change, which is a unitary transformation and expressible as

$$a_k^\dagger + \delta a_k^\dagger = \sum_j a_j^\dagger(\delta_{jk} + i\varepsilon_{jk}) \tag{22.24}$$

or

$$\delta a_k^\dagger = i\sum_j a_j^\dagger \varepsilon_{jk}$$

with transformation coefficients ε_{jk} such that

$$|\varepsilon_{jk}| \ll 1$$

The general variation of the state $|\Psi_\nu\rangle$ can be built up as a linear combination of independent variations of the form

$$|\delta\Psi_{jk}\rangle = \varepsilon_{jk} a_j^\dagger a_k|\Psi_\nu\rangle \quad \text{or} \quad \langle\delta\Psi_{jk}| = \varepsilon_{jk}^*\langle\Psi_\nu|a_k^\dagger a_j \tag{22.25}$$

where, acting on a ket on the right, a_k must annihilate a fermion in one of the *occupied* one-particle states $1, \ldots, n$, and $a_j^\dagger$ must create a particle in one of the previously *unoccupied* one-particle states $n+1, \ldots$. The unitarity of the transformation coefficients in (22.24) requires that the ε_{jk} form a Hermitian matrix. Since the variation $|\delta\Psi_{kj}\rangle$, with subscripts reversed from (22.25), vanishes owing to the exclusion principle, the condition $\varepsilon_{kj} = \varepsilon_{jk}^*$ can be ignored, and the independence of the ε-variations is assured. (Variations with $j = k$ do not change the state and are therefore irrelevant.)

We may thus confine our attention to variations $|\delta\Psi\rangle$ of the form (22.25) which are orthogonal to the "best" state $|\Psi_\nu\rangle$ of the form (22.22). The variational theorem,

$$\delta\langle\mathcal{H}\rangle = 0$$

in conjunction with the Hermitian property of $\mathcal{H}$, requires that

$$\langle\Psi_\nu|\Psi_\nu\rangle\langle\delta\Psi|\mathcal{H}|\Psi_\nu\rangle - \langle\Psi_\nu|\mathcal{H}|\Psi_\nu\rangle\langle\delta\Psi|\Psi_\nu\rangle = 0$$

The orthogonality of $|\delta\Psi\rangle$ and $|\Psi_\nu\rangle$ guarantees that the variation preserves the normalization of the state, and according to the last equation, makes it necessary that

$|\delta\Psi\rangle$ also be orthogonal to $\mathcal{H}|\Psi_\nu\rangle$. Hence, the variational condition is (*Brillouin's* theorem)

$$\boxed{\langle\delta\Psi|\mathcal{H}|\Psi_\nu\rangle = 0} \tag{22.26}$$

If the Hamiltonian (22.23) and the variation (22.25) are substituted into this condition, we obtain

$$\sum_{i\ell}\langle i|H_0|\ell\rangle\langle\Psi_\nu|a_k^\dagger a_j a_i^\dagger a_\ell|\Psi_\nu\rangle + \frac{1}{2}\sum_{qrst}\langle qr|V|ts\rangle\langle\Psi_\nu|a_k^\dagger a_j a_q^\dagger a_r^\dagger a_s a_t|\Psi_\nu\rangle = 0$$

Since k labels an occupied one-particle state in $|\Psi_\nu\rangle$, and j labels an unoccupied one, the last relation is seen to be equivalent to the equation

$$\langle j|H_0|k\rangle + \sum_{t=1}^{n}[\langle jt|V|kt\rangle - \langle jt|V|tk\rangle] = 0 \tag{22.27}$$

The sum over t is to be taken only over the occupied one-particle states.

Condition (22.27) suggests the introduction of an *effective one-particle Hamiltonian*, H_{HF}, defined by its action on the (as yet undetermined) one-particle energy eigenstates $|m\rangle$:

$$H_{HF}|m\rangle \equiv H_0|m\rangle + \sum_{p=1}^{\infty}|p\rangle\sum_{t=1}^{n}[\langle pt|V|mt\rangle - \langle pt|V|tm\rangle] = \varepsilon_m|m\rangle \tag{22.28}$$

With this definition, condition (22.27) can be construed as expressing the orthogonality between the occupied and unoccupied eigenkets of H_{HF}:

$$\langle j|H_{HF}|k\rangle = 0 \tag{22.29}$$

If in the original one-particle basis $b_\alpha^\dagger$ the interaction between the fermions is diagonal and represented as

$$\langle\alpha\beta|V|\alpha'\beta'\rangle = V_{\alpha\beta}\delta_{\alpha\alpha'}\delta_{\beta\beta'} \tag{22.30}$$

Eq. (22.28) takes the form

$$H_{HF}|m\rangle \equiv H_0|m\rangle + \sum_{t=1}^{n}\sum_{\alpha\beta}|\alpha\rangle\langle t|\beta\rangle V_{\alpha\beta}[\langle\beta|t\rangle\langle\alpha|m\rangle - \langle\alpha|t\rangle\langle\beta|m\rangle] = \varepsilon_m|m\rangle \tag{22.31}$$

The summation over the Greek indices extends over the complete set of one-particle states. Equations (22.28) and (22.31) are known as the *Hartree-Fock equations*. From (22.28) we immediately infer that

$$\langle m|H_0|m\rangle + \sum_{t=1}^{n}[\langle mt|V|mt\rangle - \langle mt|V|tm\rangle] = \varepsilon_m \tag{22.32}$$

Exercise 22.7. Verify that the one-particle Hartree-Fock Hamiltonian H_{HF} is Hermitian.

The occupied states $|t\rangle$ in Eqs. (22.27) and (22.31) are not at our discretion. Since $|\Psi_\nu\rangle$ is to be a trial vector approximating the ground states, they must be chosen from among the eigenkets of (22.31) in a manner that will minimize the expectation value of the Hamiltonian. Frequently, the best choice corresponds to the use of those eigenkets that belong to the n lowest eigenvalues ε_k, although, perhaps

contrary to expectations, the variationally minimal value of $\langle \mathcal{H} \rangle$ is not just the sum of the Hartree-Fock one-particle energies, $\sum_{k=1}^{n} \varepsilon_k$. Rather, the Hartree-Fock approximation E_ν to the ground state energy is

$$E_\nu = \langle \mathcal{H} \rangle = \langle \Psi_\nu | \mathcal{H} | \Psi_\nu \rangle = \sum_{k=1}^{n} \langle k | H_0 | k \rangle + \frac{1}{2} \sum_{\ell,k=1}^{n} | \langle \ell k | V | \ell k \rangle - \langle \ell k | V | k \ell \rangle]$$
$$= \frac{1}{2} \sum_{k=1}^{n} [\varepsilon_k + \langle k | H_0 | k \rangle] = \sum_{k=1}^{n} \varepsilon_k - \frac{1}{2} \sum_{\ell,k=1}^{n} | \langle \ell k | V | \ell k \rangle - \langle \ell k | V | k \ell \rangle] \tag{22.33}$$

For the "excited" state $a_j^\dagger a_k | \Psi_\nu \rangle$, we obtain

$$\langle \mathcal{H} \rangle = \langle \Psi_\nu | a_k^\dagger a_j \mathcal{H} a_j^\dagger a_k | \Psi_\nu \rangle = E_\nu + \varepsilon_j - \varepsilon_k - \langle jk | V | jk \rangle + \langle jk | V | kj \rangle \tag{22.34}$$

If the last two terms can be neglected, $\varepsilon_j - \varepsilon_k$ represents an excitation energy of the system.

Exercise 22.8. Verify expression (22.34).

Exercise 22.9. Prove that the expectation value of $\mathcal{H}$ in the "ionized" state $a_k | \Psi_\nu \rangle$ with $n - 1$ particles is

$$\langle \mathcal{H} \rangle = E_\nu - \varepsilon_k \quad (\textit{Koopmans}' \text{ theorem}) \tag{22.35}$$

The practical task of solving the Hartree-Fock equations is far from straightforward. The equations have the appearance of a common eigenvalue problem, but the matrix elements of the interaction V, which enter the construction of the effective one-particle Hamiltonian H_{HF}, cannot be computed without foreknowledge of the appropriate n eigensolutions $|t\rangle$ of the coupled equations (22.31). These equations are nonlinear and require an iteration technique for their solution. One starts out by guessing a set of occupied one-particle states $|t\rangle$; using these, one calculates the matrix elements of V, and one then solves the Hartree-Fock equations (22.31). If the initial guess, based on insight and experience, was fortuitously good, n of the eigensolutions of (22.31) will be similar to the initially chosen kets. If, as is more likely, the eigensolutions of the Hartree-Fock equations fail to reproduce the starting kets, the eigensolutions corresponding to the lowest n eigenvalues ε_k are used to recalculate the matrix elements of V. This procedure is repeated until a self-consistent set of solutions is obtained. Sufficiently good initial guesses of the one-particle trial states are usually available, so that in actual practice fairly rapid convergence of the iteration process is the rule rather than the exception.

In the representation that diagonalizes V, the Hartree-Fock equations can be rewritten in matrix form as

$$\sum_{\beta} [\langle \alpha | H_0 | \beta \rangle \langle \beta | m \rangle + \sum_{t=1}^{n} \langle t | \beta \rangle V_{\alpha\beta} \langle \beta | t \rangle \langle \alpha | m \rangle - \langle t | \beta \rangle V_{\alpha\beta} \langle \alpha | t \rangle \langle \beta | m \rangle]$$
$$= \varepsilon_k \langle \alpha | m \rangle \tag{22.36}$$

As an application of these equations, we consider an atom with a nuclear charge Ze and with n electrons. Then

$$H_0 = \frac{\mathbf{p}^2}{2m} - \frac{Ze^2}{r}$$

The electron–electron interaction V is diagonal in the coordinate representation and has the form

$$V(\mathbf{r}\sigma, \mathbf{r}'\sigma') = \frac{e^2}{|\mathbf{r} - \mathbf{r}'|}$$

We choose the coordinate representation with spin as the basis $|\alpha\rangle$ and $|\beta\rangle$, and we denote the Hartree-Fock eigenfunctions as

$$\langle \mathbf{r}\sigma | m \rangle = \psi_m(\mathbf{r}\sigma)$$

With this notation, the Hartree-Fock equations (22.36) are transcribed as

$$\begin{aligned} -\frac{\hbar^2}{2m}\nabla^2\psi_m(\mathbf{r}\sigma) - \frac{Ze^2}{r}\psi_m(\mathbf{r}\sigma) + e^2\sum_{t=1}^{n}\sum_{\sigma'}\int \psi_t^*(\mathbf{r}'\sigma')\frac{1}{|\mathbf{r} - \mathbf{r}'|}\psi_t(\mathbf{r}'\sigma')\psi_m(\mathbf{r}\sigma)\,d^3r' \\ -e^2\sum_{t=1}^{n}\sum_{\sigma'}\int \psi_t^*(\mathbf{r}'\sigma')\frac{1}{|\mathbf{r} - \mathbf{r}'|}\psi_t(\mathbf{r}\sigma)\psi_m(\mathbf{r}'\sigma')\,d^3r' = \varepsilon_m\psi_m(\mathbf{r}\sigma) \end{aligned} \tag{22.37}$$

These coupled nonlinear differential-integral equations constitute the most familiar realization of the Hartree-Fock theory. The first sum on the left-hand side (without the term $t = m$ if m is an occupied state) represents the average effect of the interaction between all the other electrons in occupied one-particle states. The last sum on the left-hand side is attributable to the exchange matrix elements of the interaction.

Exercise 22.10. Show that the configuration space wave function corresponding to the independent particle state (22.22) can be expressed as the *Slater determinant*

$$\psi(\mathbf{r}_1\sigma_1, \ldots, \mathbf{r}_n\sigma_n) = \frac{1}{\sqrt{n!}}\begin{vmatrix} \psi_1(\mathbf{r}_1\sigma_1) & \psi_1(\mathbf{r}_2\sigma_2) & \cdots & \psi_1(\mathbf{r}_n\sigma_n) \\ \psi_2(\mathbf{r}_1\sigma_1) & \psi_2(\mathbf{r}_2\sigma_2) & \cdots & \psi_2(\mathbf{r}_n\sigma_n) \\ \vdots & \vdots & \cdots & \vdots \\ \psi_n(\mathbf{r}_1\sigma_1) & \psi_n(\mathbf{r}_2\sigma_2) & \cdots & \psi_n(\mathbf{r}_n\sigma_n) \end{vmatrix} \tag{22.38}$$

5. ***Quantum Statistics and Thermodynamics.*** The many-body operator formalism of Chapter 21 is ideally suited for treating statistical ensembles of identical particles. Here we will derive the quantum distribution functions for a system of noninteracting particles in thermal equilibrium.

If ρ denotes the *density* (or *statistical*) operator for an ensemble with fixed values for the averages of $\mathcal{H}$ and N, statistical thermodynamics requires that the von Neumann entropy,

$$S = -k\,\text{trace}(\rho \ln \rho) \tag{22.39}$$

be made a maximum subject to the constraints

$$\langle N \rangle = \text{trace}(\rho N) = n, \qquad \langle \mathcal{H} \rangle = \text{trace}(\rho\mathcal{H}) = E, \qquad \text{trace}(\rho) = 1 \tag{22.40}$$

The entropy principle is based on the probability and information concepts introduced in Section 2 in the Appendix and Section 15.5. Except for the multiplication by Boltzmann's constant k, the entropy S is that defined in Eq. (15.128).

Using the Lagrangian multipliers α and β, we see that the variational principle takes the form

$$\delta(S - k\alpha\langle N\rangle - k\beta\langle\mathcal{H}\rangle) = 0 \tag{22.41}$$

The normalization constraint in (22.40) requires that the variations of the density operator be restricted to

$$\text{trace}(\delta\rho) = 0 \tag{22.42}$$

and, therefore,

$$\delta S = -k \text{ trace}(\delta\rho \ln \rho + \delta\rho) = -k \text{ trace}(\delta\rho \ln \rho)$$

Substituting all the variations into (22.41), we obtain

$$\text{trace}[\delta\rho(\ln \rho + \alpha N + \beta\mathcal{H})] = 0$$

which is consistent with (22.42) only if

$$\ln \rho + \alpha N + \beta\mathcal{H} = -\ln Z\, 1$$

where Z is a number. We thus arrive at the grand canonical form of the density operator:

$$\boxed{\rho = \frac{e^{-\alpha N - \beta\mathcal{H}}}{Z}} \tag{22.43}$$

The normalization condition gives us

$$\boxed{Z = \text{trace } e^{-\alpha N - \beta\mathcal{H}}} \tag{22.44}$$

which is called the *grand partition function*. The parameters α and β must be determined from the first two constraint conditions (22.40). By thermodynamic arguments, $\beta = 1/kT$ is a measure of the temperature and $\mu = -\alpha/\beta$ is identified as the chemical potential.

Exercise 22.11. Evaluate the entropy for the equilibrium state (22.43), and show that

$$-kT \ln Z = \langle\mathcal{H}\rangle - \mu\langle N\rangle - TS = E - TS - \mu n \tag{22.45}$$

which is the grand canonical potential (or generalized free energy), suitable for relating thermodynamic variables to the underlying microscopic description.[6]

For a system of noninteracting identical particles with one-particle energies ε_i, known in thermodynamics as a generalized *ideal gas*,

$$\mathcal{H} = \sum_i \varepsilon_i a_i^\dagger a_i = \sum_i \varepsilon_i N_i \tag{22.46}$$

The ensemble average of any physical quantity represented by an operator Q may be computed by application of the formula

$$\langle Q\rangle = \text{trace } \rho Q \tag{22.47}$$

[6]Callen (1985), Section 5.3, and Reif (1965), Section 6.6.

We apply this relation to the evaluation of the average occupation numbers N_i:

$$\langle N_i \rangle = \langle a_i^\dagger a_i \rangle = \text{trace}(e^{-\alpha N - \beta \mathcal{H}} a_i^\dagger a_i)/Z \tag{22.48}$$

Using Eqs. (21.31)–(21.33) and the identity (3.59), we find that

$$\text{trace}(e^{-\alpha N - \beta \mathcal{H}} a_i^\dagger a_i) = e^{-(\alpha + \beta \varepsilon_i)} \text{trace}(e^{-\alpha N - \beta \mathcal{H}} a_i a_i^\dagger) \tag{22.49}$$

Exercise 22.12. Verify Eq. (22.49).

If the commutation relations for bosons or anticomutation relations for fermions are used, we obtain (with the upper sign for bosons and the lower sign for fermions)

$$\text{trace}(e^{-\alpha N - \beta \mathcal{H}} a_i^\dagger a_i) = e^{-(\alpha + \beta \varepsilon_i)} \text{trace}[e^{-\alpha N - \beta \mathcal{H}}(1 \pm a_i^\dagger a_i)]$$

Combining this relation with (22.43), we obtain

$$\boxed{\langle N_i \rangle = \langle a_i^\dagger a_i \rangle = \frac{1}{e^{\alpha + \beta \varepsilon_i} \mp 1}} \tag{22.50}$$

which is the familiar formula for the distribution of particles with Bose-Einstein (− sign) and Fermi-Dirac (+ sign) statistics, respectively.

The connection with the more conventional method for deriving the distribution (22.50) is established by introducing the occupation numbers n_i as the eigenvalues of $N_i = a_i^\dagger a_i$ and the corresponding eigenstates $|n_1, n_2, \ldots n_i, \ldots\rangle$ as basis states of the ideal gas. In this representation, the grand partition function becomes

$$Z = \sum_{n_1, n_2, \ldots} \prod_i e^{-(\alpha + \beta \varepsilon_i) n_i} \tag{22.51}$$

The distribution (22.50) is recovered by computing

$$\langle N_i \rangle = -\frac{1}{\beta} \frac{\partial \ln Z}{\partial \varepsilon_i} \tag{22.52}$$

which follows from (22.44) and (22.48). The two kinds of quantum statistics are distinguished and their partition functions are different, because in the Bose-Einstein case the occupation numbers assume all nonnegative integers as eigenvalues, whereas for the Fermi-Dirac case, $n_i = 0, 1$ are the only possible values.

The derivation of $\langle N_i \rangle$, using operators rather than the occupation-number basis, is intended to exhibit as plainly as possible the connection between the commutation relations for bosons and the anticommutation relations for fermions and the − and + signs, respectively, which characterize the denominator of the two distribution laws.

The Maxwell-Boltzmann distribution,

$$\langle N_i \rangle = e^{-\alpha - \varepsilon_i / kT} \tag{22.53}$$

is an approximation for the quantum distributions (22.50), valid if $\langle N_i \rangle \ll 1$. This may be regarded as a low-density or high-temperature approximation.

Exercise 22.13. Using operator algebra, show that the square of the fractional deviation from the mean occupation number is

$$\frac{\langle (N_i - \langle N_i \rangle)^2 \rangle}{\langle N_i \rangle^2} = \frac{\langle N_i^2 \rangle - \langle N_i \rangle^2}{\langle N_i \rangle^2} = e^{\alpha + \beta \varepsilon_i} = \frac{1}{\langle N_i \rangle} \pm 1 \tag{22.54}$$

where the + sign pertains to Bose-Einstein statistics and the − sign to Fermi-Dirac statistics. Also consider the Maxwell-Boltzmann limit.

In the next chapter, the formalism developed here will be applied to the derivation of the Planck distribution for photons in thermal equilibrium.

Problems

1. Consider a system of identical bosons with only two one-particle basis states, $a^\dagger_{1/2}\Psi^{(0)}$ and $a^\dagger_{-1/2}\Psi^{(0)}$. Define the Hermitian operators x, p_x, y, p_y by the relations

$$a_{1/2} = \frac{1}{\sqrt{2\hbar}}\left(cx + i\frac{p_x}{c}\right), \qquad a_{-1/2} = \frac{1}{\sqrt{2\hbar}}\left(cy + i\frac{p_y}{c}\right)$$

where c is an arbitrary real constant, and derive the commutation relations for these Hermitian operators. Express the angular momentum operator (22.6) in terms of these "coordinates" and "momenta," and also evaluate $\mathcal{J}^2$. Relate $\mathcal{J}^2$ to the square of the Hamiltonian of an isotropic two-dimensional harmonic oscillator by making the identification $c = \sqrt{m\omega}$, and show the connection between the eigenvalues of these operators.

2. (a) Using the fermion creation operators $a^\dagger_{jm}$, appropriate to particles with angular momentum j, form the closed-shell state in which all one-particle states $m = -j$ to $+j$ are occupied.

(b) Prove that the closed shell has zero total angular momentum.

(c) If a fermion with magnetic quantum number m is missing from a closed shell of particles with angular momentum j, show that, for coupling angular momenta, the hole state may be treated like a one-particle state with magnetic quantum number $-m$ and an effective creation operator $(-1)^{j-m}a_{jm}$.

3. Consider the unperturbed states $a^\dagger_{nm_n}\cdots a^\dagger_{km_k}\cdots a^\dagger_{1m_1}|\mathbf{0}\rangle$ of n spin one-half particles, each occupying one of n equivalent, degenerate orthogonal orbitals labeled by the quantum number k, and with $m_k = \pm 1/2$ denoting the spin quantum number associated with the orbital k. Show that in the space of the 2^n unperturbed states a spin-independent two-body interaction may, in first-order perturbation theory, be replaced by the effective exchange (or *Heisenberg*) Hamiltonian

$$\mathcal{H}_{\text{eff}} = -\frac{1}{\hbar^2}\sum_{k\ell}\langle k\ell|V|\ell k\rangle \mathbf{S}_k\cdot\mathbf{S}_\ell + \text{const.}$$

where $\mathbf{S}_k$ is the *localized* spin operator

$$\mathbf{S}_k = \frac{\hbar}{2}\sum_{m_k m_k'} a^\dagger_{km_k}a_{km_k'}\langle m_k|\boldsymbol{\sigma}|m_k'\rangle$$

4. For a Fermi gas of free particles with Fermi momentum p_F, calculate the ground state expectation value of the pair density operator

$$\sum_{\sigma',\sigma''}\psi^\dagger_{\sigma'}(\mathbf{r}')\psi^\dagger_{\sigma''}(\mathbf{r}'')\psi_{\sigma''}(\mathbf{r}'')\psi_{\sigma'}(\mathbf{r}')$$

in coordinate space and show that there is a repulsive interaction that would be absent if the particles were not identical. Show that there is no spatial correlation between particles of opposite spin.

5. Calculate in first order the energies of the ^1S, ^3P, and ^1D states arising from the atomic configuration p^2 (two electrons with $\ell = 1$ in the same shell). Use the multipole expansion

$$\frac{e^2}{|\mathbf{r}' - \mathbf{r}''|} = e^2 \sum_{k=0}^{\infty} \frac{4\pi}{2k+1} \gamma_k(r', r'') \sum_{q=-k}^{k} (-1)^q \, Y_k^q(\hat{\mathbf{r}}') Y_k^{-q}(\hat{\mathbf{r}}'')$$

for the interaction energy between the electrons, and show that the term energies may be expressed as

$$E(^1S) = E_0 + \langle \gamma_0 \rangle + \frac{10}{25} \langle \gamma_2 \rangle$$

$$E(^3P) = E_0 + \langle \gamma_0 \rangle - \frac{5}{25} \langle \gamma_2 \rangle$$

$$E(^1D) = E_0 + \langle \gamma_0 \rangle + \frac{1}{25} \langle \gamma_2 \rangle$$

where $\langle \gamma_k \rangle$ is the radial integral

$$\langle \gamma_k \rangle = e^2 \iint \gamma_k(r', r'')[R(r')]^2[R(r'')]^2 r'^2 r''^2 \, dr' \, dr''$$

6. Apply the Hartree-Fock method to a system of two ''electrons'' which are attracted to the coordinate origin by an isotropic harmonic oscillator potential $m\omega^2 r^2/2$ and which interact with each other through a potential $V = C(\mathbf{r}' - \mathbf{r}'')^2$. Solve the Hartree-Fock equations for the ground state and compare with the exact result and with first-order perturbation theory.

CHAPTER 23

Photons and the Electromagnetic Field

The formalism developed in Chapter 21 for the quantum description of identical particles is sufficiently general to be applicable also to *photons*, the particles associated with the electromagnetic field, even though photons are relativistic particles. Instead of deriving the quantum theory of the electromagnetic field by applying quantization postulates to the classical theory, we build the elements of the quantum theory of radiation on the foundations of ordinary quantum mechanics for particles and show that in the correspondence limit the classical Maxwell theory is recovered. Our goal is to acquire enough background for a first-order account of the interaction between photons and charged particles (mainly electrons in atoms) and the principles of quantum optics. With these tools it is then possible to deepen our understanding of optical coherence, correlation, and interference phenomena.

1. ***Fundamental Notions.*** The theory is based on empirical facts about photons: they are bosons with zero mass; in free space they can carry linear momentum $\hbar\mathbf{k}$ and energy $\hbar\omega_k = c\hbar k$; and simultaneously with sharp linear momentum, they can have a definite value $+\hbar$ or $-\hbar$ for the component $\mathbf{J}\cdot\hat{\mathbf{k}}$ of angular momentum along the direction of propagation.

Photons with a sharp value for $\mathbf{J}\cdot\hat{\mathbf{k}}$ are said to possess definite positive or negative *helicity*. Since the component of orbital angular momentum parallel to $\mathbf{k}$ is necessarily zero, *spin* 1 may be attributed to the photon, provided that the eigenvalue zero of the projection $\mathbf{J}\cdot\hat{\mathbf{k}}$ is excluded. The consistent omission from the theory of the eigenstate for which $\mathbf{J}\cdot\hat{\mathbf{k}}$ has the value zero is possible only because the photon has vanishing mass and cannot be brought to rest in any Lorentz frame of reference for the purpose of measuring its intrinsic spin separately from the orbital angular momentum that it carries.

By definition, the operators $a_+^\dagger(\mathbf{k})$ and $a_-^\dagger(\mathbf{k})$ correspond to the creation of a photon with momentum $\hbar\mathbf{k}$ and with helicity $\mathbf{J}\cdot\hat{\mathbf{k}} = +\hbar$ and $-\hbar$, respectively. The momentum states, or photon *modes*, are subjected to periodic boundary conditions defining a discrete set of $\mathbf{k}$ vectors with components $2\pi n_x/L$, $2\pi n_y/L$, $2\pi n_z/L$, where n_x, n_y, n_z are integers $(0, \pm1, \pm2, \ldots)$ as in (4.53). The transition to the continuum may be made by letting the arbitrary length $L \to \infty$, or formally by simply setting $L = 2\pi$ and replacing all sums over $\mathbf{k}$ by integrals and Kronecker deltas by delta functions. All spatial integrations extend over the cube of volume L^3.

The commutation relations for the photon creation and annihilation operators are

$$[a_+(\mathbf{k}), a_+^\dagger(\mathbf{k}')] = [a_-(\mathbf{k}), a_-^\dagger(\mathbf{k}')] = \delta_{\mathbf{k}\mathbf{k}'} \tag{23.1}$$

All other commutators vanish.

The important physical operators for a system of free photons are the total energy,

$$\mathcal{H} = \sum_{\mathbf{k}} c\hbar k[a_+^\dagger(\mathbf{k})a_+(\mathbf{k}) + a_-^\dagger(\mathbf{k})a_-(\mathbf{k})] \tag{23.2}$$

the total linear momentum,

$$\mathbf{P} = \sum_{\mathbf{k}} \hbar\mathbf{k}[a_+^\dagger(\mathbf{k})a_+(\mathbf{k}) + a_-^\dagger(\mathbf{k})a_-(\mathbf{k})] \tag{23.3}$$

and the total number of photons,

$$N = \sum_{\mathbf{k}} [a_+^\dagger(\mathbf{k})a_+(\mathbf{k}) + a_-^\dagger(\mathbf{k})a_-(\mathbf{k})] \tag{23.4}$$

The operator $\mathcal{J} \cdot \hat{\mathbf{k}}$, representing the total angular momentum component along the direction $\mathbf{k}$, is an additive one-particle operator and thus a bilinear combination of $a_+^\dagger(\mathbf{k}')$, $a_-^\dagger(\mathbf{k}')$, $a_+(\mathbf{k}'')$, $a_-(\mathbf{k}'')$. Owing to the helicity property of photons, the operators creating and annihilating photons of momentum $\hbar\mathbf{k}$ occur in this linear combination only in the diagonal terms $a_+^\dagger(\mathbf{k})a_+(\mathbf{k})$ and $a_-^\dagger(\mathbf{k})a_-(\mathbf{k})$. These terms, which represent the contribution to $\mathcal{J} \cdot \hat{\mathbf{k}}$ of photons with linear momentum $\mathbf{k}$, may be written as

$$[\mathcal{J} \cdot \hat{\mathbf{k}}, a_+^\dagger(\mathbf{k})]a_+(\mathbf{k}) + [\mathcal{J} \cdot \hat{\mathbf{k}}, a_-^\dagger(\mathbf{k})]a_-(\mathbf{k}) = \hbar[a_+^\dagger(\mathbf{k})a_+(\mathbf{k}) - a_-^\dagger(\mathbf{k})a_-(\mathbf{k})] \tag{23.5}$$

This operator generates rotation of the states with definite $\mathbf{k}$ about the direction $\mathbf{k}$. According to Chapter 17 and Section 21.2, we expect that such a rotation changes $a_+^\dagger(\mathbf{k})$ into $a_+^\dagger(\mathbf{k})e^{-i\alpha}$ and $a_-^\dagger(\mathbf{k})$ into $a_-^\dagger(\mathbf{k})e^{i\alpha}$, if α is the angle of rotation.

Exercise 23.1. Using the commutation relations, prove that

$$U_R a_\pm^\dagger U_R^\dagger = e^{-i\alpha[a_+^\dagger(\mathbf{k})a_+(\mathbf{k}) - a_-^\dagger(\mathbf{k})a_-(\mathbf{k})]} a_\pm^\dagger(\mathbf{k}) e^{i\alpha[a_+^\dagger(\mathbf{k})a_+(\mathbf{k}) - a_-^\dagger(\mathbf{k})a_-(\mathbf{k})]} = e^{\mp i\alpha} a_\pm^\dagger(\mathbf{k})$$

It is apparent that under the simple rotation considered, $a_+^\dagger(\mathbf{k})$ and $a_-^\dagger(\mathbf{k})$ transform like the T_1^1 and T_1^{-1} components of an irreducible tensor of rank one defined in Section 17.7. Using Eqs. (17.85), we introduce the new operators

$$\begin{aligned} a_1^\dagger(\mathbf{k}) &= \frac{1}{\sqrt{2}}[-a_+^\dagger(\mathbf{k}) + a_-^\dagger(\mathbf{k})] \\ a_2^\dagger(\mathbf{k}) &= \frac{i}{\sqrt{2}}[a_+^\dagger(\mathbf{k}) + a_-^\dagger(\mathbf{k})] \end{aligned} \tag{23.6}$$

Under a rotation about $\mathbf{k}$ by an angle α, these operators are transformed into

$$a_1^\dagger(\mathbf{k}) \cos\alpha + a_2^\dagger(\mathbf{k}) \sin\alpha \quad \text{and} \quad -a_1^\dagger(\mathbf{k}) \sin\alpha + a_2^\dagger(\mathbf{k}) \cos\alpha$$

just as two orthogonal components of a vector in a plane perpendicular to $\mathbf{k}$. If we introduce two perpendicular unit vectors $\hat{\mathbf{e}}_\mathbf{k}^{(1)}$ and $\hat{\mathbf{e}}_\mathbf{k}^{(2)}$ in the plane perpendicular to $\mathbf{k}$, such that

$$\hat{\mathbf{e}}_\mathbf{k}^{(1)} \times \hat{\mathbf{e}}_\mathbf{k}^{(2)} = \hat{\mathbf{k}}, \qquad \hat{\mathbf{e}}_\mathbf{k}^{(2)} \times \hat{\mathbf{k}} = \hat{\mathbf{e}}_\mathbf{k}^{(1)}, \qquad \hat{\mathbf{k}} \times \hat{\mathbf{e}}_\mathbf{k}^{(1)} = \hat{\mathbf{e}}_\mathbf{k}^{(2)} \tag{23.7}$$

it follows that

$$a_1^\dagger(\mathbf{k})\hat{\mathbf{e}}_\mathbf{k}^{(1)} + a_2^\dagger(\mathbf{k})\hat{\mathbf{e}}_\mathbf{k}^{(2)} = \frac{1}{\sqrt{2}} a_+^\dagger(\mathbf{k})(-\hat{\mathbf{e}}_\mathbf{k}^{(1)} + i\hat{\mathbf{e}}_\mathbf{k}^{(2)}) + \frac{1}{\sqrt{2}} a_-^\dagger(\mathbf{k})(\hat{\mathbf{e}}_\mathbf{k}^{(1)} + i\hat{\mathbf{e}}_\mathbf{k}^{(2)})$$

transforms like a three-vector that is perpendicular to $\mathbf{k}$.

The corresponding Hermitian conjugate operator is

$$a_1(\mathbf{k})\hat{\mathbf{e}}_\mathbf{k}^{(1)} + a_2(\mathbf{k})\hat{\mathbf{e}}_\mathbf{k}^{(2)} = -\frac{1}{\sqrt{2}} a_+(\mathbf{k})(\hat{\mathbf{e}}_\mathbf{k}^{(1)} + i\hat{\mathbf{e}}_\mathbf{k}^{(2)}) + \frac{1}{\sqrt{2}} a_-(\mathbf{k})(\hat{\mathbf{e}}_\mathbf{k}^{(1)} - i\hat{\mathbf{e}}_\mathbf{k}^{(2)}) \tag{23.8}$$

The right-hand side can be expressed more simply is we introduce the complex unit vectors,

$$\hat{\mathbf{e}}_{\mathbf{k}}^{(+)} = -\frac{1}{\sqrt{2}}(\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} + i\hat{\mathbf{e}}_{\mathbf{k}}^{(2)})$$
$$\hat{\mathbf{e}}_{\mathbf{k}}^{(-)} = \frac{1}{\sqrt{2}}(\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} - i\hat{\mathbf{e}}_{\mathbf{k}}^{(2)}) \tag{23.9}$$

Equation (23.8) then takes the form

$$a_1(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} + a_2(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(2)} = a_+(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(+)} + a_-(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(-)} \tag{23.10}$$

This relation constitutes a unitary transformation from a helicity representation to a linear polarization representation.

Exercise 23.2. From the definition (23.9), calculate the Hermitian scalar and vector products of the unit vectors $\hat{\mathbf{e}}_{\mathbf{k}}^{(+)}$, $\hat{\mathbf{e}}_{\mathbf{k}}^{(-)}$, and $\hat{\mathbf{k}}$.

Exercise 23.3. Show that

$$[a_1(\mathbf{k}), a_1^\dagger(\mathbf{k})] = [a_2(\mathbf{k}), a_2^\dagger(\mathbf{k})] = 1 \tag{23.11}$$

It is tempting now to consider a unitary transformation from the *momentum* representation, which we have used to introduce photons, to a *coordinate* representation by the application of the one-particle transformation coefficients

$$\langle \mathbf{r}|\mathbf{k}\rangle = L^{-3/2}\, e^{i\mathbf{k}\cdot\mathbf{r}} \tag{23.12}$$

In this way, it is possible, as in Section 21.5, to define vector operators that describe the creation (and annihilation) of a photon at a position $\mathbf{r}$, such as

$$\boldsymbol{\psi}^\dagger(\mathbf{r}) = \frac{1}{L^{3/2}}\sum_{\mathbf{k}}[a_1^\dagger(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} + a_2^\dagger(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(2)}]e^{-i\mathbf{k}\cdot\mathbf{r}} \tag{23.13}$$

In conformity with Eq. (21.56), one could then also introduce a photon wave function in configuration space. However, such a representation—straightforward for the case of nonrelativistic particles—is of limited usefulness for photons. This is most easily seen by examining an observable like the energy of the system which, after transformation to coordinate space, will appear in the form

$$\mathcal{H} = \frac{\hbar c}{L^3}\sum_{\mathbf{k}}\int\int \boldsymbol{\psi}^\dagger(\mathbf{r}')\cdot\boldsymbol{\psi}(\mathbf{r})\, k\, e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\, d^3r\, d^3r'$$

Since $k = \sqrt{k_x^2 + k_y^2 + k_z^2}$, the integrand cannot be transformed into a local expression involving a finite number of field derivatives and interpretable as an energy *density*.

Nor is it possible to construct a simple *local current density* operator from the fields $\boldsymbol{\psi}(\mathbf{r})$ and $\boldsymbol{\psi}^\dagger(\mathbf{r})$, although $\int \boldsymbol{\psi}^\dagger(\mathbf{r})\boldsymbol{\psi}(\mathbf{r})\, d^3r$ is a constant of the motion. Last, but not least, the field $\boldsymbol{\psi}(\mathbf{r})$ defined by (23.13) does not have simple transformation properties under Lorentz transformations. We will see in Chapter 24 that a transformation analogous to (23.13) can be carried out without any of these stumbling blocks for particles of *half-integral* spin. It succeeds for electrons and neutrinos (fermions) but not for bosons!

For photons, these deficiencies are remedied by introducing a slightly modified set of *field operators*:

$$\mathbf{A}^{(+)}(\mathbf{r}) = \sqrt{4\pi\hbar c^2}\,\frac{1}{L^{3/2}}\sum_{\mathbf{k}}\frac{1}{\sqrt{2\omega_k}}[a_1(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} + a_2(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(2)}]e^{i\mathbf{k}\cdot\mathbf{r}} \tag{23.14}$$

The significant difference between this vector operator and $\psi(\mathbf{r})$ is the appearance of the coefficient $\sqrt{2\omega_k}$. We will see that its insertion permits us to write the energy density and other important observables as simple *local* bilinear functions of the field operators. The constants in the definition (23.14) were chosen with the foreknowledge of the ultimate identification of this operator with the *vector potential* of the electromagnetic field. Equation (23.14) has a Hermitian adjoint companion:

$$\begin{aligned}\mathbf{A}^{(-)}(\mathbf{r}) &= [\mathbf{A}^{(+)}(\mathbf{r})]^\dagger \\ &= \sqrt{4\pi\hbar c^2}\,\frac{1}{L^{3/2}}\sum_{\mathbf{k}}\frac{1}{\sqrt{2\omega_k}}[a_1^\dagger(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} + a_2^\dagger(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(2)}]e^{-i\mathbf{k}\cdot\mathbf{r}}\end{aligned} \tag{23.15}$$

All field operators are periodic functions of the coordinates x, y, z with period L.

If we work in the Heisenberg picture, the operators $a(\mathbf{k})$ have a simple time dependence, which can be inferred from the equation of motion for the free photon system:

$$i\hbar\frac{da(\mathbf{k}, t)}{dt} = [a(\mathbf{k}, t), \mathcal{H}]$$

When the energy operator $\mathcal{H}$ defined in (23.2) is substituted in this equation, we obtain

$$i\frac{da(\mathbf{k}, t)}{dt} = \omega_k a(\mathbf{k}, t)$$

which has the solution

$$a(\mathbf{k}, t) = a(\mathbf{k})e^{-i\omega_k t} \qquad a^\dagger(\mathbf{k}, t) = a^\dagger(\mathbf{k})e^{i\omega_k t} \tag{23.16}$$

It follows that $\mathbf{A}^{(+)}(\mathbf{r}, t)$ is obtained from (23.14) if we replace $e^{i\mathbf{k}\cdot\mathbf{r}}$ by $e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}$. Similarly, to obtain $\mathbf{A}^{(-)}(\mathbf{r}, t)$, we replace $e^{-i\mathbf{k}\cdot\mathbf{r}}$ in (23.15) by $e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}$. These forms explain why $\mathbf{A}^{(+)}$ is called the *positive frequency part* and $\mathbf{A}^{(-)}$ the *negative frequency part* of the Hermitian operator $\mathbf{A}(\mathbf{r}, t)$, which is defined as the sum of the two:

$$\begin{aligned}\mathbf{A}(\mathbf{r}, t) = \sqrt{4\pi\hbar c^2}\,\frac{1}{L^{3/2}}\sum_{\mathbf{k}}\frac{1}{\sqrt{2\omega_k}}\{&[a_1(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} + a_2(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(2)}]e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} \\ &+ [a_1^\dagger(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} + a_2^\dagger(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(2)}]e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}\}\end{aligned} \tag{23.17}$$

Owing to the orthogonality of the polarization vectors $\hat{\mathbf{e}}_{\mathbf{k}}$ to $\mathbf{k}$, this operator satisfies the *transversality* or *Coulomb gauge* condition

$$\nabla\cdot\mathbf{A}(\mathbf{r}, t) = 0 \tag{23.18}$$

In Cartesian coordinates, it also obeys the equation

$$\left(\nabla^2 - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\mathbf{A}(\mathbf{r}, t) = 0 \tag{23.19}$$

because $\omega_k = ck$. The Hermitian operators for the electric field $\mathbf{E}$ and the magnetic field $\mathbf{B}$, defined by

$$\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \quad \text{and} \quad \mathbf{B} = \nabla \times \mathbf{A} \tag{23.20}$$

respectively, are local field operators that satisfy the Maxwell equations for the free radiation field in the *Coulomb gauge*, excluding the electrostatic contribution $-\nabla\phi$ to the electric field which arises from the Coulomb interaction between charges.

The electric and magnetic fields are usually also partitioned into positive and negative frequency parts:

$$\mathbf{E} = \mathbf{E}^{(+)} + \mathbf{E}^{(-)} \quad \text{and} \quad \mathbf{B} = \mathbf{B}^{(+)} + \mathbf{B}^{(-)} \tag{23.21}$$

From (23.20) and (23.17) we deduce, explicitly,

$$\begin{aligned}
\mathbf{E}^{(+)}(\mathbf{r}, t) &= i\sqrt{2\pi\hbar}\,\frac{1}{L^{3/2}}\sum_{\mathbf{k}}\sqrt{\omega_k}\,[a_1(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} + a_2(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(2)}]e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} \\
\mathbf{E}^{(-)}(\mathbf{r}, t) &= -i\sqrt{2\pi\hbar}\,\frac{1}{L^{3/2}}\sum_{\mathbf{k}}\sqrt{\omega_k}\,[a_1^\dagger(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} + a_2^\dagger(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(2)}]e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}
\end{aligned} \tag{23.22}$$

$$\begin{aligned}
\mathbf{B}^{(+)}(\mathbf{r}, t) &= i\sqrt{2\pi\hbar c^2}\,\frac{1}{L^{3/2}}\sum_{\mathbf{k}}\frac{1}{\sqrt{\omega_k}}\,[a_1(\mathbf{k})\mathbf{k}\times\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} + a_2(\mathbf{k})\mathbf{k}\times\hat{\mathbf{e}}_{\mathbf{k}}^{(2)}]e^{i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)} \\
\mathbf{B}^{(-)}(\mathbf{r}, t) &= -i\sqrt{2\pi\hbar c^2}\,\frac{1}{L^{3/2}}\sum_{\mathbf{k}}\frac{1}{\sqrt{\omega_k}}\,[a_1^\dagger(\mathbf{k})\mathbf{k}\times\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} + a_2^\dagger(\mathbf{k})\mathbf{k}\times\hat{\mathbf{e}}_{\mathbf{k}}^{(2)}]e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}
\end{aligned} \tag{23.23}$$

We also need the Fourier transforms of these vector functions, such as

$$a_1(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} + a_2(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(2)} = -\frac{i}{\sqrt{2\pi\hbar}}\frac{1}{\sqrt{\omega_k}}\frac{1}{L^{3/2}}\int e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}\mathbf{E}^{(+)}(\mathbf{r}, t)\,d^3r \tag{23.24}$$

and

$$a_1(\mathbf{k})\mathbf{k}\times\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} + a_2(\mathbf{k})\mathbf{k}\times\hat{\mathbf{e}}_{\mathbf{k}}^{(2)} = -\frac{i}{\sqrt{2\pi\hbar c^2}}\sqrt{\omega_k}\,\frac{1}{L^{3/2}}\int e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega_k t)}\mathbf{B}^{(+)}(\mathbf{r}, t)\,d^3r \tag{23.25}$$

The identification of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{E}$ as the vector potential in the Coulomb gauge, the magnetic field, and the electric field, respectively, will be complete only if we can verify that the expressions for physical observables, such as the energy and momentum of the field, have the correct classical form in the correspondence limit. We outline the proofs in the next section.

2. ***Energy, Momentum, and Angular Momentum of the Radiation Field.*** We may now transform the total energy from its energy-momentum form, (23.2), to the standard spacetime expression by the use of Eqs. (23.8) and (23.24):

$$\begin{aligned}
\mathcal{H} &= \sum_{\mathbf{k}}\hbar\omega_k[a_+^\dagger(\mathbf{k})a_+(\mathbf{k}) + a_-^\dagger(\mathbf{k})a_-(\mathbf{k})] \\
&= \sum_{\mathbf{k}}\hbar\omega_k[a_1^\dagger(\mathbf{k})a_1(\mathbf{k}) + a_2^\dagger(\mathbf{k})a_2(\mathbf{k})] \\
&= \sum_{\mathbf{k}}\hbar\omega_k[a_1^\dagger(\mathbf{k})\hat{\mathbf{e}}_k^{(1)} + a_2^\dagger(\mathbf{k})\hat{\mathbf{e}}_k^{(2)}]\cdot[a_1(\mathbf{k})\hat{\mathbf{e}}_k^{(1)} + a_2(\mathbf{k})\hat{\mathbf{e}}_k^{(2)}] \\
&= \frac{1}{2\pi L^3}\sum_{\mathbf{k}}\iint e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}')}\mathbf{E}^{(-)}(\mathbf{r}, t)\cdot\mathbf{E}^{(+)}(\mathbf{r}', t)\,d^3r\,d^3r'
\end{aligned}$$

and finally,

$$\boxed{\mathcal{H} = \frac{1}{2\pi}\int \mathbf{E}^{(-)}(\mathbf{r}, t)\cdot\mathbf{E}^{(+)}(\mathbf{r}, t)\, d^3r} \tag{23.26}$$

This expression for the energy operator is local in the sense that an energy density can be assigned to every point in space, and the integrand sum of all these contributions makes up the total field energy.

Exercise 23.4. By expressing the field operators as sums of positive and negative frequency parts, show that

$$\boxed{\frac{1}{8\pi}\int [\mathbf{E}^2(\mathbf{r}, t) + \mathbf{B}^2(\mathbf{r}, t)]\, d^3r = \mathcal{H} + E_0} \tag{23.27}$$

where E_0 is the constant (infinite) *zero-point energy*

$$\boxed{E_0 = \frac{1}{2}\sum_{\mathbf{k}} \hbar\omega_k} \tag{23.28}$$

Noting that the left-hand side of (23.27) is formally the correct relativistic definition of the energy of the classical electromagnetic field, show that the photon annihilation and creation operators are, in the classical limit, proportional to the Fourier amplitudes of the classical radiation field. Determine the constant of proportionality.[1]

The additive constant zero-point energy E_0 in (23.27) may seem irrelevant, since it merely shifts the origin of the energy scale. However, the frequency spectrum ω_k of the photon modes, and thus the zero-point energy, depends on the boundary conditions imposed on the fields. If, instead of being periodic in a very large cube, with the ultimate limit $L \to \infty$, the fields are confined to a cavity or the space between two conducting plates, the frequency spectrum is altered and depends on the dimensions of the confining boundaries. Under these circumstances, the full expression (23.27) for the energy is needed to calculate the *Casimir* (or *van der Waals*) forces on the boundary surfaces, which are present even in the vacuum state. If L is the linear dimension involved, e.g., the distance between the conducting plates, the force must in magnitude be of the order $\hbar c/L^2$. Unlike the ''uncertainty pressure'' caused by a particle in a box, the Casimir effect represents an attractive interaction.[2]

Exercise 23.5. Show that the total linear momentum (23.3) may be transformed into

$$\boxed{\mathbf{P} = \frac{1}{2\pi c}\int \mathbf{E}^{(-)}(\mathbf{r}, t)\times\mathbf{B}^{(+)}(\mathbf{r}, t)\, d^3r} \tag{23.29}$$

and establish the correspondence with the classical expression for the momentum of the field.

[1]Jackson (1975), Section 12.10, and Sakurai (1967), Chapter 2. For more detail on the quantum theory of radiation, see Loudon (1983) and Milonni (1995).

[2]Itzykson and Zuber (1980), Section 3-2-4.

Equations (23.27) and (23.29) may be cast into a compact relativistic formula for the energy-momentum four-vector

$$P^\mu = \left(\frac{E}{c}, \mathbf{P}\right) \tag{23.30}$$

if the contravariant spacetime coordinate four-vector

$$x^\mu = (x^0, x^1, x^2, x^3) = (ct, x, y, z) = (ct, \mathbf{r})$$

and the metric

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 = c^2dt^2 - dx^2 - dy^2 - dz^2$$

are introduced. We thus choose the metric tensor

$$g^{11} = g^{22} = g^{33} = -g^{00} = -1 \quad \text{and} \quad g^{\mu\nu} = 0 \quad \text{if} \quad \mu \neq \nu \tag{23.31}$$

We note that the covariant four-vector x_μ is

$$x_\mu = g_{\mu\nu}x^\nu = (ct, -x, -y, -z) \tag{23.32}$$

and that the four-gradient is given by

$$\partial^\mu \equiv \frac{\partial}{\partial x_\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\nabla\right) \qquad \partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, \nabla\right) \tag{23.33}$$

The notation and the metric used here for relativistic quantum mechanics is generally the same as is customary in particle physics and quantum field theory, but we retain $\hbar$ and c explicitly and employ (unrationalized) Gaussian electromagnetic units rather than Heaviside-Lorentz units, as can be seen by the form of the field energy, (23.27). (For more discussion of units, see Appendix Section 4.)

Using the relation $\nabla \cdot \mathbf{E} = 0$, we may transform the integrals (23.26) and (23.29) into

$$\boxed{P^\mu = -\frac{1}{2\pi c}\int \mathbf{E}^{(-)}(\mathbf{r}, t) \cdot \partial^\mu \mathbf{A}^{(+)}(\mathbf{r}, t)\, d^3r} \tag{23.34}$$

This formula has the usual structure of an additive one-particle operator like (21.62), except that the differential operator stands between $\mathbf{E}^{(-)}$ and $\mathbf{A}^{(+)}$ rather than between $\psi^\dagger$ and ψ.

The classical expression for the *angular momentum* of the electromagnetic field is[3]

$$\mathcal{J} = \frac{1}{4\pi c}\int \mathbf{r} \times (\mathbf{E} \times \mathbf{B})\, d^3r \tag{23.35}$$

This may be shown to be the classical limit of the operator equation

$$\mathcal{J} = \frac{1}{2\pi c}\int \sum_{i=0}^{3} E_i^{(-)}(\mathbf{r} \times \nabla)A_i^{(+)}\, d^3r + \frac{1}{2\pi c}\int \mathbf{E}^{(-)} \times \mathbf{A}^{(+)}\, d^3r \tag{23.36}$$

Comparison with expression (23.34) for the momentum shows that the first term on the right-hand side of (23.36) has the form of an *orbital angular momentum* operator. The second term accounts for the *spin* of the photons. If the expansions

[3]Jackson (1975), p. 333.

(23.14) for $\mathbf{A}^{(+)}$ and (23.22) for $\mathbf{E}^{(-)}$ are substituted, the second term is seen to contribute precisely the value given by Eq. (23.5) to the component of total angular momentum along $\hat{\mathbf{k}}$ for each photon momentum state $\mathbf{k}$. The orbital angular momentum has, as usual, zero projection along the direction of the propagation.

The photon spin operator may also be written as

$$\mathcal{S} = \frac{i}{2\pi\hbar c}\int (E_x^{(-)}E_y^{(-)}E_z^{(-)})\,\mathsf{S}\begin{pmatrix}A_x^{(+)}\\A_y^{(+)}\\A_z^{(+)}\end{pmatrix} d^3r \tag{23.37}$$

The 3×3 matrices S satisfy the angular momentum commutation relations and have eigenvalues 0 and $\pm\hbar$, thus confirming that photons are particles with spin one.

Exercise 23.6. Prove the connection between (23.35) and (23.36) by using the identity

$$[\mathbf{r}\times(\mathbf{E}\times\mathbf{B})]_x = \sum_{i=1}^{3} E_i(\mathbf{r}\times\nabla)_x A_i + (\mathbf{E}\times\mathbf{A})_x + \nabla\cdot[\mathbf{E}(zA_y - yA_z)] \tag{23.38}$$

Exercise 23.7. Construct the matrices S in Eq. (23.37) explicitly and verify their commutation relations.

From the behavior of the various physical quantities associated with the field, we infer that the annihilation and creation operators $a_1(\mathbf{k})$ and $a_1^\dagger(\mathbf{k})$ become in the classical limit proportional to the Fourier amplitudes of the electric radiation field linearly polarized in one direction, while the operators $a_2(\mathbf{k})$ and $a_2^\dagger(\mathbf{k})$ become proportional to the Fourier amplitudes with perpendicular linear polarization. According to the relation (23.8), the positive-helicity operators $a_+(\mathbf{k})$ and $a_+^\dagger(\mathbf{k})$ correspond in the classical limit to the amplitudes of an electric field with left-circular polarization. Similarly the negative-helicity operators $a_-(\mathbf{k})$ and $a_-^\dagger(\mathbf{k})$ correspond to right-circular polarization in optics.[4]

Exercise 23.8. Show that if viewed by an observer toward whom the light wave is propagating, the real and imaginary parts of the positive-helicity basis vector

$$-(\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} + i\hat{\mathbf{e}}_{\mathbf{k}}^{(2)})e^{-i\omega_k t}$$

represent a counterclockwise rotation of the polarization vector, provided that $\hat{\mathbf{e}}_{\mathbf{k}}^{(1)}\times\hat{\mathbf{e}}_{\mathbf{k}}^{(2)} = \mathbf{k}$.

Exercise 23.9. Work out the equal-time commutation relations between the components of $\mathbf{A}(\mathbf{r}, t)$ and $\mathbf{E}(\mathbf{r}, t)$.

3. ***Interactions with Charged Particles.*** So far, only free photons have been considered, and it has been shown that these particles are the quanta of the free electromagnetic field. With this background, we can now introduce the interaction between photons and electrons or other charged particles. Although in many applications it is necessary to use a proper relativistic treatment of the massive particles, a nonrelativistic approximation will be used in this section for describing the electron.

[4]Jackson (1975), Section 7.2.

Following the example established in Section 4.6 for the dynamics of a charged particle in an external electromagnetic field, we assume that the Hamiltonian operator is obtained from the Hamiltonian for the free particle system by the substitution

$$\mathbf{p} \to \mathbf{p} + \frac{e}{c}\mathbf{A}$$

if the particles have charge $-e$. This prescription leads to the Hamiltonian

$$\mathcal{H} = \frac{1}{2\pi}\int \mathbf{E}^{(-)} \cdot \mathbf{E}^{(+)}\, d^3r + \frac{1}{2m}\int \psi^\dagger\left(\frac{\hbar}{i}\nabla + \frac{e}{c}\mathbf{A}\right)^2\psi\, d^3r + \int V\psi^\dagger\psi\, d^3r \quad (23.39)$$

for the interacting field operators.

The first term on the right-hand side of (23.39) refers to the electromagnetic radiation field alone. The last term describes the external forces as well as the static interactions between the particles, such as the Coulomb repulsion between electrons, which has been conveniently separated from the total electromagnetic field for inclusion with the unperturbed Hamiltonian. The middle term, when expanded, contributes one term that refers solely to the electrons. The remaining terms represent explicitly the interaction between radiation and matter fields, which is our main focus here. The Hamiltonian is naturally divided into two parts, $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1$, with $\mathcal{H}_0$ including the pure radiation and matter terms, while $\mathcal{H}_1$ symbolizes the interaction terms:

$$\mathcal{H}_0 = \frac{1}{2\pi}\int \mathbf{E}^{(-)} \cdot \mathbf{E}^{(+)}\, d^3r - \frac{\hbar^2}{2m}\int \psi^\dagger\nabla^2\psi\, d^3r + \int V\psi^\dagger\psi\, d^3r \quad (23.40)$$

$$\mathcal{H}_1 = \frac{e\hbar}{mci}\int \psi^\dagger\mathbf{A}\cdot\nabla\psi\, d^3r + \frac{e^2}{2mc^2}\int \mathbf{A}\cdot\mathbf{A}\psi^\dagger\psi\, d^3r \quad (23.41)$$

In deriving (23.40) and (23.41) from (23.39), we have assumed that the Coulomb gauge condition (23.18) has been imposed, even if the fields are interacting and represent particles that are not free.

The dynamics of *interacting fields*, governed by the Hamiltonian (23.39), is described by field operators whose structure is far more complicated than the free fields introduced in Section 23.1. The theory of interacting quantum fields, with its impressive achievements in the high-precision calculations of radiative corrections in quantum electrodynamics, lies outside the scope of this book.[5] We restrict ourselves to two simplified approaches: problems for which a first-order approximation is appropriate, and problems in which the prescribed motion of the charges may be treated by classical methods.

Since the electromagnetic interaction between the radiation field and an electron is comparatively weak, the use of perturbation theory is appropriate in many practical problems. As an illustration, we will derive the formulas for the intensity of *emission* and *absorption* of photons by a one-electron atom, neglecting the spin of the electron. The operator $\mathcal{H}_0$ describing the free photons and the atom is the unperturbed Hamiltonian, and the interaction $\mathcal{H}_1$ is the perturbation to which the general methods of Chapter 19 will be applied.

In first approximation, the first term in the interaction (23.41), which is linear in photon creation and annihilation operators, is responsible for transitions in which *one* photon is emitted or absorbed. If the electron is initially in an atomic state

[5]Among many excellent references, we mention Bjorken and Drell (1964 and 1965), Mandl and Shaw (1984), and Weinberg (1995).

defined by the Schrödinger wave function $\psi_i(\mathbf{r})$, the initial state of the electron-photon system is, in a somewhat hybrid notation, given by

$$|\Psi_i\rangle = \int \psi_i(\mathbf{r})\psi^\dagger(\mathbf{r})\, d^3r\, |\mathbf{0}\rangle_{el} \otimes |\ldots n_i^{(\lambda)}(\mathbf{k}) \ldots\rangle \tag{23.42}$$

Here $|\mathbf{0}\rangle_{el}$ denotes the no-electron state (electron vacuum), and $|\ldots n_i^{(\lambda)}(\mathbf{k}) \ldots\rangle$ symbolizes the state of the electromagnetic field in terms of photon occupation numbers. The superscript λ specifies the polarization of a photon with wave vector $\mathbf{k}$. In the final state there must be one photon more or one photon less than in the initial state, and the electron is annihilated in state $\psi_i(\mathbf{r})$ and recreated in state $\psi_f(\mathbf{r})$.

In order to evaluate the matrix element of the perturbation term for the transition from state i to f, we substitute the expression for the field operator,

$$\begin{aligned}\mathbf{A}(\mathbf{r}, 0) = \sqrt{4\pi\hbar c^2}\,\frac{1}{L^{3/2}} \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_k}} \{&[a_1(\mathbf{k})\hat{e}_{\mathbf{k}}^{(1)} + a_2(\mathbf{k})\hat{e}_{\mathbf{k}}^{(2)}]e^{i\mathbf{k}\cdot\mathbf{r}} \\ &+ [a_1^\dagger(\mathbf{k})\hat{e}_{\mathbf{k}}^{(1)} + a_2^\dagger(\mathbf{k})\hat{e}_{\mathbf{k}}^{(2)}]e^{-i\mathbf{k}\cdot\mathbf{r}}\}\end{aligned} \tag{23.43}$$

obtained from (23.17), into the interaction (23.41). With this interaction, we calculate the matrix element $\langle\Psi_f|\mathcal{H}_1|\Psi_i\rangle$ for the initial state (23.42) and the final state

$$|\Psi_f\rangle = \int \psi_f(\mathbf{r})\psi^\dagger(\mathbf{r})\, d^3r |\mathbf{0}\rangle_{el} \otimes |\ldots, n_i^{(\lambda)}(\mathbf{k}) \mp 1, \ldots\rangle \tag{23.44}$$

The result is

$$\langle\Psi_f|\mathcal{H}|\Psi_i\rangle = \frac{e}{mc}\sqrt{4\pi\hbar c^2}\,\frac{1}{L^{3/2}}\,\frac{1}{\sqrt{2\omega_k}} \begin{Bmatrix}\sqrt{n_i^{(\lambda)}} \\ \sqrt{n_i^{(\lambda)} + 1}\end{Bmatrix} \int \psi_f^*(\mathbf{r})e^{\pm i\mathbf{k}\cdot\mathbf{r}}\, \hat{\mathbf{e}}_{\mathbf{k}}^{(\lambda)} \cdot \frac{\hbar}{i}\nabla\psi_i(\mathbf{r})\, d^3r \tag{23.45}$$

The upper option in braces applies to absorption of a photon, and the lower option refers to emission. If E_i and E_f denote the initial and final energy of the unperturbed atom, transitions are appreciable only if the photon energy $\hbar\omega_k = E_f - E_i$ or $E_i - E_f$, depending on whether absorption or emission of a photon takes place.

Exercise 23.10. Reproduce the derivation of the matrix element (23.45) for absorption or emission of a photon by a charged-particle system.

To illustrate the use of the typical matrix element (23.45) for evaluating transition rate with the Golden Rule (19.99), we assume that the photon modes in the initial state (23.42) are not selectively occupied. If, as in masers and lasers, the occupation number, n_i, for some modes is large, but small or zero for neighboring modes, the Golden Rule, which presupposes a weak dependence of the matrix element on the occupation number, may be wholly inappropriate. We already remarked on this possibility in Section 19.5.

Under the usual conditions of observation, the state of the electromagnetic field corresponds to a superposition of occupied photon states in a quasicontinuum, and it is appropriate to characterize this state by an initial average photon number $\bar{n}_i^{(\lambda)}(\mathbf{k})$ for photons in the polarization mode λ and with approximate momentum $\hbar\mathbf{k}$ pointing in a solid angle $d\Omega_{\mathbf{k}}$. Since the number of photon modes for each polarization per unit energy interval is given by the density of states,

$$\rho(E) = \frac{\omega^2 L^3}{8\pi^3\hbar c^3}\, d\Omega_{\mathbf{k}} \tag{23.46}$$

the transition probability per unit time is according to the Golden Rule:

$$\boxed{w = \frac{2\pi}{\hbar}\rho(\hbar\omega_k)\left|\langle\Psi_f|\mathcal{H}_1|\Psi_i\rangle\right|^2 = \frac{\alpha}{2\pi c^2}\,\omega_k \bar{n}_{i,f}^{(\lambda)}(\mathbf{k})\left|\langle f|e^{\pm i\mathbf{k}\cdot\mathbf{r}}\,\hat{\mathbf{e}}_k^{(\lambda)}\cdot\mathbf{v}|i\rangle\right|^2 d\Omega_{\mathbf{k}}} \tag{23.47}$$

where $\alpha = e^2/\hbar c$ is the fine structure constant, and

$$\langle f|e^{\pm i\mathbf{k}\cdot\mathbf{r}}\hat{\mathbf{e}}_{\mathbf{k}}^{(\lambda)}\cdot\mathbf{v}|i\rangle = \frac{1}{m}\int \psi_f^*(\mathbf{r})\,e^{\pm i\mathbf{k}\cdot\mathbf{r}}\,\hat{\mathbf{e}}_{\mathbf{k}}^{(\lambda)}\cdot\frac{\hbar}{i}\nabla\psi_i(\mathbf{r})\,d^3r \tag{23.48}$$

is the one-particle matrix element that determines the intensity of the transition. In formula (23.47) $\bar{n}_i$ is to be used for absorption and $\bar{n}_f$ for emission of a photon.

The incident photon flux dI_0 is the number of photons incident on the atomic target per unit area and unit time in the frequency interval $d\omega$. As explained in Section 13.1,

$$dI_0 = \frac{\bar{n}}{L^3}\,c\hbar\rho(\hbar\omega)\,d\omega = \frac{\bar{n}\omega^2\,d\omega}{8\pi^3c^2}\,d\Omega_{\mathbf{k}} \tag{23.49}$$

We define a cross section $\sigma(\omega)$ that is sharply peaked at $\omega = (E_f - E_i)/\hbar$ [or $\omega = (E_i - E_f)/\hbar$] in an interval $\Delta\omega$ for photon absorption (or stimulated emission) by the relation

$$w = \int \sigma(\omega)\,dI_0 \approx \frac{dI_0}{d\omega}\int_{\Delta\omega}\sigma(\omega)\,d\omega \tag{23.50}$$

If (23.47) and (23.49) are substituted in (23.50), the cross section formulas obtained in Sections 19.5, and 19.7 by treating the radiation field as classical and prescribed are recovered exactly. The rates of transitions of higher multipolarity may be computed by similar methods.

Exercise 23.11. Check that the integrated cross section (19.62) is obtained from (23.47), (23.49), and (23.50).

If $n_i = 0$, formula (23.47) accounts for the rate of *spontaneous* emission of photons in the transition from an excited state of the radiating atom. If the transition is assumed to be of the *electric dipole* type ($e^{i\mathbf{k}\cdot\mathbf{r}} \approx 1$) and the substitution (19.66) is made for the matrix element, the transition rate for spontaneous emission of light with a polarization vector $\mathbf{e}_{\mathbf{k}}^{(\lambda)}$ and frequency ω into the solid angle $d\Omega_{\mathbf{k}}$ becomes

$$w = \frac{\alpha}{2\pi c^2}\,\omega^3\left|\langle f|\hat{\mathbf{e}}_{\mathbf{k}}^{(\lambda)}\cdot\mathbf{r}|i\rangle\right|^2 d\Omega_{\mathbf{k}} = \frac{\alpha}{2\pi c^2}\,\omega^3\langle f|\hat{\mathbf{e}}_{\mathbf{k}}^{(\lambda)}\cdot\mathbf{r}|i\rangle\langle i|\hat{\mathbf{e}}_{\mathbf{k}}^{(\lambda)}\cdot\mathbf{r}|f\rangle\,d\Omega_{\mathbf{k}}$$

The total integrated emission rate $\sum w$ is obtained by summing over the two polarizations and integrating over all angles of emission. Noting that in dyadic notation,

$$\int\left[\hat{\mathbf{e}}_{\mathbf{k}}^{(1)}\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} + \hat{\mathbf{e}}_{\mathbf{k}}^{(2)}\hat{\mathbf{e}}_{\mathbf{k}}^{(2)}\right]d\Omega_{\mathbf{k}} = \int\left[I - \hat{\mathbf{k}}\hat{\mathbf{k}}\right]d\Omega_{\mathbf{k}} = \frac{8\pi}{3}\,I \tag{23.51}$$

we obtain

$$\begin{aligned}\sum w &= \frac{4\alpha}{3c^2}\,\omega^3\left|\langle f|\mathbf{r}|i\rangle\right|^2 \\ &= \frac{4\alpha}{3c^2}\,\omega^3\{\left|\langle f|\mathbf{r}\cdot\hat{\mathbf{x}}|i\rangle\right|^2 + \left|\langle f|\mathbf{r}\cdot\hat{\mathbf{y}}|i\rangle\right|^2 + \left|\langle f|\mathbf{r}\cdot\hat{\mathbf{z}}|i\rangle\right|^2\}\end{aligned} \tag{23.52}$$

In atomic spectroscopy, this expression is for historical reasons often referred to as the Einstein A coefficient.

If the initial state of the radiating one-electron atom is characterized by the quantum numbers $n\ell m$, the expression (23.52) can be summed over the magnetic quantum numbers of the final state, using the relations (19.72) and (19.78). Since

$$|Y_1^m(\hat{\mathbf{x}})|^2 + |Y_1^m(\hat{\mathbf{y}})|^2 + |Y_1^m(\hat{\mathbf{z}})|^2 = \frac{3}{4\pi} \tag{23.53}$$

we obtain for the spontaneous emission rate from the energy level $n\ell$ to $n'\ell \pm 1$,

$$\boxed{\sum w = \frac{4\alpha}{3c^2}\,\omega^2 \times \begin{cases} \dfrac{\ell+1}{2\ell+1}\,(R_{n\ell}^{n'\ell+1})^2 \\ \dfrac{\ell}{2\ell+1}\,(R_{n\ell}^{n'\ell-1})^2 \end{cases}} \tag{23.54}$$

corresponding to the selection rule $\Delta\ell = \ell' - \ell = \pm 1$, which is the same here as in Section 19.5.

Exercise 23.12. Using the hydrogenic wave functions (12.92), compute the spontaneous emission rate for the $2p \rightarrow 1s$ transition in the H atom. Evaluate its reciprocal, the mean lifetime of the $2p$ state.

4. ***Elements of Quantum Optics.*** Proper higher-order perturbation calculations on interacting radiation and matter fields, beyond the first-order processes treated in the last section, belong in the domain of quantum electrodynamics. However, if the matter field that creates and annihilates photons can be approximated by a *prescribed* classical current distribution, the dynamics of the radiation field is greatly simplified, and it is not always necessary to resort to perturbation expansions. This strategy is successful because the free electromagnetic field is representable as a system of independent harmonic oscillators whose interaction with the matter field is linear in the photon creation and annihilation operators if the second, quadratic, term in the interaction (23.41) is neglected. Such a semiclassical theory is thus equivalent to a system of *forced* harmonic oscillators, for which we have already worked out exact solutions of the dynamical equation in Section 14.6. The theory that treats the matter field classically, but the electromagnetic field as a quantum system, complements the approach taken in Chapter 19, where we considered the effect of a classical radiation field on an atom, albeit using perturbation theory.

Here we suppose that the Hamiltonian of the radiation field coupled to a prescribed classical current density $\mathbf{j}(\mathbf{r}, t)$ is obtained from (23.39) by letting the material carrier of the current become very massive ($m \rightarrow \infty$) while keeping $\mathbf{j}(\mathbf{r}, t)$ finite:

$$\mathcal{H} = \frac{1}{2\pi}\int \mathbf{E}^{(-)} \cdot \mathbf{E}^{(+)}\, d^3r - \frac{1}{c}\int \mathbf{j}(\mathbf{r}, t) \cdot \mathbf{A}(\mathbf{r})\, d^3r \tag{23.55}$$

This expression is transformed in the momentum representation as

$$\mathcal{H} = \sum_{\mathbf{k}} \hbar\omega_k [a_1^\dagger(\mathbf{k})a_1(\mathbf{k}) + a_2^\dagger(\mathbf{k})a_2(\mathbf{k})]$$
$$- \sqrt{4\pi\hbar} \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_k}} \mathbf{j}(\mathbf{k}, t) \cdot [a_1(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} + a_2\hat{\mathbf{e}}_{\mathbf{k}}^{(2)}] \tag{23.56}$$
$$- \sqrt{4\pi\hbar} \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_k}} \mathbf{j}^*(\mathbf{k}, t) \cdot [a_1^\dagger(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(1)} + a_2^\dagger(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(2)}]$$

where

$$\mathbf{j}(\mathbf{k}, t) = \frac{1}{L^{3/2}} \int \mathbf{j}(\mathbf{r}, t) e^{i\mathbf{k}\cdot\mathbf{r}}\, d^3r \tag{23.57}$$

is the Fourier transform of the applied current.

The Hamiltonian (23.56) is a sum of independent contributions from each photon mode with wave vector $\mathbf{k}$ and definite polarization. Omitting, for brevity, all reference to the particular mode, each addend is of the form

$$\hbar\omega a^\dagger a + f(t)a + f^*(t)a^\dagger$$

which, except for the missing zero-point energy, is the same as the Hamiltonian (14.106) for the forced linear harmonic oscillator. The complex-valued function of time $f(t)$ stands for

$$f(t) = -\sqrt{\frac{2\pi\hbar}{\omega_k}}\, \mathbf{j}(\mathbf{k}, t) \cdot \hat{\mathbf{e}}_{\mathbf{k}}^{(\lambda)} \tag{23.58}$$

Our interest will be focused on the changes produced by a pulsed current distribution which is effective only during a finite time interval but vanishes in the distant past and the remote future as $t \to \pm\infty$. It is convenient to assume that $f(t) \neq 0$ only for $T_2 > t > T_1 > 0$. Therefore, the electromagnetic field is free before T_1 and after T_2.

Since the current is assumed to be known and susceptible to a classical description, each photon mode develops independently in time, and the calculations of Section 14.6 become directly applicable. If the electromagnetic field is in its ground (or vacuum) state before the onset of the interaction, the final state will be

$$S|\mathbf{0}\rangle = \exp\left[-\pi \sum_{\mathbf{k},\lambda} \frac{1}{\hbar\omega_k} |j(\mathbf{k}, \omega_k, \lambda)|^2\right] \exp\left[i \sum_{k,\lambda} \sqrt{\frac{2\pi}{\hbar\omega_k}}\, j(\mathbf{k}, \omega_k, \lambda) a_\lambda^\dagger(\mathbf{k})\right] |\mathbf{0}\rangle \tag{23.59}$$

in the interaction picture. Here we define the Fourier integral,

$$j(\mathbf{k}, \omega_k, \lambda) = \frac{1}{L^{3/2}} \int_{-\infty}^{+\infty} dt \int d^3r\, e^{i(\omega t - \mathbf{k}\cdot\mathbf{r})} \mathbf{j}(\mathbf{r}, t) \cdot \hat{\mathbf{e}}_{\mathbf{k}}^{(\lambda)} \tag{23.60}$$

Exercise 23.13. Verify the result (23.59) by factoring S in (14.143) and using it for each independent mode.

The final state (23.59) is a direct product of *coherent states*—one for each excited mode—since for each single mode,

$$e^{\alpha a^\dagger}|0\rangle = e^{|\alpha|^2/2}|\alpha\rangle$$

The eigenvalues α are proportional to the Fourier transforms of the current density, (23.60). The state is also an eigenstate of the positive-frequency operators $\mathbf{A}^{(+)}(\mathbf{r})$ and $\mathbf{E}^{(+)}(\mathbf{r})$. It can be verified that the expectation value of $\mathbf{A}(\mathbf{r})$ in the state (23.59) is identical with the classical retarded vector potential generated, in the Coulomb gauge, by the prescribed current distribution. This is as it should be according to the correspondence principle.

As we saw in Section 14.6 on the forced oscillator, each coherently excited mode, with specified $\mathbf{k}$ and polarization, is populated by photons according to a Poisson distribution. The successive emissions of photons from a prescribed current distribution may, therefore, be regarded as statistically independent events, and this behavior is relevant in quantum optics and quantum electronics.

Exercise 23.14. Show that the total probability for emitting from a prescribed classical current distribution $\mathbf{j}(\mathbf{r}, t)$ a net total number of n photons into a specified set R of photon modes, and none into any other mode, is given again by a Poisson distribution.

$$P_n(R) = e^{-\bar{n}} \frac{r^n}{n!} \tag{23.61}$$

where $\bar{n}$ is the mean number of photons emitted into all modes. Use the multinomial expansion for the calculation, and evaluate $\bar{n}$ and r in terms of the Fourier components of the applied current density.

For large mean photon numbers, coherent states like (23.59) are good approximations to the states of the electromagnetic field excited by an ideal laser. The most characteristic feature of any coherent state $|\Psi_{\text{coh}}\rangle$ like (23.59) is the factorization property for the expectation value of any normally ordered product of field operators:

$$\begin{aligned}
&\langle\Psi_{\text{coh}}| \prod_i \mathbf{E}^{(-)}(\mathbf{r}_i, t_i) \prod_j \mathbf{E}^{(+)}(\mathbf{r}_j, t_j)|\Psi_{\text{coh}}\rangle \\
&\quad = \prod_i \langle\Psi_{\text{coh}}|\mathbf{E}^{(-)}(\mathbf{r}_i, t_i)|\Psi_{\text{coh}}\rangle \prod_j \langle\Psi_{\text{coh}}|\mathbf{E}^{(+)}(\mathbf{r}_j, t_j)|\Psi_{\text{coh}}\rangle \\
&\quad = \prod_i \mathbf{E}'^{(-)}(\mathbf{r}_i, t_i) \prod_j \mathbf{E}'^{(-)*}(\mathbf{r}_j, t_j)
\end{aligned} \tag{23.62}$$

Here the $\mathbf{E}'^{(-)}(\mathbf{r}, t)$ are eigenvalues of the field operators, corresponding to the coherent eigenstate $|\Psi_{\text{coh}}\rangle$ and propagating freely after the interaction has ceased. In particular, the energy density of the radiation field for a coherent state is proportional to

$$\begin{aligned}
\langle\Psi_{\text{coh}}|\mathbf{E}^{(-)}(\mathbf{r}) \cdot \mathbf{E}^{(+)}(\mathbf{r})|\Psi_{\text{coh}}\rangle &= \langle\Psi_{\text{coh}}|\mathbf{E}^{(-)}(\mathbf{r})|\Psi_{\text{coh}}\rangle \cdot \langle\Psi_{\text{coh}}|\mathbf{E}^{(+)}(\mathbf{r})|\Psi_{\text{coh}}\rangle \\
&= \mathbf{E}'^{(-)}(\mathbf{r}) \cdot \mathbf{E}'^{(+)}(\mathbf{r})
\end{aligned} \tag{23.63}$$

showing that in a coherent, or quasiclassical, state the field intensity may be calculated by treating the expectation values of $\mathbf{E}^{(+)}$ and $\mathbf{E}^{(-)}$ like classical fields.

For a single plane wave mode, the field is according to (23.22) effectively

$$\mathbf{E}^{(+)}(\mathbf{r}, t) \to \frac{i\sqrt{2\pi\hbar\omega_k}}{L^{3/2}} a_\lambda(\mathbf{k})\hat{\mathbf{e}}_{\mathbf{k}}^{(\lambda)} e^{i(\mathbf{k}\cdot\mathbf{r} - \omega_k t)} \tag{23.64}$$

This is a single-mode field operator in the Heisenberg picture. By setting $t = 0$, we obtain its Schrödinger picture version, assuming the field to be free.

Exercise 23.15. Calculate the expectation value of the full electric field operator

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}^{(+)}(\mathbf{r}, t) + \mathbf{E}^{(-)}(\mathbf{r}, t)$$

in a single-mode coherent state. Show that the result resembles a classical plane wave. Also work out the square root of the variance (or uncertainty) of the field amplitude $\mathbf{E} \cdot \hat{\mathbf{e}}_{\mathbf{k}}^{(\lambda)}$.

Expectation values, like (23.62), of products of field operators for general states of the radiation field are significant for characterizing the measurable properties of a system of photons. The rate at which photons are usually detected at a point in spacetime through absorption by atoms or other systems of charges in the ground state is to a good approximation proportional to the squared transition matrix element

$$|\langle \Psi_f | \mathbf{E}^{(+)}(\mathbf{r}, t) | \Psi_i \rangle|^2 \tag{23.65}$$

Here Ψ_i and Ψ_f denote the initial and final states of the radiation field, with Ψ_f having one photon less than Ψ_i. An ideal one-photon detector is equally sensitive to all photons present at point $(\mathbf{r}, t)$ in the initial state. For such a detector, the sum of (23.65) over *all* final states of the radiation field may, by use of the closure property for a complete set of states, be written as

$$\begin{aligned} \sum_f |\langle \Psi_f | \mathbf{E}^{(+)}(\mathbf{r}, t) | \Psi_i \rangle|^2 &= \sum_f \langle \Psi_i | \mathbf{E}^{(-)}(\mathbf{r}, t) | \Psi_f \rangle \langle \Psi_f | \mathbf{E}^{(+)}(\mathbf{r}, t) | \Psi_i \rangle \\ &= \langle \Psi_i | \mathbf{E}^{(-)}(\mathbf{r}, t) \cdot \mathbf{E}^{(+)}(\mathbf{r}, t) | \Psi_i \rangle \end{aligned} \tag{23.66}$$

which is the expected field intensity at spacetime point $\mathbf{r}$, t. In the next section, we apply these concepts to a quantum mechanical analysis of optical interference phenomena, and especially the familiar two-slit interference experiment.

5. ***Coherence, Interference, and Statistical Properties of the Field.*** In an interference experiment, two (or more) field modes are simultaneously excited, creating a state of the photon field as a superposition of several photon-number basis states $|n_1, n_2, \ldots, n_{\mathbf{k}}, \ldots\rangle$, and the state of the field is probed at various points in space or time. We speak of interference when the observed photon detection rate differs from the sum of the individual intensities expected for the separate modes, and we attribute interference to *coherence* properties of the system. The term *coherence* has acquired many different meanings in physics, and caution is in order when using it. For example, a pure state that is a linear combination of two or more component states is sometimes redundantly called a *coherent* superposition. On the other hand, as a technical term we use the name *coherent state* generally for an eigenstate of an annihilation operator and, in the case of the radiation field, specifically an eigenstate of $\mathbf{E}^{(+)}$. In this section, the degree of first-order (spacetime) coherence of the field will be defined.

Following the classical treatment of interference, two or more field modes that are excited are often thought of as distinct radiation fields that are being superposed. In quantum mechanics, such a view can easily and inadvertently lead to the erroneous conclusion that the two or more components of a superposition always represent distinct photons. In the interest of conceptual clarity, we regard the radiation field as a single spacetime-dependent operator, which is a superposition of all modes defined by appropriate boundary conditions, whether or not they are excited. If in-

terference is observed, the particular state as much as the omnipresent field operator is responsible for its occurrence.

The art of (photon, neutron, electron, atom, or molecular) interferometry requires considerable skill to produce the desired superpositions of simple modes, and much effort must go into this first stage of an interference experiment. In a typical (Young-type) two-slit experiment, two neighboring modes are excited and made to interfere with amplitudes whose relative magnitude depends on the widths of the slits and whose relative phase is a function of the geometric arrangement (see Figure 1.1). Up to this point, the analysis involving the field, with its modes determined by the solutions of Maxwell's equations subject to appropriate boundary conditions, is identical with the procedure followed in classical optics. Coherent superpositions of field modes are also readily produced by many other techniques, such as amplitude or beam splitters, and nonlinear optical devices that can change the number of photons in the state.

Assuming that the desired photon state has been generated, here we are concerned with the second stage of an interference experiment, the detection of photons at a spacetime point. We first consider the common situation of effective *two-mode excitation* and write the relevant basis states as $|n_1, n_2\rangle$, omitting for brevity all reference to the remaining inert modes that are not being observed and may be assumed to have zero photon occupation. If just *one photon* is present in the two modes of interest, the normalized state vector has the general form

$$|\Psi\rangle = \cos\theta|1, 0\rangle + e^{i\phi}\sin\theta|0, 1\rangle \tag{23.67}$$

The amplitudes $\cos\theta$ and $e^{i\phi}\sin\theta$ are determined by the experimental conditions, which define the preparation of the state. For example, in a two-slit interference setup with equal slits,

$$\cos\theta = \sin\theta = \frac{1}{\sqrt{2}}$$

and the phase ϕ is determined by the angle of incidence of the plane-wave photon onto the plane of the slits. The relevant part of the field operator is

$$\mathbf{E}^{(+)}(\mathbf{r}t) \to u_1(\mathbf{r}t)e^{i\delta_1(\mathbf{r}t)}a_1 + u_2(\mathbf{r}t)e^{i\delta_2(\mathbf{r}t)}a_2 \tag{23.68}$$

It is determined by the boundary conditions and the mode frequency, and is a function of the spacetime point $\mathbf{r}t$ at which the field is being probed. For simplicity in evaluating the intensity, the vector polarization character of the field is ignored on the right-hand side and we pretend that the field is a scalar. The amplitudes $u_{1,2}(\mathbf{r}t)$ are real and nonnegative, and the $\delta_{1,2}(\mathbf{r}t)$ are real phases. (To emphasize the parallelism of space and time coordinates in the present context, we omit the comma between $\mathbf{r}$ and t. This convention will be continued in the remainder of this chapter.) Again, we should resist the temptation of thinking of the terms in (23.68) as two different radiation fields.

The observed intensity is, apart from an overall normalization factor,

$$\begin{aligned} I(\mathbf{r}t) = \langle\Psi|\mathbf{E}^{(-)}(\mathbf{r}t)\cdot\mathbf{E}^{(+)}(\mathbf{r}t)|\Psi\rangle &\to u_1^2\langle\Psi|a_1^\dagger a_1|\Psi\rangle + u_2^2\langle\Psi|a_2^\dagger a_2|\Psi\rangle \\ &\quad + 2u_1u_2\,\mathrm{Re}[e^{i\delta}\langle\Psi|a_1^\dagger a_2|\Psi\rangle] \\ &= (\langle 1, 0|\cos\theta + \langle 0, 1|e^{-i\phi}\sin\theta)[u_1^2a_1^\dagger a_1 + u_2^2a_2^\dagger a_2 \\ &\quad + u_1u_2(e^{i\delta}a_1^\dagger a_2 + e^{-i\delta}a_2^\dagger a_1)](\cos\theta|1, 0\rangle + e^{i\phi}\sin\theta|0, 1\rangle) \\ &= u_1^2\cos^2\theta\langle 1, 0|a_1^\dagger a_1|1, 0\rangle + u_2^2\sin^2\theta\langle 0, 1|a_2^\dagger a_2|0, 1\rangle \\ &\quad + u_1u_2\sin 2\theta\,\mathrm{Re}e^{i(\phi+\delta)}\langle 1, 0|a_1^\dagger a_2|0, 1\rangle \end{aligned}$$

The last term accounts for interference. The relative intensity can be simply written in the form:

$$\boxed{I(\mathbf{r}t) = 1 + \frac{u_1 u_2}{u_1^2 \cos^2 \theta + u_2^2 \sin^2 \theta} \sin 2\theta \cos(\phi + \delta)} \tag{23.69}$$

The quantity $\delta = \delta_2 - \delta_1$ is the relative phase of the two modes that are relevant in this intensity measurement. In the conventional two-slit experiment, where owing to the method of photon production from a single source both relevant modes have the same frequency, the observed intensity is determined by the position of the detector that eventually destroys the state. As Figure 1.1 shows schematically, $\delta = k(s_1 - s_2)$ is proportional to the difference of the path lengths, s_1 and s_2, between the detector and the two slits. To observe the intensity variations of the phase $\phi + \delta$ as a function of $\mathbf{r}t$, the experimental arrangement must ensure that $(u_1/u_2)\, e^{-i\delta}$ varies sufficiently slowly in space and time. In a Young-type interference experiment, this is achieved by making the distance between the slits small and the distance between the slits and the detector large. The preceding analysis shows that in the simple architecture considered here a *single* photon can be said to be responsible for interference, reminding us of Dirac's famous remark that (under single-photon two-slit conditions) a photon interferes with itself.

In general, however, if the state of the radiation field is suitably generated, more than one photon is involved in causing interference. For example, if we assume that two modes are excited into coherent, or semiclassical, many-photon states:

$$|\Psi\rangle = |\alpha_1\rangle \otimes |\alpha_2\rangle \tag{23.70}$$

the intensity is

$$\begin{aligned} I(\mathbf{r}t) &= \langle \alpha_1, \alpha_2 | (u_1 e^{-i\delta_1} a_1^\dagger + u_2 e^{-i\delta_2} a_2^\dagger)(u_1 e^{i\delta_1} a_1 + u_2 e^{i\delta_2} a_2) | \alpha_1, \alpha_2 \rangle \\ &= u_1^2 |\alpha_1|^2 + u_2^2 |\alpha_2|^2 + 2u_1 u_2 |\alpha_1 \alpha_2| \cos(\delta + \phi) \end{aligned} \tag{23.71}$$

where ϕ is now understood to be the phase difference between the coherent state eigenvalues α_1 and α_2. If α_1 and α_2 are thought of as the classical radiation fields, formula (23.71) is the familiar result of classical wave theory.

The intensity pattern is measured at spacetime point $\mathbf{r}t$, but since two modes are involved, the interference measurement conveys information about the normalized *correlation functions* for the field at two appropriately chosen spacetime reference points $\mathbf{r}_1 t_1$ and $\mathbf{r}_2 t_2$ (such as the locations of the two slits). The first-order correlation function is defined as the dimensionless quantity,

$$\boxed{g^{(1)}(\mathbf{r}_1 t_1; \mathbf{r}_2 t_2) = \frac{\langle \Psi | \mathbf{E}^{(-)}(\mathbf{r}_1 t_1) \cdot \mathbf{E}^{(+)}(\mathbf{r}_2 t_2) | \Psi \rangle}{\sqrt{\langle \Psi | \mathbf{E}^{(-)}(\mathbf{r}_1 t_1) \cdot \mathbf{E}^{(+)}(\mathbf{r}_1 t_1) | \Psi \rangle \langle \Psi | \mathbf{E}^{(-)}(\mathbf{r}_2 t_2) \cdot \mathbf{E}^{(+)}(\mathbf{r}_2 t_2) | \Psi \rangle}}} \tag{23.72}$$

In a well-designed interference measurement, the active modes are experimentally distinguishable. Bearing in mind that the mode functions are idealizations approximating classical wave packets of finite extent, we may suppose that the reference points effectively satisfy the separability condition

$$u_1(\mathbf{r}_2 t_2) = u_2(\mathbf{r}_1 t_1) \approx 0 \tag{23.73}$$

In the two-slit experiment, the separate slits are conveniently taken as the reference points in space, and the reference times have the retarded values

$$t_1 = t - \frac{s_1}{c} \quad \text{and} \quad t_2 = t - \frac{s_2}{c} \tag{23.74}$$

Effectively, the expression (23.68) can then be replaced by

$$\mathbf{E}^{(+)}(\mathbf{r}t) = \mathbf{E}_1^{(+)}(\mathbf{r}_1 t_1) + \mathbf{E}_2^{(+)}(\mathbf{r}_2 t_2) \to u_1(\mathbf{r}_1 t_1) e^{i\delta(\mathbf{r}_1 t_1)} a_1 + u_2(\mathbf{r}_2 t_2) e^{i\delta(\mathbf{r}_2 t_2)} a_2 \tag{23.75}$$

Under the conditions specified for Young-type interference, the interference term in the expression for the intensity is seen to be proportional to the normalized correlation function (23.72) for the field at the two slits at the reference times (23.74).

The magnitude $|g^{(1)}(\mathbf{r}_1 t_1, \mathbf{r}_2 t_2)|$ is called the *degree of first-order coherence* of the state of the radiation field. If we set $\langle \Psi_a | = \langle \Psi | \mathbf{E}^{(-)}(\mathbf{r}_1 t_1)$ and $|\Psi_b\rangle = \mathbf{E}^{(+)}(\mathbf{r}_2 t_2)|\Psi\rangle$, the Schwarz inequality (9.75) shows that $0 \leq |g^{(1)}| \leq 1$. If $|g^{(1)}| = 1$, the field is said to possess *first-order coherence*. If $0 < |g^{(1)}| < 1$, we have *partial coherence*, and $g^{(1)} = 0$ corresponds to an *incoherent* field.

If the excitation of all modes is represented by coherent photon states, as in the two-mode example (23.70), the factorization property (23.63) shows that the field has first-order coherence for all pairs of points in spacetime, justifying the term *coherent* for the eigenstates of the annihilation operators.

Exercise 23.16. Check that the one-photon two-mode states (23.67) are first-order coherent for all pairs of spacetime points.

Exercise 23.17. Prove that the two-mode photon-number basis state $|n_1, n_2\rangle$, with $n_{1,2} \neq 0$, has zero first-order coherence at points that satisfy the condition (23.73).

A more complicated higher-order *correlation function* arises when we consider the *joint probability* for detecting a photon in each of two detectors at spacetime points $\mathbf{r}_1 t_1$ and $\mathbf{r}_2 t_2$. In lowest order of perturbation theory, the coincidence rate is proportional to

$$\langle \Psi | \mathbf{E}^{(-)}(\mathbf{r}_1 t_1) \mathbf{E}^{(-)}(\mathbf{r}_2 t_2) \mathbf{E}^{(+)}(\mathbf{r}_2 t_2) \mathbf{E}^{(+)}(\mathbf{r}_1 t_1) | \Psi \rangle$$

When the field operators in such a correlation function are expanded in terms of photon modes, we are always faced with the need to evaluate expectation values of normally ordered products of creation and annihilation operators in equal numbers. These correlation functions measure higher-order degrees of coherence.

The measurement of second-order (temporal and spatial) photon correlations was pioneered by Hanbury Brown and Twiss.[6] For example, in a two-mode configuration they experimentally determined, by photon counting in state Ψ, the difference between the average of the product of the photon numbers, $\langle \Psi | a_1^\dagger a_2^\dagger a_2 a_1 | \Psi \rangle$, and the product of the averages of the photon numbers for the two modes, $\langle \Psi | a_1^\dagger a_1 | \Psi \rangle \langle \Psi | a_2^\dagger a_2 | \Psi \rangle$. If the state of the field is coherent as specified in (23.70), these products are equal and there is no correlation. By introducing the photon creation and annihilation operators,

$$a^\dagger = \cos\theta\, a_1^\dagger + \sin\theta e^{i\phi} a_2^\dagger, \qquad a = \cos\theta\, a_1 + \sin\theta e^{-i\phi} a_2 \tag{23.76}$$

we define, as the generalization of the one-photon state (23.67), an n-photon state $|\Psi_n\rangle$:

$$|\Psi_n\rangle = \frac{1}{\sqrt{n!}} (a^\dagger)^n |\mathbf{0}\rangle = \frac{1}{\sqrt{n!}} (\cos\theta\, a_1^\dagger + \sin\theta e^{i\phi} a_2^\dagger)^n |0, 0\rangle \tag{23.77}$$

[6]Loudon (1983), Section 6.4, and Baym (1969, 1981), Chapter 19.

Such a state can be produced by beamsplitting techniques from an n-photon state in a single mode. Using the binomial expansion, we obtain

$$\begin{aligned} |\Psi_n\rangle &= \sum_{k=0}^{n} \frac{\sqrt{n!}}{k!(n-k)!} \cos^k \theta \sin^{n-k} \theta \, e^{i(n-k)\phi} (a_1^\dagger)^k (a_2^\dagger)^{n-k} |\mathbf{0}\rangle \\ &= \sum_{k=0}^{n} \sqrt{\frac{n!}{k!(n-k)!}} \cos^k \theta \sin^{n-k} \theta \, e^{i(n-k)\phi} |k, n-k\rangle \end{aligned} \tag{23.78}$$

The probability of finding k photons in mode 1 and $n - k$ photons in mode 2 is thus given by the binomial distribution,

$$|\langle k, n-k|\Psi_n\rangle|^2 = \frac{n!}{k!(n-k)!} \cos^{2k} \theta \sin^{2(n-k)} \theta \tag{23.79}$$

Exercise 23.18. Show that

$$\langle \Psi_n | a_1^\dagger a_1 | \Psi_n \rangle = n \cos^2 \theta \quad \text{and} \quad \langle \Psi_n | a_2^\dagger a_1 | \Psi_n \rangle = n \sin^2 \theta \tag{23.80}$$

and

$$\langle \Psi_n | a_1^\dagger a_2^\dagger a_2 a_1 | \Psi_n \rangle = n(n-1) \cos^2 \theta \sin^2 \theta \tag{23.81}$$

From (23.80) and (23.81), the normalized second-order correlation function for the state $|\Psi_n\rangle$ defined by (23.77) is found to be

$$\frac{\langle \Psi_n | a_1^\dagger a_2^\dagger a_2 a_1 | \Psi_n \rangle}{\langle \Psi_n | a_1^\dagger a_1 | \Psi_n \rangle \langle \Psi_n | a_2^\dagger a_2 | \Psi_n \rangle} = 1 - \frac{1}{n} \tag{23.82}$$

Since this expression equals unity for a coherent state, (23.82) shows the degree to which the state $|\Psi_n\rangle$ lacks second-order coherence.

Generally, the state of the radiation field is not a pure state $|\Psi\rangle$, but is a mixture described by a density operator ρ, as defined in Section 15.5. The correlation functions characterizing the radiation field are then given by the averages

$$\begin{aligned} &\langle \mathbf{E}^{(-)}(\mathbf{r}_1 t_1) \ldots \mathbf{E}^{(-)}(\mathbf{r}_n t_n) \mathbf{E}^{(+)}(\mathbf{r}_n t_n) \ldots \mathbf{E}^{(+)}(\mathbf{r}_1 t_1) \rangle \\ &\quad = \text{trace}[\rho \mathbf{E}^{(-)}(\mathbf{r}_1 t_1) \ldots \mathbf{E}^{(-)}(\mathbf{r}_n t_n) \mathbf{E}^{(+)}(\mathbf{r}_n t_n) \ldots \mathbf{E}^{(+)}(\mathbf{r}_1 t_1)] \end{aligned} \tag{23.83}$$

The definition (23.72) of the two-point correlation function and the corresponding degree of first-order coherence must be appropriately amended by using traces to express the expectation values.

As a concrete example (Figure 23.1), we consider light of the same frequency emitted from two point sources, A and B, and the detection of photons at points a and b. The four channels (Aa), (Ab), (Ba), and (Bb) linking the sources and detectors are labeled 1 through 4, respectively, and represent four orthogonal field modes, with their creation and annihilation operators, $a_i^\dagger$ and a_i $(i = 1, 2, 3, 4)$. Source A is assumed to emit single photons into channels 1 and 2 with complex amplitude α. Similarly, source B produces photons in channels 3 and 4 with amplitude β. Thus, the sources are assumed to produce the four one-photon states, $\alpha a_1^\dagger |\mathbf{0}\rangle$, $\alpha a_2^\dagger |\mathbf{0}\rangle$, $\beta a_3^\dagger |\mathbf{0}\rangle$, $\beta a_4^\dagger |\mathbf{0}\rangle$. (In a more general case, the four emission amplitudes could all be different.)

The (unnormalized) density matrix needed to calculate the photon number counted by an ideal detector a is, in the basis spanned by states 1 and 3,

$$\rho_a = \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{pmatrix} \tag{23.84}$$

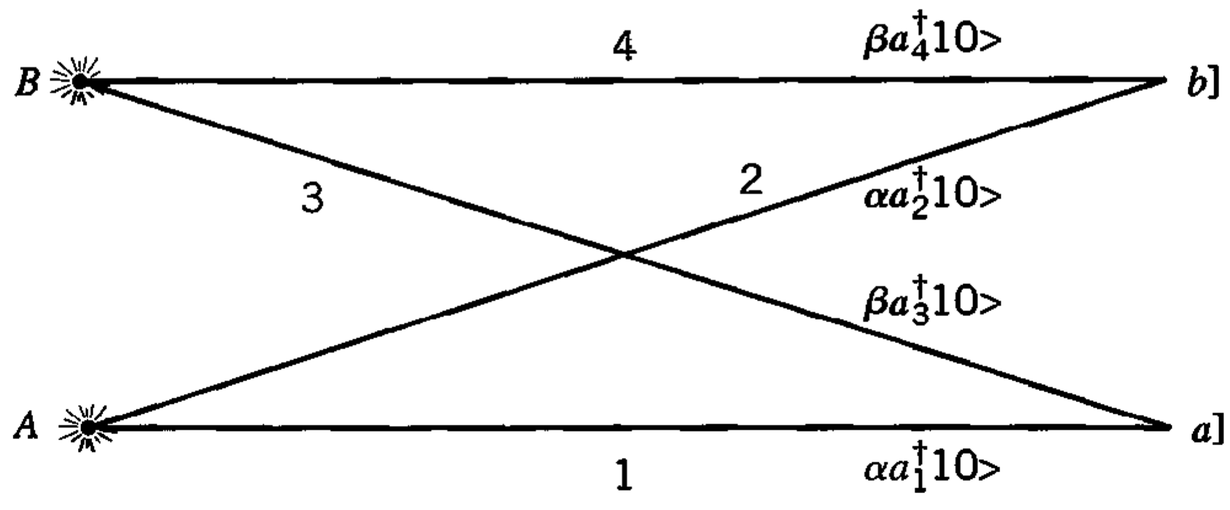


Figure 23.1. Schematic diagram illustrating a spatial Hanbury Brown-Twiss type intensity interference experiment. The two independent light sources are A and B, and the two detectors are labeled a and b. The four channels connecting the sources and detectors are 1 through 4.

If the sources A and B are statistically independent (''incoherent'') and we plan to calculate ensemble averages, (23.84) must be replaced by the density matrix for the ensemble. Since the phases of α and β are random and uncorrelated, the off-diagonal matrix elements of ρ_a average to zero, and we have

$$\rho_a = \begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix} \tag{23.85}$$

The field operator that pertains to detector a is proportional to

$$\mathbf{E}^{(+)}(a) \rightarrow a_1 + e^{i\delta_a} a_3 \tag{23.86}$$

Here we have assumed, merely for simplicity, that the amplitudes $u_{1,3}$ in the notation of (23.64) of the two modes are equal in magnitude (and have been absorbed in the overall normalization), but that the relative phase δ_a is sensitively dependent on the geometry and the position of detector a, just as in classical optics. As before, we also ignore the vector polarization properties of the field. The 2×2 matrix representing the intensity operator at a is

$$\mathbf{E}^{(-)}(a) \cdot \mathbf{E}^{(+)}(a) \rightarrow \begin{pmatrix} 1 & e^{i\delta_a} \\ e^{-i\delta_a} & 1 \end{pmatrix} \tag{23.87}$$

Combining (23.85) with (23.87), we obtain for the suitably normalized average intensity:

$$\begin{aligned} \langle I_a \rangle &= \langle \mathbf{E}^{(-)}(a) \cdot \mathbf{E}^{(+)}(a) \rangle \\ &= \text{trace}\left[\begin{pmatrix} |\alpha|^2 & 0 \\ 0 & |\beta|^2 \end{pmatrix} \begin{pmatrix} 1 & e^{i\delta_a} \\ e^{-i\delta_a} & 1 \end{pmatrix} \right] = |\alpha|^2 + |\beta|^2 \end{aligned} \tag{23.88}$$

which is scarcely a surprising result. Under the simplifying assumptions we have made, the average intensity (or photon detection rate) at detector b is the same: $\langle I_b \rangle = \langle I_a \rangle$.

We now extend the calculation to the less trivial evaluation of the second-order intensity correlation function, or average *coincidence rate*, for the two detectors:

$$\langle I_a I_b \rangle = \langle \mathbf{E}^{(-)}(a) \mathbf{E}^{(-)}(b) \mathbf{E}^{(+)}(b) \mathbf{E}^{(+)}(a) \rangle \tag{23.89}$$

Since we are interested in the state that represents the presence of one photon at each of the two detectors, the four relevant two-photon states are:

$$\alpha^2 a_2^\dagger a_1^\dagger |\mathbf{0}\rangle,\ \beta^2 a_4^\dagger a_3^\dagger |\mathbf{0}\rangle,\ \alpha\beta a_4^\dagger a_1^\dagger |\mathbf{0}\rangle,\ \alpha\beta a_3^\dagger a_2^\dagger |\mathbf{0}\rangle$$

The (unnormalized) density matrix for the statistical ensemble (assuming independent sources) in this representation is

$$\rho_{ab} = \begin{pmatrix} |\alpha|^4 & 0 & 0 & 0 \\ 0 & |\beta|^4 & 0 & 0 \\ 0 & 0 & |\alpha|^2|\beta|^2 & |\alpha|^2|\beta|^2 \\ 0 & 0 & |\alpha|^2|\beta|^2 & |\alpha|^2|\beta|^2 \end{pmatrix} \tag{23.90}$$

The field operator pertaining to detector b is analogous to (23.86):

$$\mathbf{E}^{(+)}(b) \to a_2 + e^{i\delta_b} a_4 \tag{23.91}$$

We use the operators (23.86) and (23.91) to construct the matrix representing the correlation operator for substitution in (23.83):

$$\mathbf{E}^{(-)}(a)\mathbf{E}^{(-)}(b)\mathbf{E}^{(+)}(b)\mathbf{E}^{(+)}(a) \to \begin{pmatrix} 1 & e^{i(\delta_a+\delta_\beta)} & e^{i\delta_b} & e^{i\delta_a} \\ e^{-i(\delta_a+\delta_b)} & 1 & e^{-i\delta_a} & e^{-i\delta_b} \\ e^{-i\delta_b} & e^{i\delta_a} & 1 & e^{i(\delta_a-\beta_b)} \\ e^{-i\delta_a} & e^{i\delta_b} & e^{-i(\delta_a-\delta_b)} & 1 \end{pmatrix} \tag{23.92}$$

The trace of the product of the matrices (23.90) and (23.92) is the correlation function (23.83):

$$\boxed{\begin{aligned} \langle I_a I_b \rangle &= |\alpha|^4 + |\beta|^4 + 2|\alpha|^2|\beta|^2[1 + \cos(\delta_a - \delta_b)] \\ &= \langle I_a \rangle \langle I_b \rangle + 2|\alpha|^2|\beta|^2 \cos(\delta_a - \delta_b) \end{aligned}} \tag{23.93}$$

Here the intensity correlation is seen to depend on the phase difference $\delta_a - \delta_b$, which is determined by the wavelength and by the geometry of the source-detector configuration (Figure 23.1). The result (23.93) forms the basis for photon counting and intensity interferometry and is generally applicable to boson systems.

Exercise 23.19. Check the correlation matrix (23.92).

A particularly important mixed state is the state of thermal equilibrium, which is a special case of a *chaotic state*. A chaotic state is defined as the state of minimum available information or maximum randomness and entropy, subject to the constraint that it must reproduce a given value $\langle N_i \rangle$ for the mean number of photons occupying the mode i. The theory of Section 22.5 may be applied to the calculation of the von Neumann entropy S for each mode, but now there is a Lagrangian multiplier λ_i for each mode. Since we are considering only a single mode here, we temporarily omit all indices identifying the mode. Following the same arguments as before, we maximize

$$S - k\lambda\langle N \rangle = -k \operatorname{trace}(\rho \ln \rho) - k\lambda\langle N \rangle \tag{23.94}$$

and obtain for the chaotic state the density operator

$$\rho = \frac{e^{-\lambda a^\dagger a}}{Z} \tag{23.95}$$

with the partition function,

$$Z = \operatorname{trace} e^{-\lambda a^\dagger a} = \sum_n \langle n | e^{-\lambda a^\dagger a} | n \rangle = \frac{1}{1 - e^{-\lambda}} \tag{23.96}$$

and the bosonic constraint

$$\boxed{\langle N\rangle = \frac{1}{e^{\lambda} - 1}} \tag{23.97}$$

or

$$e^{-\lambda} = \frac{\langle N\rangle}{\langle N\rangle + 1} \tag{23.98}$$

Since $\langle N\rangle \geq 0$, relation (23.98) requires that $\lambda \geq 0$.

Exercise 23.20. Prove that the partition function (23.96) for photons is

$$Z = \langle N\rangle + 1 \tag{23.99}$$

Evaluate the entropy as a function of $\langle N\rangle$.

The density operator for a chaotic photon state may thus be expressed as

$$\rho = \frac{\langle N\rangle^N}{(\langle N\rangle + 1)^{N+1}} \tag{23.100}$$

Exercise 23.21. Show that

$$\rho = \frac{e^{-\lambda}}{\langle N\rangle} \sum_{n=0}^{\infty} e^{-\lambda n} |n\rangle\langle n| = \frac{1}{\pi\langle N\rangle} \int e^{-|\alpha|^2/\langle N\rangle} |\alpha\rangle \, d^2\alpha \, \langle\alpha| \tag{23.101}$$

are equivalent representations for the chaotic-state density operator in terms of photon-number and coherent states, respectively.

Exercise 23.22. Prove that any pure or mixed *single-mode* state, represented by a density operator ρ, has first-order coherence at all pairs of spacetime points.

If a radiation field is in thermal equilibrium, as in a black-body cavity, all field modes (now labeled by the index i) have the same temperature, which is determined by the average *total energy* of the system.

$$\langle H\rangle = \sum_i \hbar\omega_i \langle a_i^\dagger a_i\rangle$$

On thermodynamic grounds, we argue as in Section 22.5 that in thermal equilibrium at temperature T the Langrangian multiplier for each mode must be identified with

$$\lambda_i = \frac{\hbar\omega_i}{kT} \tag{23.102}$$

Substituting this value in (23.97), we find that the average number of photons in the mode $(\mathbf{k}, \lambda)$ in thermal equilibrium is

$$\boxed{\langle N_{\mathbf{k}}\rangle = \frac{1}{e^{\hbar\omega_{\mathbf{k}}/kT} - 1}} \tag{23.103}$$

This is Planck's distribution law for *black-body radiation*, from which the quantum theory took its start.

Problems

1. Consider a system of photons with fixed propagation vector in the positive z direction. The one-photon basis states are linear polarization states in x and y directions, $a_x^\dagger|\mathbf{0}\rangle$ and $a_y^\dagger|\mathbf{0}\rangle$.

(a) If $a_L^\dagger$ and $a_R^\dagger$ are the creation operators for positive and negative helicity photons, respectively, express the photon spin operator

$$S_z = \hbar(a_L^\dagger a_L - a_R^\dagger a_R)$$

in terms of the creation and annihilation operators for linear polarization.

(b) Work out the commutation relations of S_z with $a_x^\dagger$ and $a_y^\dagger$, and compare them with the commutation relations for the components of a vector operator. Show that S_z commutes with the number-of-photons operator.

(c) Construct the general two-photon state for this system. Derive the two-photon eigenstates of S_z, and interpret them in terms of the helicity of the photons.

(d) For an ensemble of unpolarized two-photon states, obtain the density matrix. How would the density matrix look if the two photons were distinguishable?

2. Prove that in a hydrogen atom the radiative transition from the $2S$ excited to the $1S$ ground state cannot occur by emission of one photon. Outline (but do not attempt to carry through in detail) the calculation of the transition rate for two-photon emission. In the dipole approximation, show that the two photons are preferentially emitted in the same direction or in opposite directions and that the angular correlation function is proportional to $1 + \cos^2\theta$, if θ is the angle between the photon momenta and if the polarization of the emitted light is not observed. Estimate the order of magnitude of the lifetime of the metastable $2S$ state.

3. Evaluate the peak value of the cross section for electric dipole absorption by a linear harmonic oscillator in its ground state, assuming that the excited state is depleted only by spontaneous emission. Use Eq. (19.134).

4. Compute the total rate of photon emission from the initial $n = 3$ level of a hydrogen atom to the $n = 2$ level (Balmer alpha line) and the mean lifetime of the atom, assuming that initially the substates of the $n = 3$ level are equally populated.

CHAPTER 24

Relativistic Electron Theory

In this final chapter, we develop a condensed account of the relativistic theory of electrons, or other fermions with spin one-half and mass m. In the interaction with the electromagnetic field, positrons of charge $e > 0$ are created simultaneously with electrons of charge $-e$. Turning history upside down, we begin with a consistent unified description of both particles and antiparticles in terms of a common *electron-positron field*, based on the free-particle Dirac equation. We extend the global symmetries of Chapter 17 to the relativistic theory and consider the discrete symmetries (spatial reflection, time reversal, and charge conjugation) and their interconnection. In a one-particle approximation to the quantum field theory, the Dirac equation for a (four-component) spinor wave function in an external electromagnetic field is seen to fit into the standard scheme of ordinary quantum mechanics. In the nonrelativistic limit, the Dirac equation reduces to the Schrödinger equation. The story ends with the relativistic theory of the fine structure of the spectrum of the hydrogenic atom.

1. ***The Electron-Positron Field.*** A free relativistic electron or positron is characterized by its linear momentum $\mathbf{p}$ and energy E_p, which are related by

$$E_p = \sqrt{c^2p^2 + (mc^2)^2} \tag{24.1}$$

As in the case of the photon, only a measurement of the component of the particle's spin in the direction of the momentum is compatible with a sharp energy-momentum vector. Hence, the electron or positron may have definite positive (R) or negative (L) *helicity*.[1]

We introduce creation and annihilation operators for electrons ($a^\dagger$ and a) and positrons ($b^\dagger$ and b) in the two helicity states, subject to the *anticommutation* relations:

$$\begin{aligned} a_R(\mathbf{p})a_R^\dagger(\mathbf{p}') + a_R^\dagger(\mathbf{p}')a_R(\mathbf{p}) &= a_L(\mathbf{p})a_L^\dagger(\mathbf{p}') + a_L^\dagger(\mathbf{p}')a_L(\mathbf{p}) \\ &= b_R(\mathbf{p})b_R^\dagger(\mathbf{p}') + b_R^\dagger(\mathbf{p}')b_R(\mathbf{p}) \\ &= b_L(\mathbf{p})b_L^\dagger(\mathbf{p}') + b_L^\dagger(\mathbf{p}')b_L(\mathbf{p}) = \delta(\mathbf{p} - \mathbf{p}') \end{aligned} \tag{24.2}$$

All other anticommutators of these eight operators are set equal to zero, partly as a consequence of the fermion theory developed in Chapter 21, and partly (namely, for anticommutators of a or $a^\dagger$ with b or $b^\dagger$) as an assumption that will be seen to be consistent with the formulation of a unified electron-positron theory.[2]

[1] *Warning:* In optics it is conventional to define positive (negative) helicity as left (right) circular polarization of light. See Exercise 23.8.

[2] In Chapter 23, the photon momentum was restricted to discrete values by the imposition of periodic boundary conditions on the radiation field. For the electron-positron field, it is convenient to let $L \to \infty$ from the beginning and allow all of momentum space for $\mathbf{p}$.

The operators for the energy, linear momentum, and charge of a system of free electrons and positrons are easily written down:

$$\mathcal{H} = \int E_p[a_R^\dagger(\mathbf{p})a_R(\mathbf{p}) + a_L^\dagger(\mathbf{p})a_L(\mathbf{p}) + b_R^\dagger(\mathbf{p})b_R(\mathbf{p}) + b_L^\dagger(\mathbf{p})b_L(\mathbf{p})]\, d^3p \tag{24.3}$$

$$\mathbf{P} = \int \mathbf{p}[a_R^\dagger(\mathbf{p})a_R(\mathbf{p}) + a_L^\dagger(\mathbf{p})a_L(\mathbf{p}) + b_R^\dagger(\mathbf{p})b_R(\mathbf{p}) + b_L^\dagger(\mathbf{p})b_L(\mathbf{p})]\, d^3p \tag{24.4}$$

$$Q = e \int [-a_R^\dagger(\mathbf{p})a_R(\mathbf{p}) - a_L^\dagger(\mathbf{p})a_L(\mathbf{p}) + b_R^\dagger(\mathbf{p})b_R(\mathbf{p}) + b_L^\dagger(\mathbf{p})b_L(\mathbf{p})]\, d^3p \tag{24.5}$$

Also, if the operator $\mathcal{J}$ represents the total angular momentum of the electron-positron system,

$$[\mathcal{J} \cdot \hat{\mathbf{p}}, a_R^\dagger(\mathbf{p})]a_R(\mathbf{p}) + [\mathcal{J} \cdot \hat{\mathbf{p}}, a_L^\dagger(\mathbf{p})]a_L(\mathbf{p}) + [\mathcal{J} \cdot \hat{\mathbf{p}}, b_R^\dagger(\mathbf{p})]b_R(\mathbf{p}) + [\mathcal{J} \cdot \hat{\mathbf{p}}, b_L^\dagger(\mathbf{p})]b_L(\mathbf{p})$$
$$= \frac{\hbar}{2}[a_R^\dagger(\mathbf{p})a_R(\mathbf{p}) - a_L^\dagger(\mathbf{p})a_L(\mathbf{p}) + b_R^\dagger(\mathbf{p})b_R(\mathbf{p}) - b_L^\dagger(\mathbf{p})b_L(\mathbf{p})] \tag{24.6}$$

is the spin component along the direction of the particle momentum $\mathbf{p}$ per unit volume in momentum space.

The objective of *local* quantum field theory is to seek ways of expressing these physical quantities as volume integrals of local (density) operators so that the operators for the total energy, momentum, charge, and other additive physical quantities eventually appear in the form

$$\mathcal{H} = \int \psi^\dagger(\mathbf{r})K\psi(\mathbf{r})\, d^3r \tag{24.7}$$

where K is an appropriate one-particle operator. The field operators $\psi(\mathbf{r})$ are again distinguished in the notation from their wave function relatives $\psi(\mathbf{r})$. As explained in Chapter 21, they are defined in the usual manner as Fourier integrals for a transformation of the creation operators from the momentum to the coordinate basis. However, care is required in the construction of the Fourier coefficients as well as in the choice of the one-particle operators representing physical quantities.

For example, it is formally possible to write the energy of the system of free electrons and positrons as

$$\mathcal{H} = \int \psi^\dagger(\mathbf{r})\sqrt{-\hbar^2c^2\nabla^2 + (mc^2)^2}\psi(\mathbf{r})\, d^3r$$

Such a choice was seen to be unsatisfactory in the case of photons because it implies a nonlocal expression for the energy density. For photons, this impasse led to the inference that a reasonable definition of a one-photon probability density in ordinary space cannot be given.

In the case of relativistic particles with mass, the same conclusion holds, although the expansion of $\sqrt{-\hbar^2c^2\nabla^2 + (mc^2)^2}$ in powers of ∇^2 shows that the nonlocal effects, which arise from the presence of arbitrarily high derivatives, disappear in the nonrelativistic approximation.[3] The goal of formulating a strictly *relativistic one-particle theory* is unattainable.

[3]The *Foldy-Wouthuysen* version of relativistic electron theory is based on the use of the square root operator for the energy, but it can be put in local form only by successive approximations. See Rose (1961), Schweber (1961), and Gross (1993), Section 5.7.

It is possible to produce a sensible field theory for particles with mass along similar lines as was done for photons, and this is customarily done for bosons such as pions (spin zero). Dirac's discovery of the relativistic theory for electrons (and positrons) showed that the field theory for fermions with spin one-half may be developed in a form that is strongly reminiscent of one-particle quantum mechanics. A straightforward relativistic one-particle *approximation* thus becomes feasible for such particles, and we will develop it in Section 24.6.

In the language of quantum field theory, the essence of Dirac's discovery is the observation that the physical quantities (24.3)–(24.5) may be reexpressed in alternate form by the use of the anticommutation relations and some simple changes of variables of integration, resulting in

$$\mathcal{H} = \int E_p[a_R^\dagger(\mathbf{p})a_R(\mathbf{p}) + a_L^\dagger(\mathbf{p})a_L(\mathbf{p}) - b_R(-\mathbf{p})b_R^\dagger(-\mathbf{p}) - b_L(-\mathbf{p})b_L^\dagger(-\mathbf{p}) + C]\, d^3p \tag{24.8}$$

$$\mathbf{P} = \int \mathbf{p}[a_R^\dagger(\mathbf{p})a_R(\mathbf{p}) + a_L^\dagger(\mathbf{p})a_L(\mathbf{p}) + b_R(-\mathbf{p})b_R^\dagger(-\mathbf{p}) + b_L(-\mathbf{p})b_L^\dagger(-\mathbf{p}) + C]\, d^3p \tag{24.9}$$

$$Q = e\int [-a_R^\dagger(\mathbf{p})a_R(\mathbf{p}) - a_L^\dagger(\mathbf{p})a_L(\mathbf{p}) - b_R(-\mathbf{p})b_R^\dagger(-\mathbf{p}) - b_L(-\mathbf{p})b_L^\dagger(-\mathbf{p}) + C]\, d^3p \tag{24.10}$$

and

$$[\mathcal{J}\cdot\hat{\mathbf{p}}, a_R^\dagger(\mathbf{p})]a_R(\mathbf{p}) + [\mathcal{J}\cdot\hat{\mathbf{p}}, a_L^\dagger(\mathbf{p})]a_L(\mathbf{p}) - [\mathcal{J}\cdot\hat{\mathbf{p}}, b_R^\dagger(-\mathbf{p})]b_R(-\mathbf{p}) - [\mathcal{J}\cdot\hat{\mathbf{p}}, b_L^\dagger(-\mathbf{p})]b_L(-\mathbf{p}) = \frac{\hbar}{2}[a_R^\dagger(\mathbf{p})a_R(\mathbf{p}) - a_L^\dagger(\mathbf{p})a_L(\mathbf{p}) - b_R(-\mathbf{p})b_R^\dagger(-\mathbf{p}) + b_L(-\mathbf{p})b_L^\dagger(-\mathbf{p})] \tag{24.11}$$

If we momentarily disregard the constant terms symbolized by C, these expressions show that the *annihilation* operator for a *positron*, $b_L(-\mathbf{p})$, can also be interpreted as an operator *creating an electron* of momentum $\mathbf{p}$ but *positive* helicity and *negative* energy, $-E_p$. Such negative energies appear quite naturally in a relativistic theory that relates energy and momentum by the equation

$$E_p^2 = c^2p^2 + (mc^2)^2 \tag{24.12}$$

allowing in addition to Eq. (24.1) the solution $-\sqrt{c^2p^2 + (mc^2)^2}$.

Using these clues, we construct a field operator

$$\psi(\mathbf{r}) = \psi^{(+)}(\mathbf{r}) + \psi^{(-)}(\mathbf{r}) \tag{24.13}$$

as the sum of positive and negative frequency (energy) parts defined as

$$\psi^{(+)}(\mathbf{r}) = \frac{1}{(2\pi\hbar)^{3/2}}\int [u^{(R)}(\mathbf{p})a_R(\mathbf{p}) + u^{(L)}(\mathbf{p})a_L(\mathbf{p})]e^{(i/\hbar)\mathbf{p}\cdot\mathbf{r}}\, d^3p \tag{24.14}$$

$$\begin{aligned}\psi^{(-)}(\mathbf{r}) &= \frac{1}{(2\pi\hbar)^{3/2}}\int [v^{(R)}(-\mathbf{p})b_R^\dagger(\mathbf{p}) + v^{(L)}(-\mathbf{p})b_L^\dagger(\mathbf{p})]e^{-(i/\hbar)\mathbf{p}\cdot\mathbf{r}}\, d^3p \\ &= \frac{1}{(2\pi\hbar)^{3/2}}\int [v^{(R)}(\mathbf{p})b_R^\dagger(-\mathbf{p}) + v^{(L)}(\mathbf{p})b_L^\dagger(-\mathbf{p})]e^{(i/\hbar)\mathbf{p}\cdot\mathbf{r}}\, d^3p\end{aligned} \tag{24.15}$$

The coefficients $u^{(R)}(\mathbf{p})$, $u^{(L)}(\mathbf{p})$, $v^{(R)}(\mathbf{p})$, and $v^{(L)}(\mathbf{p})$ are one-column matrices that must be orthogonal to each other, such that for a fixed momentum $\mathbf{p}$,

$$u^{(R)\dagger}u^{(L)} = (u_1^{(R)*}\ u_2^{(R)*}\ \ldots)\begin{pmatrix} u_1^{(L)} \\ u_2^{(L)} \\ \vdots \\ \vdots \end{pmatrix} = 0$$

and similarly,

$$u^{(R)\dagger}u^{(L)} = u^{(R)\dagger}v^{(R)} = u^{(R)\dagger}v^{(L)} = u^{(L)\dagger}v^{(R)} = u^{(L)\dagger}v^{(L)} = v^{(R)\dagger}v^{(L)} = 0 \quad (24.16)$$

Generalizing the terminology introduced in Section 16.3, we call these one-column matrices, with an as yet unspecified number of rows, *spinors*, and *Dirac spinors* on occasion when it is essential to avoid confusion with the two-component matrices of Chapter 16. We assume these spinors to be normalized according to the relations[4]

$$u^{(R)\dagger}u^{(R)} = u^{(L)\dagger}u^{(L)} = v^{(R)\dagger}v^{(R)} = v^{(L)\dagger}v^{(L)} = 1 \quad (24.17)$$

The field operators $\psi(\mathbf{r})$ and $\psi^\dagger(\mathbf{r})$ are similarly spinors, with components carrying a yet to be determined number of spin indices. If such spinors can be found, the total linear momentum and the total charge of the system can be written as

$$\mathbf{P} = \int \psi^\dagger(\mathbf{r})\frac{\hbar}{i}\nabla\psi(\mathbf{r})\,d^3r \quad (24.18)$$

$$Q = -\frac{e}{2}\int [\psi^\dagger(\mathbf{r})\psi(\mathbf{r}) - \tilde{\psi}(\mathbf{r})\tilde{\psi}^\dagger(\mathbf{r})]\,d^3r \quad (24.19)$$

where the symbol $\sim$ indicates matrix transposition.

Three comments are in order: (1) The formula (24.18) represents the expression (24.9), but we omit the constant term in the integrand, which merely ensures that the vacuum has zero momentum and is not needed if all momenta are measured relative to the vacuum. (2) Equation (24.18) for the linear momentum has the same form in the relativistic as in the nonrelativistic theory because, as indicated by Eq. (2.28), $(\hbar/i)\,\nabla$ represents the three spatial components of a relativistic four-vector. Angular momentum is made relativistic in a similar straightforward manner (Section 24.3). (3) The charge operator (24.5) was constructed to have zero expectation value in the vacuum, defined as the state in which there are no electrons and no positrons with positive energy. The peculiar form of Eq. (24.19) arises from rewriting (24.5) more symmetrically as

$$\begin{aligned} Q = -\frac{e}{2}\int [&a_R^\dagger(\mathbf{p})a_R(\mathbf{p}) - a_R(\mathbf{p})a_R^\dagger(\mathbf{p}) + a_L^\dagger(\mathbf{p})a_L(\mathbf{p}) - a_L(\mathbf{p})a_L^\dagger(\mathbf{p}) \\ &- b_R^\dagger(\mathbf{p})b_R(\mathbf{p}) + b_R(\mathbf{p})b_R^\dagger(\mathbf{p}) - b_L^\dagger(\mathbf{p})b_L(\mathbf{p}) + b_L(\mathbf{p})b_L^\dagger(\mathbf{p})]\,d^3p \end{aligned}$$

If we start from (24.10), the charge operator can be expressed in terms of the field as

$$Q = -e\int \psi^\dagger(\mathbf{r})\psi(\mathbf{r})\,d^3r + e\int \langle \mathbf{0}|\psi^\dagger(\mathbf{r})\psi(\mathbf{r})|\mathbf{0}\rangle\,d^3r \quad (24.20)$$

[4] *Warning:* A variety of different normalizations for Dirac spinors are current in the literature. The main reason for making a different choice is that one often prefers a Lorentz-invariant normalization, which (24.17) is not.

In quantum field theory, it is advantageous to define a *normal-ordered product* or *Wick product* : . . . : of the two fermion field operators $\psi(\mathbf{r})$ and $\psi^\dagger(\mathbf{r}')$ such that it has vanishing vacuum expectation value:[5]

$$-:\psi(\mathbf{r})\psi^\dagger(\mathbf{r}'): = :\psi^\dagger(\mathbf{r}')\psi(\mathbf{r}): = \psi^\dagger(\mathbf{r}')\psi(\mathbf{r}) - \langle \mathbf{0}|\psi^\dagger(\mathbf{r}')\psi(\mathbf{r})|\mathbf{0}\rangle \tag{24.21}$$

where the minus sign is due to the anticommutation relations. With this notation, we write the total charge operator simply as

$$Q = -e\int :\psi^\dagger(\mathbf{r})\psi(\mathbf{r}): d^3r \tag{24.22}$$

The equivalence of (24.19) and (24.20) can be directly established after the anticommutation relations for the field are obtained in the next section. The compact formula (24.22) shows how close we can come to our stated objective of expressing additive physical quantities as integrals over local densities. Except for the appearance of the normal-ordered product, which is an essential feature of relativistic theory, (24.22) indeed looks like (24.7).

Exercise 24.1. Show that the vacuum expectation value $\langle \mathbf{0}|\psi^\dagger(\mathbf{r})\psi(\mathbf{r})|\mathbf{0}\rangle$ is infinite.

2. ***The Dirac Equation.*** It remains to show that the energy of the system can also be written in the form

$$\mathcal{H} = \int \psi^\dagger(\mathbf{r})\, H\, \psi(\mathbf{r})\, d^3r + \text{const.} \tag{24.23}$$

Substitution of the fields (24.14) and (24.15) in this integral shows that this goal can be accomplished if we require that

$$Hu^{(R,L)}(\mathbf{p})e^{(i/\hbar)\mathbf{p}\cdot\mathbf{r}} = E_p u^{(R,L)}(\mathbf{p})e^{(i/\hbar)\mathbf{p}\cdot\mathbf{r}} \tag{24.24}$$

$$Hv^{(R,L)}(\mathbf{p})e^{(i/\hbar)\mathbf{p}\cdot\mathbf{r}} = -E_p v^{(R,L)}(\mathbf{p})e^{(i/\hbar)\mathbf{p}\cdot\mathbf{r}}\,. \tag{24.25}$$

If (24.23) is to be an integral over a localized energy density, the requirements of Lorentz invariance make it mandatory to seek a Hamiltonian that is *linear* in the differential operator ∇. Therefore, we attempt to construct H in the form

$$\boxed{H = c\boldsymbol{\alpha}\cdot\frac{\hbar}{i}\nabla + \beta mc^2} \tag{24.26}$$

leaving the constant square *matrices* α_x, α_y, α_z, and β as yet undetermined. With this choice for H, Eqs. (24.24) and (24.25) reduce to

$$(c\boldsymbol{\alpha}\cdot\mathbf{p} + \boldsymbol{\beta}mc^2)u^{(R,L)}(\mathbf{p}) = E_p u^{(R,L)}(\mathbf{p}) \tag{24.27}$$

$$(c\boldsymbol{\alpha}\cdot\mathbf{p} + \beta mc^2)v^{(R,L)}(\mathbf{p}) = -E_p v^{(R,L)}(\mathbf{p}) \tag{24.28}$$

Since the eigenvalues $\pm E_p$ are real and the eigenspinors orthogonal, the operator-matrix

$$\boxed{H_\mathbf{p} = c\boldsymbol{\alpha}\cdot\mathbf{p} + \beta mc^2} \tag{24.29}$$

[5]See Mandl and Shaw (1984).

must be Hermitian. Thus, $\boldsymbol{\alpha}$ and β are four Hermitian matrices. They must be *at least* four dimensional (four rows and four columns) if $H_{\mathbf{p}}$ is to have four orthogonal eigenspinors, and they should be *no more* than four dimensional if the description of electrons and positrons in terms of momentum, energy, and helicity is complete.

Since the eigenvalues of $H_{\mathbf{p}}$ are to be E_p and $-E_p$, with each of these being doubly degenerate, all four eigenvalues of $(H_{\mathbf{p}})^2$ must be equal to E_p^2, hence $(H_{\mathbf{p}})^2 = E_p^2\, 1$ and Trace $H_{\mathbf{p}} = 0$ is required. If we take the square of (24.26) and use the relation (24.12), we thus obtain the conditions

$$\alpha_x^2 = \alpha_y^2 = \alpha_z^2 = \beta^2 = 1 \tag{24.30}$$

$$\begin{aligned} \alpha_x\alpha_y + \alpha_y\alpha_z = \alpha_y\alpha_z + \alpha_z\alpha_y = \alpha_z\alpha_x + \alpha_x\alpha_x \\ = \alpha_x\beta + \beta\alpha_x = \alpha_y\beta + \beta\alpha_y = \alpha_z\beta + \beta\alpha_z = 0 \end{aligned} \tag{24.31}$$

Our problem thus reduces to a purely algebraic one of finding *four-dimensional* Hermitian matrices with the properties (24.30) and (24.31). Pauli proved that all matrix solutions to these equations for $\boldsymbol{\alpha}$ and β are reducible by unitary transformation to one another. Hence, it is sufficient to determine one particular 4×4 solution and show that all traces vanish.

Exercise 24.2. Using only the conditions (24.30) and (24.31), prove that the trace of α_x, α_y, and α_z, and β vanishes, and show that each of these matrices has n eigenvalues $+1$ and n eigenvalues -1, where $2n$ is the dimension of the matrices.

Exercise 24.3. From (24.30) and (24.31), prove that $\boldsymbol{\alpha}$ and β are at least four dimensional.

The most widely used representation of the $\boldsymbol{\alpha}$ and β matrices are the 4×4 matrices specified in terms of the 2×2 Pauli matrices of Section 16.4:

$$\boxed{\alpha = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}, \qquad \beta = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}} \tag{24.32}$$

Every element in these 2×2 matrices is itself to be understood as a 2×2 matrix, so the matrices $\boldsymbol{\alpha}$ and β are 4×4. We refer to (24.32) as the *standard representation.*

Exercise 24.4. Verify the validity of the solutions (24.32) to the problem posed by conditions (24.30) and (24.31).

The discussion of this section so far leaves unidentified the Hermitian matrix that represents the helicity. Such a matrix must commute with $H_{\mathbf{p}}$ and distinguish, by its eigenvalues, the two helicity states R and L. It will be readily identified after the angular momentum operator is obtained (Sections 24.3 and 24.4).

The anticommutation relations for the field operators can now be derived from Eq. (24.2) and the remarks following this equation. The four eigenspinors $u^{(R)}(\mathbf{p})$, $u^{(L)}(\mathbf{p})$, $v^{(R)}(\mathbf{p})$, and $v^{(L)}(\mathbf{p})$ of the 4×4 matrix $H_{\mathbf{p}}$ are orthonormal. Hence, they form a complete set of spinors, and the *closure* relation

$$u^{(R)}u^{(R)\dagger} + u^{(L)}u^{(L)\dagger} + v^{(R)}v^{(R)\dagger} + v^{(L)}v^{(L)\dagger} = 1 \tag{24.33}$$

holds. Using this relation, we can easily verify that

$$\boxed{\begin{aligned}\psi_\alpha(\mathbf{r})\psi_\beta(\mathbf{r}') + \psi_\beta(\mathbf{r}')\psi_\alpha(\mathbf{r}) &= 0\\ \psi_\alpha^\dagger(\mathbf{r})\psi_\beta^\dagger(\mathbf{r}') + \psi_\beta^\dagger(\mathbf{r}')\psi_\alpha^\dagger(\mathbf{r}) &= 0\\ \psi_\alpha(\mathbf{r})\psi_\beta^\dagger(\mathbf{r}') + \psi_\beta^\dagger(\mathbf{r}')\psi_\alpha(\mathbf{r}) &= \delta(\mathbf{r}-\mathbf{r}')\delta_{\alpha\beta}\end{aligned}} \tag{24.34}$$

Exercise 24.5. Verify Eqs. (24.34).

Exercise 24.6. Using (24.18) and (24.34), prove that

$$[\psi(\mathbf{r}), \mathbf{P}] = \frac{\hbar}{i}\nabla\psi(\mathbf{r}) \tag{24.35}$$

which is the spatial companion of the Heisenberg equation of motion for the field operator.

From the equations of motion for the creation and annihilation operators, the time development of the *free* Dirac field is deduced by use of the Hamiltonian (24.3). We obtain in the *Heisenberg picture*

$$\psi^{(+)}(\mathbf{r}, t) = \frac{1}{(2\pi\hbar)^{3/2}}\int [u^{(R)}(\mathbf{p})a_R(\mathbf{p}) + u^{(L)}(\mathbf{p})a_L(\mathbf{p})]e^{(i/\hbar)(\mathbf{p}\cdot\mathbf{r}-E_p t)}\,d^3p \tag{24.36}$$

$$\psi^{(-)}(\mathbf{r}, t) = \frac{1}{(2\pi\hbar)^{3/2}}\int [v^{(R)}(\mathbf{p})b_R^\dagger(-\mathbf{p}) + v^{(L)}(\mathbf{p})b_L^\dagger(-\mathbf{p})]e^{(i/\hbar)(\mathbf{p}\cdot\mathbf{r}+E_p t)}\,d^3p \tag{24.37}$$

If Eqs. (24.24), (24.25), and (24.26) are applied, we see that both frequency components of ψ and the total field itself satisfy the *field equation*

$$\boxed{i\hbar\frac{\partial\psi(\mathbf{r}, t)}{\partial t} = c\left(\frac{\hbar}{i}\boldsymbol{\alpha}\cdot\nabla + \beta mc\right)\psi(\mathbf{r}, t)} \tag{24.38}$$

This equation, which is the analogue of the time-dependent Schrödinger equation of nonrelativistic quantum mechanics and of Maxwell's equations for the electromagnetic field, is known as the *Dirac equation of the electron*. The Dirac equation (24.38) can be cast in a more appealing form, particularly suitable for discussion of Lorentz covariance, by the introduction of a new set of 4×4 matrices, known as the *Dirac γ matrices*:

$$\gamma^0 = \beta, \qquad \gamma^1 = \beta\alpha_x \qquad \gamma^2 = \beta\alpha_y, \qquad \gamma^3 = \beta\alpha_z \tag{24.39}$$

Exercise 24.7. Show that the three "spatial" matrices, γ^1, γ^2, γ^3, are antihermitian. In a transparent notation: $\boldsymbol{\gamma}^\dagger = -\boldsymbol{\gamma}$.

Using relativistic notation, the metric introduced in Section 23.2 and the summation convention, with Greek indices running from 0 to 3, we may rewrite Eq. (24.38) in the compact form

$$\boxed{\gamma^\mu\frac{\partial\psi}{\partial x^\mu} + i\kappa\psi = 0} \tag{24.40}$$

We have abbreviated the inverse of the Compton wavelength of the electron as

$$\kappa = mc/\hbar$$

and denoted

$$\gamma^\mu \frac{\partial}{\partial x^\mu} = \gamma^0 \frac{\partial}{\partial x^0} + \boldsymbol{\gamma} \cdot \boldsymbol{\nabla} = \beta \frac{1}{c} \frac{\partial}{\partial t} + \beta \boldsymbol{\alpha} \cdot \boldsymbol{\nabla}$$

The conditions (24.30) and (24.31) may be summarized as anticommutation relations for the γ matrices:

$$\boxed{\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}\, 1} \tag{24.41}$$

The one-particle differential operator that represents energy-momentum is given by

$$p^\mu = \left(\frac{E}{c}, \mathbf{p}\right) = i\hbar \frac{\partial}{\partial x_\mu} = \left(\frac{i\hbar}{c}\frac{\partial}{\partial t}, \frac{\hbar}{i}\boldsymbol{\nabla}\right)$$

$$p_\mu = \left(\frac{E}{c}, -\mathbf{p}\right) = i\hbar \frac{\partial}{\partial x^\mu} = \left(\frac{i\hbar}{c}\frac{\partial}{\partial t}, -\frac{\hbar}{i}\boldsymbol{\nabla}\right)$$

For electrons with charge $q = -e$ ($e > 0$), the presence of an external electromagnetic field, acting on the matter field, is as usual taken into account by the replacement

$$p^\mu \to p^\mu + \frac{e}{c} A^\mu$$

This gauge principle defines a *minimal interaction* of the Dirac spinor field with the vector field

$$A^\mu = (\phi, \mathbf{A}) \quad \text{and} \quad A_\mu = (\phi, -\mathbf{A})$$

The substitution changes the Dirac equation from its free field form into

$$\boxed{\gamma^\mu \left(\frac{\partial}{\partial x^\mu} - \frac{ie}{\hbar c} A_\mu\right)\psi + i\kappa\psi = 0} \tag{24.42}$$

or in the noncovariant form, analogous to Eq. (24.38),

$$\boxed{i\hbar \frac{\partial \psi(\mathbf{r}, t)}{\partial t} = \left[c\boldsymbol{\alpha} \cdot \left(\frac{\hbar}{i}\boldsymbol{\nabla} + \frac{e}{c}\mathbf{A}\right) - e\phi + \beta mc^2\right]\psi(\mathbf{r}, t)} \tag{24.43}$$

It is useful to define an *adjoint* Dirac field operator by the relation

$$\overline{\psi} = \psi^\dagger \gamma^0$$

Since $\boldsymbol{\gamma}$ is antihermitian and γ^0 Hermitian, Hermitian conjugation of Eq. (24.42) and multiplication on the right by γ^0 leads to

$$\boxed{\left(\frac{\partial}{\partial x^\mu} + \frac{ie}{\hbar c} A_\mu\right)\overline{\psi}\gamma^\mu - i\kappa\overline{\psi} = 0} \tag{24.44}$$

If this equation is multiplied on the right by ψ and Eq. (24.42) on the left by $\overline{\psi}$, and if the resulting equations are added to one another, the *continuity equation*

$$\boxed{\frac{\partial}{\partial x^\mu}(\overline{\psi}\gamma^\mu\psi) = 0} \tag{24.45}$$

is obtained. Similarly, it is easy to prove the further continuity equation

$$\frac{\partial}{\partial x^\mu}(\tilde{\psi}\tilde{\gamma}^\mu\tilde{\bar{\psi}}) = 0 \tag{24.46}$$

Comparing these expressions with the total charge operator (24.19) or (24.22), we infer that the *electric current density* four-vector of the electron-positron system is defined by

$$j^\mu = (c\rho, \mathbf{j}) = -\frac{ec}{2}(\bar{\psi}\gamma^\mu\psi - \tilde{\psi}\tilde{\gamma}^\mu\tilde{\bar{\psi}}) = -ec{:}\bar{\psi}\gamma^\mu\psi{:} \tag{24.47}$$

This operator is often simply referred to as the *four-current*. Conservation of charge is ensured by the continuity Eqs. (24.45) and (24.46), or

$$\frac{\partial j^\mu}{\partial x^\mu} = 0 \tag{24.48}$$

Unlike the nonrelativistic current density, which explicitly depends on the vector potential **A** (see Exercise 4.17), **A** does not appear in the definition (24.47). Of course, it affects the current indirectly, since the field operator is a solution of the Dirac equation (24.42), which includes the electromagnetic potential. We will relate the relativistic formulation to the nonrelativistic limit in Section 24.8.

Exercise 24.8. Derive the continuity equations. Show that the current is a Hermitian operator, and, using the anticommutation relations (24.34), verify the equality of the two expressions for the conserved current in (24.47).

3. ***Relativistic Invariance.*** Unlike the relativistic invariance of Maxwell's equations for the free radiation field, even in quantized form, which needs no proof since the Lorentz transformations were designed to accomplish just this purpose, it is necessary to demonstrate that the Dirac theory is in consonance with the demands of special relativity. Specifically, the requirement of invariance of the theory under inhomogeneous Lorentz (or Poincaré) transformations will serve as a guide in establishing the transformation properties of the electron-positron field. The general theory of the irreducible representations of the *Lorentz group* contains all the relevant information, but if nothing more than the transformation properties of a special field is desired, the mathematical structure may be deduced from simple physical considerations.

Einstein's restricted principle of special relativity postulates the equivalence of physical systems that are obtained from each other by geometrical translation or rotation or that differ from one another only by being in uniform relative motion. According to Section 17.1, such equivalent systems can be connected by a unitary transformation of the respective state vectors.

The principle of relativity is implemented by constructing the coordinate transformation

$$x'^\mu = a^\mu{}_\nu x^\nu + b^\mu \tag{24.49}$$

with real coefficients $a^\mu{}_\nu$ and b^μ, subject to the orthogonality condition

$$dx'^\mu dx'_\mu = dx^\nu dx_\nu \quad \text{or} \quad a^\mu{}_\lambda a_\mu{}^\nu = \delta^\nu_\lambda \tag{24.50}$$

In addition to the *proper orthochronous* Lorentz transformation for which

$$\det a^{\mu}{}_{\nu} = 1, \qquad a^{0}{}_{0} \geq 1$$

the orthogonality condition allows *improper* Lorentz transformations such as space reflections and time reversal, as well as combinations of these with proper orthochronous transformations. Although there is no compelling reason to expect that the coverage of the principle of relativity extends to the improper Lorentz transformations and those reversing the sense of time, it is important to investigate whether the proposed theory is invariant under the totality of the transformations licensed by the orthogonality condition (24.50).

It is a fundamental assertion of local quantum field theory that if an active Lorentz transformation takes the point $(\mathbf{r}, t)$ into $(\mathbf{r}', t')$ and changes the state Ψ into a state $U\Psi$, where U is unitary, the components of $\psi(\mathbf{r}', t')U\Psi$ must be related by a linear transformation to the components of $U\psi(\mathbf{r}, t)\Psi$. Hence, the field must transform as

$$\psi(\mathbf{r}', t')U\Psi = \mathsf{S}U\psi(\mathbf{r}, t)\Psi$$

or

$$U^{\dagger}\psi(\mathbf{r}', t')U = \mathsf{S}\psi(\mathbf{r}, t) \tag{24.51}$$

The 4×4 matrix S defines the geometrical transformation properties of the spinor whose components, like those of a vector or tensor, are reshuffled in this symmetry operation. [Compare (24.51) to Eq. (17.83).] It is assumed that the vacuum state is left unchanged by a symmetry transformation: $U\Psi^{(0)} = \Psi^{(0)}$.

We first consider three-dimensional rotations as a subgroup of the Lorentz transformations. From the definition of rotations, it follows that we must expect the relations

$$Ua_{R,L}(\mathbf{p}) = a_{R,L}(\mathbf{p}')U, \qquad Ub_{R,L}(\mathbf{p}) = b_{R,L}(\mathbf{p}')U \tag{24.52}$$

to hold, with $\mathbf{p}'$ being the momentum vector that is obtained from $\mathbf{p}$ by the rotation. Since $\mathbf{p} \cdot \mathbf{r} = \mathbf{p}' \cdot \mathbf{r}'$ and since the integral over the entire momentum space is invariant under rotations, it follows from Eqs. (24.13), (24.14), (24.15), and (24.52) that condition (24.51) will be satisfied if we determine the matrix S such that

$$u^{(R,L)}(\mathbf{p}') = \mathsf{S}u^{(R,L)}(\mathbf{p}) \tag{24.53}$$

$$v^{(R,L)}(\mathbf{p}') = \mathsf{S}v^{(R,L)}(\mathbf{p}) \tag{24.54}$$

Since E_p is invariant under rotations, the last two equations in conjunction with (24.27) and (24.28) imply the condition

$$H_{\mathbf{p}'}\mathsf{S} = \mathsf{S}H_{\mathbf{p}} \tag{24.55}$$

where

$$H_{\mathbf{p}} = c\boldsymbol{\alpha} \cdot \mathbf{p} + \beta mc^2 = \beta(c\boldsymbol{\gamma} \cdot \mathbf{p} + mc^2)$$

If we write

$$p'^{k} = a^{k}{}_{j}p^{j}, \qquad p_k = p'_{\ell}a^{\ell}{}_{k}$$

with summations over repeated Latin indices extending from 1 to 3 only, substitution into (24.55) produces the conditions

$$\beta\gamma^{\ell}\mathsf{S} = \mathsf{S}\beta\gamma^{k}a^{\ell}{}_{k} \qquad \beta\mathsf{S} = \mathsf{S}\beta$$

or

$$a^{\ell}{}_{k}\gamma^{k} = \mathsf{S}^{-1}\gamma^{\ell}\mathsf{S} \tag{24.56}$$

and

$$\gamma^{0} = \mathsf{S}^{-1}\gamma^{0}\mathsf{S} \tag{24.57}$$

The conditions (24.56) and (24.57) for the matrix S are included as special cases in the general condition that S must satisfy if the electron-positron field theory is to be invariant under all (homogeneous) Lorentz transformations:

$$\boxed{a^{\lambda}{}_{\mu}\gamma^{\mu} = \mathsf{S}^{-1}\gamma^{\lambda}\mathsf{S}} \tag{24.58}$$

Although this condition may be obtained by generalizing the argument that we have given for spatial rotations as active transformations, it is easier to derive it by taking the passive point of view and requiring that the Dirac equation (24.40) must be invariant under the transformation:

$$x'^{\mu} = a^{\mu}{}_{\nu}x^{\nu}, \qquad \frac{\partial}{\partial x^{\mu}} = \frac{\partial}{\partial x'^{\nu}}\,a^{\nu}{}_{\mu} \tag{24.59}$$

and

$$\psi'(\mathbf{r}', t') = \mathsf{S}\psi(\mathbf{r}, t) \tag{24.60}$$

Exercise 24.9. Derive condition (24.58) from the Lorentz invariance of the Dirac equation.

The demonstration of the Lorentz invariance of the theory will be complete if the matrix S can be exhibited for each possible Lorentz transformation. The explicit construction of S for proper orthochronous Lorentz transformations, which can be obtained continuously from the identity operation, is most easily accomplished by considering the condition (24.58) in an *infinitesimal* neighborhood of the identity. We may write

$$a^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \varepsilon^{\mu}{}_{\nu} \tag{24.61}$$

with the condition $\varepsilon^{\mu}{}_{\nu} = -\varepsilon_{\nu}{}^{\mu}$ as an immediate consequence of the orthogonality condition (24.50). For the case of spatial rotations, see Exercises 16.3 and 17.21.

If an arbitrary Lorentz transformation represented by S is followed by an infinitesimal one, the composite transformation is represented by $\mathsf{S} + d\mathsf{S}$, and the infinitesimal transformation is represented by

$$(\mathsf{S} + d\mathsf{S})\mathsf{S}^{-1} = 1 + d\mathsf{S}\cdot\mathsf{S}^{-1}$$

Applying Eq. (24.61) to Eq. (24.58), we get

$$(1 - d\mathsf{S}\cdot\mathsf{S}^{-1})\gamma^{\lambda}(1 + d\mathsf{S}\cdot\mathsf{S}^{-1}) = \gamma^{\lambda} + \varepsilon^{\lambda}{}_{\mu}\gamma^{\mu}$$

or

$$[\gamma^{\lambda}, d\mathsf{S}\cdot\mathsf{S}^{-1}] = \varepsilon^{\lambda}{}_{\mu}\gamma^{\mu} \quad \text{for all } \lambda \tag{24.62}$$

The solution of this commutation relation is seen to be

$$d\mathsf{S}\cdot\mathsf{S}^{-1} = \frac{1}{4}\,\varepsilon_{\mu\nu}\gamma^{\mu}\gamma^{\nu} \tag{24.63}$$

A three-dimensional rotation by an angle $\delta\phi$ about an axis along the unit vector $\hat{\mathbf{n}}$ takes the position vector $\mathbf{r}$, according to (11.10), into

$$\mathbf{r}' = \mathbf{r} + \delta\phi\, \hat{\mathbf{n}} \times \mathbf{r} \tag{24.64}$$

By comparison with Eqs. (24.59) and (24.61), the identification

$$\begin{aligned} \varepsilon^1{}_2 &= -\varepsilon^2{}_1 = -\delta\phi\, n_3 = -\varepsilon_{12} = \varepsilon_{21} \\ \varepsilon^2{}_3 &= -\varepsilon^3{}_2 = -\delta\phi\, n_1 = -\varepsilon_{23} = \varepsilon_{32} \\ \varepsilon^3{}_1 &= -\varepsilon^1{}_3 = -\delta\phi\, n_2 = -\varepsilon_{31} = \varepsilon_{13} \end{aligned} \tag{24.65}$$

emerges.

Exercise 24.10. Check one of the three equations (24.65).

If we substitute the infinitesimal displacements in (24.63) and define the matrix Σ,

$$\Sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu] \tag{24.66}$$

Eq. (24.63) reduces to

$$d\mathsf{S} \cdot \mathsf{S}^{-1} = -\frac{i}{2}\,\delta\phi(n_1\Sigma^{23} + n_2\Sigma^{31} + n_3\Sigma^{12})]$$

This matrix differential equation has the simple unitary solution

$$\mathsf{S} = \exp\left[-\frac{i}{2}\,\phi(n_1\Sigma^{23} + n_2\Sigma^{31} + n_3\Sigma^{12})\right] = e^{-i\phi\hat{\mathbf{n}}\cdot\Sigma/2} \tag{24.67}$$

We have used the notation

$$\Sigma_x = \Sigma^{23}, \qquad \Sigma_y = \Sigma^{31}, \qquad \Sigma_z = \Sigma^{12} \tag{24.68}$$

for the four-dimensional analogues of the *Pauli spin matrices*.

Exercise 24.11. Show that the 4×4 matrices Σ defined by (24.66) and (24.68) satisfy the usual commutation relations for Pauli spin matrices. Show that in the standard representation (24.32),

$$\Sigma = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} \tag{24.69}$$

If (24.64) is substituted into (24.51) and the integration over the rotation angle ϕ is performed, the behavior of the spinor field under finite rotations (with $t' = t$) is obtained:

$$\begin{aligned} U^\dagger\boldsymbol{\psi}(\mathbf{r}, t)U &= \exp\left[-i\phi\, \hat{\mathbf{n}} \cdot \left(\mathbf{r} \times \frac{1}{i}\nabla + \frac{1}{2}\boldsymbol{\Sigma}\right)\right]\boldsymbol{\psi}(\mathbf{r}, t) \\ &= \exp\left[-\frac{i}{\hbar}\,\phi\, \hat{\mathbf{n}} \cdot \left(\mathbf{L} + \frac{\hbar}{2}\boldsymbol{\Sigma}\right)\right]\boldsymbol{\psi}(\mathbf{r}, t) \end{aligned} \tag{24.70}$$

where $\mathbf{L} = \mathbf{r} \times (\hbar/i)\nabla$, as in the one-particle coordinate representation. If the unitary operator U is expressed as

$$U = \exp[-(i/\hbar)\phi\, \hat{\mathbf{n}} \cdot \mathcal{J}] \tag{24.71}$$

it follows from Eq. (24.70) that the Hermitian operator $\mathcal{J}$ must satisfy the commutation relations

$$[\psi(\mathbf{r}, t), \mathcal{J}] = \left(\mathbf{r} \times \frac{\hbar}{i}\nabla + \frac{\hbar}{2}\boldsymbol{\Sigma}\right)\psi(\mathbf{r}, t) \tag{24.72}$$

which are the rotational analogue of (24.35). The *total angular momentum operator*

$$\boxed{\mathcal{J} = \int \psi^\dagger(\mathbf{r})\left(\mathbf{r} \times \frac{\hbar}{i}\nabla + \frac{\hbar}{2}\boldsymbol{\Sigma}\right)\psi(\mathbf{r})\, d^3r} \tag{24.73}$$

satisfies this equation. Similar to the total angular momentum for photons (Section 23.2), the two terms on the right-hand side of (24.73) can be interpreted as orbital and spin angular momentum.

Exercise 24.12. Verify that Eq. (24.73) is consistent with Eq. (24.72) and with the defining relation for helicity, (24.6).

Since any proper orthochronous homogeneous Lorentz transformation may be obtained as a succession of spatial rotations and special Lorentz transformations, it suffices for the invariance proof to show the existence of S for special Lorentz transformations.

Exercise 24.13. For a special Lorentz transformation corresponding to uniform motion with velocity $v = c \tanh \chi$ along the x axis, show that

$$\mathsf{S} = \exp\left(\frac{i}{c}\chi\,\Sigma^{10}\right) = \exp\left(\frac{1}{2}\chi\,\alpha_x\right) \tag{24.74}$$

Note that since Σ^{10} is antihermitian, the matrix S is not unitary in this case. [The *unitary* operator U, which effects this transformation in accordance with Eq. (24.51), can again be constructed by starting from the infinitesimal transformation.]

Exercise 24.14. Discuss coordinate translations in the theory of the Dirac field.

From (24.67) and (24.74) it is easy to deduce that the matrix S for all Lorentz transformations has the property

$$\mathsf{S}^\dagger\gamma^0\mathsf{S} = \gamma^0 \tag{24.75}$$

Exercise 24.15. Verify Eq. (24.75). For three-rotations, reconcile this result with (24.57).

Combining (24.58) and (24.75), we obtain

$$\boxed{\mathsf{S}^\dagger\gamma^0\gamma^\mu\mathsf{S} = a^\mu{}_\nu\gamma^0\gamma^\nu} \tag{24.76}$$

If the unitary operator U, induced by a Lorentz transformation, is applied to the current density (24.47), use of (24.51) and (24.76) shows that the current density is a four-vector operator and satisfies the transformation equation

$$U^\dagger j^\mu(\mathbf{r}', t')U = a^\mu{}_\nu j^\nu(\mathbf{r}, t) \tag{24.77}$$

in generalization of the concept of a vector operator defined in Section 17.7.

The study of proper orthochronous Lorentz transformations must be supplemented by consideration of the fundamental *improper transformations*. Spatial reflections will be discussed in the remainder of this section. The study of time reversal is left to Section 24.5.

If *spatial reflection* of all three coordinates, or inversion, is a symmetry operation for the Dirac theory, condition (24.58) implies that

$$\mathsf{S}^{-1}\gamma^0\mathsf{S} = \gamma^0, \quad \text{and} \quad \mathsf{S}^{-1}\boldsymbol{\gamma}\mathsf{S} = -\boldsymbol{\gamma} \tag{24.78}$$

From these equations and (24.51) for $\mathbf{r}' = -\mathbf{r}$, it follows that the current density (24.47) behaves as a four-vector under the action of the unitary inversion operator U_P:

$$U_P^\dagger j^\mu(-\mathbf{r}, t)U_P = j_\mu(\mathbf{r}, t) \tag{24.79}$$

only if S is unitary, $\mathsf{S}^\dagger\mathsf{S} = 1$. Except for an arbitrary phase factor, all the conditions imposed if S is to represent an inversion are solved by

$$\mathsf{S} = \gamma^0 = \beta \tag{24.80}$$

and the inversion is thus accomplished by the relation

$$U_P\psi(\mathbf{r}, t)U_P^\dagger = \gamma^0\psi(-\mathbf{r}, t) \tag{24.81}$$

The unitary operator U_P defined by this equation is the *parity operator* for the Dirac field. It is conventional to assume that the vacuum state is an eigenstate of U_P with even parity.

Exercise 24.16. Show that the current is a (polar) vector under coordinate inversion only if the matrix S is unitary. Attempt an explicit construction of the parity operator in terms of the field operators.

It is convenient to define an additional Hermitian 4×4 Dirac matrix,

$$\gamma^5 \equiv \gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \tag{24.82}$$

which has the properties

$$\gamma^\mu\gamma^5 + \gamma^5\gamma^\mu = 0 \quad \text{and} \quad (\gamma^5)^2 = 1 \tag{24.83}$$

Exercise 24.17. Construct the matrix for γ^5 in the standard representation. Derive the eigenvalues and eigenspinors of γ^5. Prove Eqs. (24.83) as well as the property

$$[\gamma^5, \Sigma^{\mu\nu}] = 0 \tag{24.84}$$

Exercise 24.18. Verify the following transformation properties for the designated bilinear functions of the field operators under proper orthochronous Lorentz transformations and under reflections:

$\bar{\psi}(\mathbf{r}, t)\psi(\mathbf{r}, t)$	scalar	
$\bar{\psi}(\mathbf{r}, t)\gamma^\mu\psi(\mathbf{r}, t)$	vector	
$\bar{\psi}(\mathbf{r}, t)\Sigma^{\mu\nu}\psi(\mathbf{r}, t)$	antisymmetric tensor of rank two	(24.85)
$\bar{\psi}(\mathbf{r}, t)\gamma^5\gamma^\mu\psi(\mathbf{r}, t)$	axial (pseudo-)vector	
$\bar{\psi}(\mathbf{r}, t)\gamma^5\psi(\mathbf{r}, t)$	pseudoscalar	

The five kinds of bilinear field operators (24.85) provide the building blocks for constructing different interactions in relativistic quantum field theories. It can

be shown[6] that the set of matrices 1, γ^μ, $\Sigma^{\mu\nu}$, $\gamma^5\gamma^\mu$, γ^5, which glue $\overline{\psi}$ and ψ together in these bilinear products, is complete in the sense that any arbitrary 4×4 matrix can be expanded in terms of these 16. They form the basis of an algebra of 4×4 matrices.

4. ***Solutions of the Free Field Dirac Equation.*** In this section, we derive explicit solutions of the Eqs. (24.27) and (24.28) which combine to:

$$H_p u(\mathbf{p}) \equiv (c\boldsymbol{\alpha} \cdot \mathbf{p} + \beta mc^2)u(\mathbf{p}) = Eu(\mathbf{p}) \tag{24.86}$$

For this purpose, we will employ the standard representation (24.32) of the Dirac matrices.

Exercise 24.19. Write out the four linear homogeneous equations implied by (24.86) in full detail, and show that the vanishing of their determinant is assured by the condition $E^2 = E_p^2 = c^2p^2 + (mc^2)^2$. Prove that all 3×3 minors of the scalar determinant also vanish (but not all 2×2 minors), and interpret this result.

The simplest solutions are obtained if the momentum vector $\mathbf{p}$ points in the direction of the positive z axis. In this case, Eq. (24.27) reduces for $E = E_p$ to

$$\begin{pmatrix} mc^2 & 0 & cp & 0 \\ 0 & mc^2 & 0 & -cp \\ cp & 0 & -mc^2 & 0 \\ 0 & -cp & 0 & -mc^2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = E_p \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix}$$

or

$$\begin{aligned} mc^2u_1 + cpu_3 &= E_pu_1 \\ cpu_1 - mc^2u_3 &= E_pu_3 \\ mc^2u_2 - cpu_4 &= E_pu_2 \\ -cpu_2 - mc^2u_4 &= E_pu_4 \end{aligned} \tag{24.87}$$

Evidently, this system of equations possesses two linearly independent, and in fact orthogonal, solutions:

$$u^{(R)} \propto \begin{pmatrix} 1 \\ 0 \\ \dfrac{cp}{E_p + mc^2} \\ 0 \end{pmatrix} \quad \text{and} \quad u^{(L)} \propto \begin{pmatrix} 0 \\ 1 \\ 0 \\ \dfrac{-cp}{E_p + mc^2} \end{pmatrix} \tag{24.88}$$

The labels R and L have been affixed to these spinors because they are eigenspinors of the one-particle *helicity operator* $\boldsymbol{\Sigma} \cdot \hat{\mathbf{p}}$ (here reduced to Σ_z) with eigenvalues $+1$ and -1, respectively. According to the formula (24.73), this is the component of angular momentum along $\mathbf{p}$ for the particles of linear momentum $\mathbf{p}$, since orbital angular momentum contributes nothing to this projection (Exercise 24.12).

[6]See Rose (1961), Section 11.

The corresponding solutions for the eigenvalue $E = -E_p$ are

$$\nu^{(R)} \propto \begin{pmatrix} \dfrac{-cp}{E_p + mc^2} \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \nu^{(L)} \propto \begin{pmatrix} 0 \\ \dfrac{cp}{E_p + mc^2} \\ 0 \\ 1 \end{pmatrix} \tag{24.89}$$

Exercise 24.20. Show that the only component of the matrix $\boldsymbol{\Sigma}$ that commutes with $H_{\mathbf{p}}$ is the helicity operator $\boldsymbol{\Sigma} \cdot \hat{\mathbf{p}}$.

Exercise 24.21. Determine the multiplicative constants for each of the four solutions (24.88) and (24.89), ensuring the normalization (24.17).

Exercise 24.22. Show that the free-particle spinors for $\mathbf{p}$ in the z direction can also be generated by applying a Lorentz transformation (also known as a *Lorentz boost*) to the trivial solutions of (24.86) for a particle at rest ($\mathbf{p} = 0$). Refer to Exercise 24.13 and Eq. (24.74).

The eigenspinors with definite helicity but arbitrary linear momentum vector $\mathbf{p}$ are easily found by rotating the states described by (24.88) and (24.89) by an angle $\phi = \arccos(p_z/p)$ about the axis determined by the vector $(-p_y, p_x, 0)$. Such a rotation takes the z axis into the direction of $\mathbf{p}$. The matrix operator that carries out this rotation is

$$\mathsf{S} = \exp\left[-i \frac{\phi}{2} \frac{(-p_y \Sigma_x + p_x \Sigma_y)}{\sqrt{p_x^2 + p_y^2}} \right] \tag{24.90}$$

Using the generalization of the identity (16.62) to the 4×4 Pauli matrices, we may write this as

$$\mathsf{S} = I \cos \frac{\phi}{2} + i \frac{(p_y \Sigma_x - p_x \Sigma_y)}{\sqrt{p_x^2 + p_y^2}} \sin \frac{\phi}{2} \tag{24.91}$$

Hence, if the components of $\hat{\mathbf{p}}$ are denoted as (n_x, n_y, n_z), the rotated spinors are

$$u^{(R)}(\mathbf{p}) \propto \begin{pmatrix} \cos \dfrac{\phi}{2} \\ \dfrac{(p_x + ip_y)}{\sqrt{p^2 - p_z^2}} \sin \dfrac{\phi}{2} \\ \dfrac{cp}{E_p + mc^2} \cos \dfrac{\phi}{2} \\ \dfrac{cp}{E_p + mc^2} \dfrac{p_x + ip_y}{\sqrt{p^2 - p_z^2}} \sin \dfrac{\phi}{2} \end{pmatrix} \propto \begin{pmatrix} n_z + 1 \\ n_x + in_y \\ \dfrac{cp}{E_p + mc^2} (n_z + 1) \\ \dfrac{cp}{E_p + mc^2} (n_x + in_y) \end{pmatrix} \tag{24.92}$$

and

$$u^{(L)}(\mathbf{p}) \propto \begin{pmatrix} \dfrac{-p_x + ip_y}{\sqrt{p^2 - p_z^2}} \sin\dfrac{\phi}{2} \\ \cos\dfrac{\phi}{2} \\ -\dfrac{cp}{E_p + mc^2} \dfrac{-p_x + ip_y}{\sqrt{p^2 - p_z^2}} \sin\dfrac{\phi}{2} \\ -\dfrac{cp}{E_p + mc^2} \cos\dfrac{\phi}{2} \end{pmatrix} \propto \begin{pmatrix} -n_x + in_y \\ n_z + 1 \\ -\dfrac{cp}{E_p + mc^2}(-n_x + in_y) \\ -\dfrac{cp}{E_p + mc^2}(n_z + 1) \end{pmatrix} \tag{24.93}$$

Exercise 24.23. Work out similar expressions for $\nu^{(R)}$ and $\nu^{(L)}$.

Exercise 24.24. Verify the closure relation (24.33).

The matrix

$$\mathsf{B}_+(\mathbf{p}) = u^{(R)}(\mathbf{p})u^{(R)\dagger}(\mathbf{p}) + u^{(L)}(\mathbf{p})u^{(L)\dagger}(\mathbf{p}) \tag{24.94}$$

constructed from the normalized eigenspinors, gives zero when applied to an eigenspinor of $c\boldsymbol{\alpha}\cdot\mathbf{p} + \beta mc^2$ with eigenvalue $-E_p$; applied to an eigenspinor of $c\boldsymbol{\alpha}\cdot\mathbf{p} + \beta mc^2$ with eigenvalue E_p, it acts like the unit matrix. Hence, it can be expressed as the (*Casimir*) projection operator

$$\mathsf{B}_+(\mathbf{p}) = \frac{E_p + c\boldsymbol{\alpha}\cdot\mathbf{p} + \beta mc^2}{2E_p} \tag{24.95}$$

Similarly, the matrix

$$\mathsf{B}_-(\mathbf{p}) = \nu^{(R)}(\mathbf{p})\nu^{(R)\dagger}(\mathbf{p}) + \nu^{(L)}(\mathbf{p})\nu^{(L)\dagger}(\mathbf{p}) = \frac{E_p - c\boldsymbol{\alpha}\cdot\mathbf{p} - \beta mc^2}{2E_p} \tag{24.96}$$

acts as a projection operator for the eigenspinors with eigenvalue $-E_p$. These projection operators are useful in many applications.

Exercise 24.25. Show that for the eigenspinors of fixed momentum $\mathbf{p}$ and an arbitrary 4×4 matrix A,

$$u^{(R)\dagger}\mathsf{A}u^{(R)} + u^{(L)\dagger}\mathsf{A}u^{(L)} + \nu^{(R)\dagger}\mathsf{A}\nu^{(R)} + \nu^{(L)\dagger}\mathsf{A}\nu^{(L)} = \text{trace } \mathsf{A} \tag{24.97}$$

5. ***Charge Conjugation, Time Reversal, and the PCT Theorem.*** A simple relationship follows from the similarity of the equations satisfied by $\nu^{(R,L)}(\mathbf{p})$ and $u^{(R,L)}(\mathbf{p})$. If the complex conjugate of Eq. (24.28) is taken and $\mathbf{p}$ is replaced by $-\mathbf{p}$, we obtain

$$(-c\boldsymbol{\alpha}^*\cdot\mathbf{p} + mc^2\beta^*)\nu^{(R,L)*}(-\mathbf{p}) = -E_p\nu^{(R,L)*}(-\mathbf{p})$$

which is to be compared with (24.27),

$$(c\boldsymbol{\alpha}\cdot\mathbf{p} + mc^2\beta)u^{(R,L)}(\mathbf{p}) = E_p u^{(R,L)}(\mathbf{p})$$

If we can find a matrix C with the properties

$$\mathsf{C}\boldsymbol{\alpha}^* = \boldsymbol{\alpha}\mathsf{C}, \qquad \mathsf{C}\beta^* = -\beta\mathsf{C} \tag{24.98}$$

it is seen that $\mathsf{C}\nu^{(R,L)*}(-\mathbf{p})$ satisfies the same equation as $u^{(R,L)}(\mathbf{p})$.

Exercise 24.26. Establish that $\mathsf{C}^{-1}\boldsymbol{\gamma}\mathsf{C} = \tilde{\boldsymbol{\gamma}}$ as well as $\Sigma\mathsf{C} = -\mathsf{C}\Sigma^*$ and $\mathsf{C}^{-1}\gamma^0\mathsf{C} = -\tilde{\gamma}^0$.

Helicity is preserved under this transformation of solutions of the Dirac equation. Indeed, from the equations

$$\boldsymbol{\Sigma}\cdot\mathbf{p}\mathsf{C}\nu^{(R)*}(-\mathbf{p}) = -\mathsf{C}(\boldsymbol{\Sigma}\cdot\mathbf{p})^*\nu^{(R)*}(-\mathbf{p}) = \mathsf{C}\nu^{(R)*}(-\mathbf{p})$$
$$\boldsymbol{\Sigma}\cdot\mathbf{p}\mathsf{C}\nu^{(L)*}(-\mathbf{p}) = -\mathsf{C}(\boldsymbol{\Sigma}\cdot\mathbf{p})^*\nu^{(L)*}(-\mathbf{p}) = -\mathsf{C}\nu^{(L)*}(-\mathbf{p})$$

if follows that the identification

$$u^{(R)}(\mathbf{p}) = \mathsf{C}\nu^{(R)*}(-\mathbf{p}) \tag{24.99}$$

and

$$u^{(L)}(\mathbf{p}) = \mathsf{C}\nu^{(L)*}(-\mathbf{p}) \tag{24.100}$$

may be made. By using the same matrix C in both of these equations, we make a partial choice of the previously undetermined relative phases of the spinors u and ν. The normalization (24.17) requires that

$$u^{(R,L)\dagger}(\mathbf{p})u^{(R,L)}(\mathbf{p}) = \tilde{\nu}^{(R,L)}(-\mathbf{p})\mathsf{C}^\dagger\mathsf{C}\nu^{(R,L)*}(-\mathbf{p}) = \nu^{(R,L)\dagger}(-\mathbf{p})\tilde{\mathsf{C}}\mathsf{C}^*\nu^{(R,L)}(-\mathbf{p}) = 1$$

or

$$\tilde{\mathsf{C}}\mathsf{C}^* = I \tag{24.101}$$

i.e., C must be unitary. In the standard representation (24.32), the conditions (24.98) and (24.101) are satisfied by the matrix

$$\mathsf{C} = \beta\alpha_y = \gamma^2$$

Equations (24.14) and (24.15) show that $u^{(R,L)}(\mathbf{p})$ is associated with the annihilation of an electron and $\nu^{(R,L)*}(-\mathbf{p})$ with the annihilation of a positron. The connection (24.99) and (24.100) between these two amplitudes suggests that the unitary transformation that takes electrons into positrons and vice versa, without changing either momentum of helicity, may have a simple local formulation in terms of the fields. We define the unitary operator C, known as the *charge conjugation* or *particle-antiparticle conjugation* operator, by the equations

$$\begin{aligned} Ca_R(\mathbf{p})C^\dagger &= b_R(\mathbf{p}) \qquad Cb_R(\mathbf{p})C^\dagger = a_R(\mathbf{p}) \\ Ca_L(\mathbf{p})C^\dagger &= b_L(\mathbf{p}) \qquad Cb_L(\mathbf{p})C^\dagger = a_L(\mathbf{p}) \end{aligned} \tag{24.102}$$

From Eqs. (24.14), (24.15), (24.99), and (24.100), it is seen that

$$C\boldsymbol{\psi}(\mathbf{r})C^\dagger = \mathsf{C}\tilde{\boldsymbol{\gamma}}^0\overline{\boldsymbol{\psi}}(\mathbf{r}) = \mathsf{C}\tilde{\boldsymbol{\psi}}^\dagger(\mathbf{r}) \tag{24.103}$$

Exercise 24.27. Verify Eq. (24.103) and show conversely that

$$C\boldsymbol{\psi}^\dagger(\mathbf{r})C^\dagger = \tilde{\boldsymbol{\psi}}(\mathbf{r})\mathsf{C}^{-1} \tag{24.104}$$

The definition of C is supplemented by requiring that the vacuum state remain unchanged under charge conjugation: $C|\mathbf{0}\rangle = |\mathbf{0}\rangle$.

As time develops, the relations (24.103) and (24.104) remain applicable if the electron-positron field is free. This follows from the definition of charge conjugation and can be verified by showing that if $\boldsymbol{\psi}(\mathbf{r}, t)$ and $\boldsymbol{\psi}^\dagger(\mathbf{r}, t)$ are connected at all times

by (24.103), the two Dirac equations (24.42) and (24.44) with $A_\mu = 0$ imply one another.

Exercise 24.28. Prove the last statement.

In the presence of an external electromagnetic field ($A_\mu \neq 0$), the Dirac equation (24.42) is no longer invariant under charge conjugation as defined by relation (24.103). Applying this operation to Eq. (24.42), we obtain

$$\gamma^\mu\left(\frac{\partial}{\partial x^\mu} - \frac{ie}{\hbar c} A_\mu\right)\mathsf{C}\tilde{\psi}^\dagger + i\kappa\mathsf{C}\tilde{\psi}^\dagger = 0$$

If, by using the commutation properties of the matrix C, we reduce this equation, we find

$$\left(\frac{\partial}{\partial x^\mu} - \frac{ie}{\hbar c} A_\mu\right)\bar{\psi}\gamma^\mu - i\kappa\bar{\psi} = 0 \tag{24.105}$$

This equation is the same as Eq. (24.44) except for the important change of the sign in front of the vector potential. The presence of an external field thus destroys the invariance of the theory under charge conjugation. At the same time, it is apparent that the invariance is restored if the electromagnetic field is regarded as part of the dynamical system and is reversed ($A_\mu \to -A_\mu$) when charge conjugation is applied.

Exercise 24.29. Reproduce the steps leading to Eq. (24.105), using the results of Exercise 24.26 and the properties of the Dirac matrices.

Exercise 24.30. Show that under charge conjugation the current density operator, defined in Eq. (24.47), changes into its negative if the anticommutation properties of the field are used.

We now return briefly to the parity operator U_P, defined in Eq. (24.81) by its action on the field:

$$U_P\psi(\mathbf{r}, t)U_P^\dagger = \gamma^0\psi(-\mathbf{r}, t) \tag{24.106}$$

We note that together with $u^{(R,L)}(\mathbf{p})$ the spinors $\gamma^0 u^{(R,L)}(-\mathbf{p})$, obtained by reflection, are also solutions of Eq. (24.27). Since $\boldsymbol{\Sigma}\cdot\hat{\mathbf{p}}$ changes sign under reflection, the helicity is reversed, and $\gamma^0 u^{(R)}(-\mathbf{p})$ must be proportional to $u^{(L)}(\mathbf{p})$. Similarly, $\gamma^0 v^{(R)}(-\mathbf{p})$ must be proportional to $v^{(L)}(\mathbf{p})$. It is consistent with the relations (24.99) and (24.100), and the condition $\gamma^0\mathsf{C} = -\mathsf{C}\tilde{\gamma}^0$ (Exercise 24.26) to set

$$\gamma^0 u^{(R)}(-\mathbf{p}) = u^{(L)}(\mathbf{p}) \tag{24.107}$$

Since γ^0 is Hermitian, the equations

$$u^{(L)}(\mathbf{p}) = \gamma^0 u^{(R)}(-\mathbf{p}) = \gamma^0\mathsf{C}v^{(R)*}(\mathbf{p}) = -\mathsf{C}\tilde{\gamma}^0 v^{(R)*}(\mathbf{p}) = -\mathsf{C}[\gamma^0 v^{(R)}(\mathbf{p})]^*$$

and (24.100) lead to the conclusion that we must have

$$\gamma^0 v^{(R)}(-\mathbf{p}) = -v^{(L)}(\mathbf{p}) \tag{24.108}$$

From (24.106), (24.107), and (24.108), we deduce the transformation properties of the electron and positron annihilation operators under spatial reflection as

$$\begin{aligned} U_P a_R(\mathbf{p}) U_P^\dagger &= a_L(-\mathbf{p}) \qquad & U_P a_L(\mathbf{p}) U_P^\dagger &= a_R(-\mathbf{p}) \\ U_P b_R(\mathbf{p}) U_P^\dagger &= -b_L(-\mathbf{p}) \qquad & U_P b_L(\mathbf{p}) U_P^\dagger &= -b_R(-\mathbf{p}) \end{aligned} \tag{24.109}$$

The difference in sign between the equations in the first and second rows has important physical consequences, since it implies that an electron and a positron in the same orbital states have opposite parities.[7]

We conclude this discussion of discrete symmetries with some remarks about time reversal. The general concepts needed for the discussion were already presented in Section 17.9. The *antiunitary* time reversal operator Θ is defined to reverse the sign of all momenta and spins. We therefore require that

$$\begin{aligned} \Theta a_R(\mathbf{p}) \Theta^{-1} &= e^{i\alpha_R(\mathbf{p})} a_R(-\mathbf{p}) \qquad & \Theta a_L(\mathbf{p}) \Theta^{-1} &= e^{i\alpha_L(\mathbf{p})} a_L(-\mathbf{p}) \\ \Theta b_R(\mathbf{p}) \Theta^{-1} &= e^{i\beta_R(\mathbf{p})} b_R(-\mathbf{p}) \qquad & \Theta b_L(\mathbf{p}) \Theta^{-1} &= e^{i\beta_L(\mathbf{p})} b_L(-\mathbf{p}) \end{aligned} \tag{24.110}$$

Although the phases in (24.110) are arbitrary, it is possible to choose them in such a manner that the fields undergo simple transformations under time reversal. From the antiunitary property of Θ, one may derive the transformation properties of the creation operators:

$$\begin{aligned} \Theta a_R^\dagger(\mathbf{p}) \Theta^{-1} &= e^{-i\alpha_R(\mathbf{p})} a_R^\dagger(-\mathbf{p}) \qquad & \Theta a_L^\dagger(\mathbf{p}) \Theta^{-1} &= e^{-i\alpha_L(\mathbf{p})} a_L^\dagger(-\mathbf{p}) \\ \Theta b_R^\dagger(\mathbf{p}) \Theta^{-1} &= e^{-i\beta_R(\mathbf{p})} b_R^\dagger(-\mathbf{p}) \qquad & \Theta b_L^\dagger(\mathbf{p}) \Theta^{-1} &= e^{-i\beta_L(\mathbf{p})} b_L^\dagger(-\mathbf{p}) \end{aligned} \tag{24.111}$$

Exercise 24.31. Derive (24.111) from (24.110).

If we apply Θ to the fields defined in Eqs. (24.36) and (24.37), and make some trivial substitutions in the integrand, we obtain from (24.110),

$$\Theta \psi^{(+)}(\mathbf{r}, t) \Theta^{-1} = \frac{1}{(2\pi\hbar)^{3/2}} \int [u^{(R)*}(-\mathbf{p}) e^{i\alpha_R(-\mathbf{p})} a_R(\mathbf{p}) + u^{(L)*}(-\mathbf{p}) e^{i\alpha_L(-\mathbf{p})} a_L(\mathbf{p})] e^{(i/\hbar)(\mathbf{p}\cdot\mathbf{r} + E_p t)} \, d^3p$$

In arriving at this equation, the antiunitary nature of Θ is used, resulting in complex conjugation. The right-hand side of this equation becomes a local expression, $\mathsf{T}\psi^{(+)}(\mathbf{r}, -t)$, for the time-reversed field, if a 4×4 matrix T can be found such that

$$u^{(R,L)*}(-\mathbf{p}) e^{i\alpha_{R,L}(-\mathbf{p})} = \mathsf{T} u^{(R,L)}(\mathbf{p}) \tag{24.112}$$

The normalization (24.33) implies that T must be unitary:

$$\mathsf{T}\mathsf{T}^\dagger = I \tag{24.113}$$

The relations (24.112) are consistent with the Dirac equation (24.27) only if T satisfies the conditions

$$\boldsymbol{\alpha}^* \mathsf{T} = -T\boldsymbol{\alpha}, \qquad \beta^* \mathsf{T} = \mathsf{T}\beta \tag{24.114}$$

The unitary solution to these equations is unique except for an arbitrary phase factor.

[7]For illustrations of the selection rules that can be derived for interactions invariant under reflection and charge conjugation, see Sakurai (1967), Section 4-4.

Exercise 24.32. If T_1 and T_2 are two different unitary matrices that satisfy (24.114), construct $\mathsf{T}_1\mathsf{T}_2^{-1}$. Show that this commutes with all Dirac matrices and hence must be a multiple of the unit matrix.

In the standard representation (24.32), the imaginary matrix

$$\mathsf{T} = -i\alpha_z\alpha_x = \Sigma_y \tag{24.115}$$

is a solution. It has the important property

$$\mathsf{T}\,\mathsf{T}^* = -1 \tag{24.116}$$

which can be proved to be independent of the representation.

Exercise 24.33. Apply the time reversal operator to the negative frequency part of the field, and show that the same matrix T may be used to transform $\psi^{(-)}$ as $\psi^{(+)}$.

It follows that the complete electron-positron field is transformed under time reversal according to

$$\boxed{\Theta\psi(\mathbf{r}, t)\Theta^{-1} = \mathsf{T}\psi(\mathbf{r}, -t), \qquad \Theta\psi^\dagger(\mathbf{r}, t)\Theta^{-1} = \psi^\dagger(\mathbf{r}, -t)\mathsf{T}^\dagger} \tag{24.117}$$

Exercise 24.34. Show that $\mathsf{T}\boldsymbol{\Sigma}\cdot\hat{\mathbf{p}} = -\boldsymbol{\Sigma}^*\cdot\hat{\mathbf{p}}\mathsf{T}$ and that helicity is preserved under time reversal, substantiating (24.112).

The properties of T impose restrictions on the phases α and β in (24.110) and (24.111). Iteration of Eqs. (24.112), in conjunction with the requirement (24.116), gives the result

$$e^{i\alpha_{R,L}(-\mathbf{p})-i\alpha_{R,L}(\mathbf{p})} = e^{i\beta_{R,L}(-\mathbf{p})-i\beta_{R,L}(\mathbf{p})} = -1 \tag{24.118}$$

The effect of two successive time reversals can now be established. Owing to the antilinearity of Θ:

$$\Theta^2 a_R(\mathbf{p})\Theta^{-2} = e^{-i\alpha_R(\mathbf{p})}\Theta a_R(-\mathbf{p})\Theta^{-1} = e^{-i\alpha_R(\mathbf{p})+i\alpha_R(-\mathbf{p})}a_R(\mathbf{p}) = -a_R(\mathbf{p}) \tag{24.119}$$

Thus, application of Θ^2 merely changes the sign of the annihilation operator. The same conclusion holds for all other annihilation and creation operators. Hence, Θ^2 acts like $+1$ on states with an even number of Dirac particles (and, more generally, fermions), and like -1 on states with an odd number of such particles. This conclusion agrees with the discussion of Section 17.9 (Kramers degeneracy).

Double time reversal cannot have any physical consequences, and Θ^2 is a unitary operator that commutes with all observables. State vectors that are obtained by superposition of states with even and odd numbers of Dirac particles undergo an unacceptable transformation under the application of Θ^2 and cannot be physically realized. This statement is a *superselection rule*, and it is consistent with a superselection rule inferred in Section 17.2 from the commutivity of observables with rotations by 2π: *Fermions cannot be created or destroyed in odd numbers.*

In the presence of an external electromagnetic field, the time reversal operation is generally no longer a symmetry operation. However, the invariance of the Dirac equation (24.43) under time reversal as defined by (24.117) is restored if $\mathbf{A}$ is changed into $-\mathbf{A}$, while ϕ is left unchanged.

Exercise 24.35. Determine the transformation properties of the Dirac current density operator under time reversal.

In addition to angular momentum operators and other generators of the proper Lorentz group, we have now discussed three discrete symmetry operations corresponding to reflection, charge conjugation, and time reversal. Originally defined for free fields, these operations may remain symmetry operations when interactions are introduced. For instance, quantum electrodynamics is invariant under each of these three operations. Weak interactions are not invariant under the three operations separately. However, invariance still holds for the product (i.e., the successive application) of the three discrete operations (*PCT theorem*), provided only that the restricted principle of relativity is valid.[8]

6. ***The One-Particle Approximation.*** In Chapter 21, quantum field theory was developed into a consistent description of systems of identical particles from the concepts of nonrelativistic quantum mechanics for a single particle. But, as was emphasized in Chapter 23 for photons and in Section 24.1 for electrons, there are obstacles in the way to constructing a relativistic form of local one-particle quantum mechanics in a rigorous manner. Inevitably, such a theory is an approximation to a proper many-body theory.

It is tempting to identify the state

$$\psi_\alpha^\dagger(\mathbf{r})|\mathbf{0}\rangle = |\mathbf{r}, \alpha\rangle \tag{24.120}$$

in analogy to Eqs. (21.3) and (21.6) as the one-electron state that corresponds to a sharp position of the particle with discrete quantum number α. The inadequacy of this identification in the relativistic theory is seen from the fact that, owing to the anticommutation relations, the state (24.120) cannot be normalized properly. The trouble stems, of course, from the properties of the field ψ whose expansion contains *both* electron annihilation and positron creation operators, so that

$$\psi_\alpha(\mathbf{r})|\mathbf{0}\rangle \neq 0$$

An obvious possibility for remedying this difficulty is to use the positive frequency field, instead of Eq. (24.120), and to make the identification

$$\psi_\alpha^{(+)\dagger}(\mathbf{r})|\mathbf{0}\rangle = |\mathbf{r}, \alpha\rangle \tag{24.121}$$

for a one-electron state. Such a theory, if pursued, would contain one-electron wave functions that do not correspond to any state in the original field theory (e.g., the negative energy eigenstates of H). This difficulty would make its appearance whenever we encountered an operator, such as a strong potential energy V, that connects the "physical" with the "unphysical" states. Arbitrary exclusion of the "unphysical" states would violate the completeness requirements and lead to incorrect results in calculations that involve virtual intermediate states. On the other hand, their inclusion would be embarrassing, since the theory then permits transitions to "unphysical" states if a strongly fluctuating perturbation is applied.[9]

The conclusion is inescapable that in relativistic quantum mechanics there can be no one-particle state that describes a particle at position $\mathbf{r}$. Although we can say

[8]For further discussion of the *PCT* theorem, see Gross (1993), Section 8.7.
[9]For an interesting discussion of these difficulties, see Sakurai (1967), pp. 120–121.

how much charge there is in a small volume in space, we cannot specify precisely how many particles are located in this volume. On the other hand, we are able to count the number of particles with a specified value of momentum **p**, as we did in Section 24.1. (The asymmetry between the descriptions in coordinate and momentum space in the relativistic theory can ultimately be traced to the qualitative difference between time and energy: Time runs from $-\infty$ to $+\infty$, whereas energy is always bounded from below.)

Instead of modifying the field operators, we choose to formulate the one-particle approximation by introducing a new "*electron vacuum*" $|\mathbf{0}e\rangle$ such that

$$\boldsymbol{\psi}_\alpha(\mathbf{r})|\mathbf{0}e\rangle = 0 \tag{24.122}$$

holds. Obviously, this is a state in which there are no electrons but in which all available positron one-particle states are occupied. Thus, it can hardly be called the physical vacuum, since relative to the no-particle vacuum it has infinite positive charge and infinite energy. Nevertheless, departures from this state by the addition of one electron or subtraction of one positron can be treated effectively as one-particle states. We thus define a *one-electron state* as

$$\boxed{|\Psi_e\rangle = \int \boldsymbol{\psi}^\dagger(\mathbf{r}, t)\psi_e(\mathbf{r}, t)\, d^3r\, |\mathbf{0}e\rangle} \tag{24.123}$$

where $\psi_e(\mathbf{r}, t)$ is a *spinor wave function*—not an operator—with four components. From the equal-time anticommutation relation,

$$\boldsymbol{\psi}_\alpha^\dagger(\mathbf{r}, t)\boldsymbol{\psi}_\beta(\mathbf{r}', t) + \boldsymbol{\psi}_\beta(\mathbf{r}', t)\boldsymbol{\psi}_\alpha^\dagger(\mathbf{r}, t) = \delta_{\alpha\beta}\delta(\mathbf{r} - \mathbf{r}')$$

and Eq. (24.122), it follows immediately that the normalization

$$\langle\Psi_e|\Psi_e\rangle = \int \psi_e^\dagger(\mathbf{r}, t)\psi_e(\mathbf{r}, t)\, d^3r = 1$$

is applicable and that the electron wave function is

$$\psi_e(\mathbf{r}, t) = \langle\mathbf{0}e|\boldsymbol{\psi}(\mathbf{r}, t)|\Psi_e\rangle \tag{24.124}$$

in analogy to Eq. (21.57) in the nonrelativistic theory.

Exercise 24.36. Verify that the total charge of the system in state Ψ_e of Eq. (24.123) differs from the charge in the new vacuum state $|\mathbf{0}e\rangle$ by $-e$,

Similarly, a new "*positron vacuum*" $|\mathbf{0}p\rangle$ and a one-positron state $|\Psi_p\rangle$ are defined by the equations

$$\boldsymbol{\psi}^\dagger(\mathbf{r})|\mathbf{0}p\rangle = 0 \tag{24.125}$$

$$|\Psi_p\rangle = \int \tilde{\boldsymbol{\psi}}(\mathbf{r}, t)\mathbf{C}^{-1}\psi_p(\mathbf{r}, t)\, d^3r\, |\mathbf{0}p\rangle \tag{24.126}$$

From the normalization of these states to unity, the *positron wave function* $\psi_p(\mathbf{r}, t)$ is obtained as

$$\psi_p(\mathbf{r}, t) = \langle\mathbf{0}p|\mathbf{C}\tilde{\boldsymbol{\psi}}^\dagger(\mathbf{r}, t)|\Psi_p\rangle \tag{24.127}$$

In spacetime, the spinor *wave functions* ψ_e and ψ_p satisfy the same equations (24.42) and (24.44) as the field operators:

$$\boxed{\gamma^\mu\left(\frac{\partial}{\partial x^\mu} - \frac{ie}{\hbar c}A_\mu\right)\psi_e + i\kappa\psi_e = 0} \tag{24.128}$$

$$\boxed{\gamma^\mu\left(\frac{\partial}{\partial x^\mu} + \frac{ie}{\hbar c}A_\mu\right)\psi_p + i\kappa\psi_p = 0} \tag{24.129}$$

These are the relativistic generalizations of the one-particle time-dependent Schrödinger equation for particles with spin one-half. Dirac originally proposed the equation bearing his name in the one-particle form (24.128).

To recover the usual probability interpretation of one-particle quantum mechanics, we consider an additive Hermitian one-particle operator, like the linear or angular momentum, which in Sections 24.1 and 24.2 was written in the form

$$\mathcal{K} = \int \boldsymbol{\psi}^\dagger(\mathbf{r})K\boldsymbol{\psi}(\mathbf{r})\, d^3r$$

Here $K(\mathbf{r}, -i\hbar\nabla)$ is a function of the position and momentum vectors and, in addition, may also be a 4×4 matrix acting on a spinor. Using (24.123) and its inverse (24.124), we obtain for such an operator the one-electron expectation value

$$(\Psi_e, \mathcal{K}\Psi_e) = \int \psi_e^\dagger(\mathbf{r})K\psi_e(\mathbf{r})\, d^3r \tag{24.130}$$

Exercise 24.37. Show that similarly for a positron,

$$(\Psi_p, \mathcal{K}\Psi_p) = \int \psi_p^\dagger(\mathbf{r})K\psi_p(\mathbf{r})\, d^3r \tag{24.131}$$

if $\mathcal{K}$ represents a physical quantity that is invariant under charge conjugation and that therefore satisfies the condition

$$\mathsf{C}^{-1}K\mathsf{C} = -K^* \tag{24.132}$$

Verify that linear and angular momentum operators satisfy this condition, as does the free-particle energy operator.

Owing to the connection (24.99) and (24.100), if forces are absent, the wave function of a free electron or positron with momentum $\mathbf{p}$ and definite helicity is the plane wave,

$$\frac{1}{(2\pi\hbar)^{3/2}}\, u^{(R,L)}(\mathbf{p})\; e^{(i/\hbar)(\mathbf{p}\cdot\mathbf{r}-E_p t)} = \frac{1}{(2\pi\hbar)^{3/2}}\, \mathsf{C}v^{(R,L)*}(-\mathbf{p})e^{(i/\hbar)(\mathbf{p}\cdot\mathbf{r}-E_p t)} \tag{24.133}$$

The approximate nature of the one-electron or one-positron theory is apparent in many ways. For example, the equation of motion for a free one-electron wave function,

$$\gamma^\mu \frac{\partial\psi_e}{\partial x^\mu} + i\kappa\psi_e = 0$$

has, for a given momentum $\mathbf{p}$, four linearly independent solutions. Two of these, for positive energy, correspond to the two spin states of the electron. The remaining

two are eigenstates of H with negative eigenvalues and represent, according to the definition (24.123), the removal of a positron from (rather than the addition of an electron to) the "vacuum" $|\mathbf{0}e\rangle$. These solutions cannot be ignored since, for a spin one-half particle with mass, Lorentz invariance requires that ψ_e have four components, so that four linearly independent spinors are needed to specify an arbitrary initial state. Even if initially the wave function were a superposition of only electron eigenstates, the amplitudes of the positron components may, under the influence of forces, eventually become appreciable.

As a simple example, we imagine that the free Dirac electron is subjected to a perturbation during a time interval τ. According to (19.23) of time-dependent perturbation theory, the transition amplitude for an exponential perturbation is

$$-\frac{i}{\hbar}\langle k|V|s\rangle \int_{-\infty}^{+\infty} e^{i\omega_{ks}t'} e^{-|t'|/\tau}\, dt' = -\frac{2i}{\hbar}\frac{\tau}{1+\omega_{ks}^2\tau^2}\langle k|V|s\rangle$$

The optimum value of τ is evidently $\tau = 1/\omega_{ks}$. For this value, the transition amplitude has magnitude $|\langle k|V|s\rangle|/\hbar\omega_{ks}$. Since $\hbar\omega_{ks} \approx mc^2$ for transitions between positive and negative eigenstates of H, it is apparent that the one-particle approximation breaks down when interaction energies of strength $\approx mc^2$ fluctuate in times of the order $\approx \hbar/mc^2$.

Translating these considerations from time into space language, we can say that if the potential energy changes by mc^2 over a distance of the order $\hbar/mc$, an initial one-electron state may lead to *pair annihilation* of the electron with one of the positrons present in the state $|\mathbf{0}e\rangle$. Properly, it may thus be said that we are dealing with a *one-charge* rather than a *one-particle* theory.

Understanding and accepting the one-electron Dirac theory as an approximation to a more accurate description involving interacting fields, we now ignore the many-body aspects and consider the relativistic wave equation,

$$\boxed{i\hbar\frac{\partial\psi}{\partial t} = \left[c\boldsymbol{\alpha}\cdot\left(\frac{\hbar}{i}\nabla + \frac{e}{c}\mathbf{A}\right) - e\phi + \beta mc^2\right]\psi} \tag{24.134}$$

on its own merits. From here on, instead of writing ψ_e, we omit the subscript e and simply use ψ to denote the Dirac electron wave function. We emphasize again that it is a spinor function rather than a field operator.

Just as in Section 24.2, we can derive the continuity equation,

$$\frac{\partial(\psi^\dagger\psi)}{\partial t} + \nabla\cdot(c\psi^\dagger\boldsymbol{\alpha}\psi) = 0 \tag{24.135}$$

defining a probability density

$$\rho = \psi^\dagger\psi \tag{24.136}$$

and the probability current density

$$\mathbf{j} = c\psi^\dagger\boldsymbol{\alpha}\psi \tag{24.137}$$

Equation (24.134) has the usual form

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi \tag{24.138}$$

familiar from ordinary quantum mechanics, with H being a Hermitian operator. The only unusual feature of H is the fact that, unlike the nonrelativistic one-particle Hamiltonian and unlike the total field energy operator $\mathcal{H}$, the one-particle H,

$$H = c\boldsymbol{\alpha} \cdot \left(\frac{\hbar}{i}\nabla + \frac{e}{c}\mathbf{A}\right) - e\phi + \beta mc^2 \tag{24.139}$$

is not a positive definite operator. Nonetheless, we will see in Section 24.8 that a simple correspondence exists between the relativistic and nonrelativistic Hamiltonians.

The stationary state solutions of the Dirac equation for a *free electron* (or positron) need not be discussed in detail here, because this was in effect already done in Section 24.4. In the nonrelativistic limit, $E_p \approx mc^2$, in the standard representation (24.88), the third and fourth components of $u^{(R,L)}$ are small in the ratio of v/c compared to the first two components. The converse is true for the spinors $v^{(R,L)}$, as (24.89) shows. It is, therefore, customary to speak of "large" and "small" components of the Dirac wave functions, but this terminology is dependent on the representation used for the Dirac matrices. The stationary state solutions of (24.134) for a static central potential will be taken up in Section 24.9.

7. ***Dirac Theory in the Heisenberg Picture.*** In quantum field theory, the time development of the field operators $\psi(\mathbf{r}, t)$ in (24.38) was formulated in terms of the *Heisenberg* picture. When the transition to a one-electron theory is made, Eq. (24.134) for the spinor wave functions $\psi(\mathbf{r}, t)$ is expressed in the *Schrödinger* picture. We now study the one-electron theory in the Heisenberg picture, where state vectors are constant while operators that are not explicitly time-dependent evolve according to the formula

$$i\hbar\frac{dA(t)}{dt} = A(t)H(t) - H(t)A(t) \equiv [A(t), H(t)]$$

With the understanding that we are working in the Heisenberg picture, for the purposes of this section only we may simplify the notation by writing A for $A(t)$.

The Hamiltonian operator has the form

$$H = c\boldsymbol{\alpha} \cdot \left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right) - e\phi + \beta mc^2 \tag{24.140}$$

where now not only $\mathbf{r}$ and $\mathbf{p}$ but also the matrices $\boldsymbol{\alpha}$ and β must be regarded as dynamical variables, subject to time evolution.

If the time derivative of $\mathbf{r}$ is defined as the *velocity operator* $\mathbf{v}$, we obtain

$$\mathbf{v} = \frac{d\mathbf{r}}{dt} = \frac{1}{i\hbar}[\mathbf{r}, H] = c\boldsymbol{\alpha} \tag{24.141}$$

Although in classical mechanics they are equal, in the relativistic quantum theory the operator $m\mathbf{v}$ formed from (24.141) is not the same as the kinetic momentum, $\mathbf{p} + (e/c)\mathbf{A}$. To derive the equation of motion for the latter, we note that

$$\frac{d\mathbf{p}}{dt} = \frac{1}{i\hbar}[\mathbf{p}, H] = e\nabla\phi - e\nabla(\boldsymbol{\alpha} \cdot \mathbf{A})$$

and

$$\frac{d\mathbf{A}}{dt} = \frac{\partial \mathbf{A}}{\partial t} + \frac{1}{i\hbar}[\mathbf{A}, H] = \frac{\partial \mathbf{A}}{\partial t} + c\boldsymbol{\alpha} \cdot \nabla \mathbf{A}$$

Hence, combining the last two equations,

$$\frac{d}{dt}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right) = e\nabla\phi + \frac{e}{c}\frac{\partial \mathbf{A}}{\partial t} + e\boldsymbol{\alpha} \cdot \nabla \mathbf{A} - e\nabla(\boldsymbol{\alpha} \cdot \mathbf{A})$$
$$= -e\left(-\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} - \nabla\phi\right) - e\boldsymbol{\alpha} \times (\nabla \times \mathbf{A})$$

or

$$\frac{d}{dt}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right) = -e\mathbf{E} - e\boldsymbol{\alpha} \times \mathbf{B} = -e\mathbf{E} - e\frac{\mathbf{v}}{c} \times \mathbf{B} \tag{24.142}$$

This equation can be written in a form that is even more reminiscent of classical physics if we note the identity

$$\boldsymbol{\alpha}(H + e\phi) + (H + e\phi)\boldsymbol{\alpha} = 2c\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right) \tag{24.143}$$

which shows the connection between $m\mathbf{v} = mc\boldsymbol{\alpha}$ and $\mathbf{p} + (e/c)\mathbf{A}$.[10]

Exercise 24.38. Carry out the details of the calculation leading to (24.142), evaluating all requisite commutators. Verify the anticommutation property (24.143).

Combining (24.142) and (24.143) with (22.141), we obtain a quantum mechanical analogue of the *Lorentz equation*,

$$\frac{d}{dt}\left[\frac{1}{2}\left(\mathbf{v}\frac{H + e\phi}{c^2} + \frac{H + e\phi}{c^2}\mathbf{v}\right)\right] = -e\left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B}\right) \tag{24.144}$$

If expectation values are taken, $H + e\phi$ can be replaced in the lowest approximation by mc^2 or $-mc^2$, depending on whether the state is made up of positive or negative energy solutions of the Dirac equation. Effectively, therefore,

$$m\frac{d\mathbf{v}}{dt} \approx \mp e\left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B}\right) \tag{24.145}$$

The upper sign corresponds to a single electron with charge $-e$ and positive energy moving in the given external field. The lower sign corresponds to the motion of a single *hole* in the "sea of positrons," represented by the electron "vacuum" state $|\mathbf{0}e\rangle$, in the presence of the external field. Such a hole is equivalent to a particle of the opposite charge, i.e., an electron, with negative mass or energy, and Eq. (24.145) is consistent with this interpretation.

More insight into the significance of various operators can be gained if the particle is moving freely, so that we may choose $\mathbf{A} = 0$ and $\phi = 0$ everywhere. The

[10]The Foldy-Wouthuysen transformation (see note 3) sheds light on the pecularities that are encountered in quantum mechanics when nonrelativistic concepts, such as position and velocity of a particle, are generalized to the relativistic regime.

Heisenberg equations of motion may then be solved explicitly. The equation of motion (24.142) reduces to

$$\frac{d\mathbf{p}}{dt} = 0 \quad \text{and} \quad \mathbf{p} = \text{const.}$$

The free particle Hamiltonian is

$$H = c\boldsymbol{\alpha} \cdot \mathbf{p} + \beta mc^2$$

and the equation of motion for the operator $\boldsymbol{\alpha}$ becomes

$$\frac{d\boldsymbol{\alpha}}{dt} = \frac{1}{i\hbar}[\boldsymbol{\alpha}, H] = \frac{2}{i\hbar}(c\mathbf{p} - H\boldsymbol{\alpha}) \tag{24.146}$$

Since $H = \text{const.}$, this equation has a simple solution:

$$\frac{\mathbf{v}(t)}{c} = \boldsymbol{\alpha}(t) = cH^{-1}\,\mathbf{p} + e^{2(i/\hbar)Ht}[\boldsymbol{\alpha}(0) - cH^{-1}\mathbf{p}] \tag{24.147}$$

The last equation can be integrated:

$$\mathbf{r}(t) = \mathbf{r}(0) + \frac{c^2\mathbf{p}}{H}t + \frac{\hbar c}{2iH}(e^{2(i/\hbar)Ht} - 1)\left[\boldsymbol{\alpha}(0) - \frac{c\mathbf{p}}{H}\right] \tag{24.148}$$

The first two terms on the right-hand side describe simply the uniform motion of a free particle. The last term is a feature of relativistic quantum mechanics and connotes a high-frequency vibration (''*Zitterbewegung*'') of the particle with frequency $\approx mc^2/\hbar$ and amplitude $\hbar/mc$, the Compton wavelength of the particle. Since for a free particle, as a special case of (24.143),

$$(\boldsymbol{\alpha} - c\mathbf{p}H^{-1})H + H(\boldsymbol{\alpha} - c\mathbf{p}H^{-1}) = 0 \tag{24.149}$$

in a representation based on momentum and energy, the operator $\boldsymbol{\alpha} - c\mathbf{p}H^{-1}$ has nonvanishing matrix elements only between states of equal momentum and *opposite* energies. Thus, the last term in Eqs. (24.147) and (24.148) is intimately connected with the appearance of the negative energy states in a relativistic theory that simultaneously describes particles and antiparticles.[11]

It is of interest to note a few further operator identities. For a free particle Hamiltonian,

$$\beta H + H\beta = 2mc^2 \tag{24.150}$$

and

$$\gamma^5 H + H\gamma^5 = 2c\boldsymbol{\Sigma} \cdot \mathbf{p} \tag{24.151}$$

Hence, in a state of energy E, the operator β has the expectation value

$$\langle\beta\rangle = \frac{mc^2}{E} = \pm\sqrt{1 - \frac{c^2p^2}{E^2}} \tag{24.152}$$

so that $\langle\beta\rangle$ approaches ± 1 in the nonrelativistic approximation and vanishes as the speed of light is approached.

[11]Sakurai (1967), pp. 117–119 and 139–140.

Similarly,

$$\langle \gamma^5 \rangle = \langle \boldsymbol{\Sigma} \cdot \mathbf{p} \rangle \frac{c}{E} \tag{24.153}$$

showing that the expectation value of γ^5 is v/c times the expectation value of the helicity operator $\boldsymbol{\Sigma} \cdot \hat{\mathbf{p}}$. The operator γ^5 is called the *chirality*.

Exercise 24.39. Verify Eqs. (24.150)–(24.153) for free relativistic electrons. Prove that the helicity $\boldsymbol{\Sigma} \cdot \hat{\mathbf{p}}$ is conserved in free particle dynamics, but that the chirality is conserved only if the Dirac particles have zero mass. (This result has led to the term *approximate chiral symmetry* for a theory of a Dirac particle whose mass is neglected.)

The role of the spin in the one-electron Dirac theory is brought into focus if we evaluate the time derivative of $\boldsymbol{\Sigma}$ for an electron exposed to a vector potential $\mathbf{A}$ but no potential ϕ, so that

$$H = c\boldsymbol{\alpha} \cdot \left(\mathbf{p} + \frac{e}{c} \mathbf{A} \right) + \beta mc^2 \tag{24.154}$$

By a sequence of algebraic manipulations, we obtain

$$\frac{d\boldsymbol{\Sigma}}{dt} = \frac{1}{i\hbar} [\boldsymbol{\Sigma}, H] = -\frac{2c}{\hbar} \boldsymbol{\alpha} \times \left(\mathbf{p} + \frac{e}{c} \mathbf{A} \right) \tag{24.155}$$

and

$$H \frac{d\boldsymbol{\Sigma}}{dt} + \frac{d\boldsymbol{\Sigma}}{dt} H = -\frac{2c^2}{\hbar} \left\{ \left[\boldsymbol{\alpha} \cdot \left(\mathbf{p} + \frac{e}{c} \mathbf{A} \right) \right] \left[\boldsymbol{\alpha} \times \left(\mathbf{p} + \frac{e}{c} \mathbf{A} \right) \right] + \left[\boldsymbol{\alpha} \times \left(\mathbf{p} + \frac{e}{c} \mathbf{A} \right) \right] \left[\boldsymbol{\alpha} \cdot \left(\mathbf{p} + \frac{e}{c} \mathbf{A} \right) \right] \right\}$$

The contents of the brace on the right-hand side of the last equation may be reduced to $(e\hbar/c)\boldsymbol{\Sigma} \times (\nabla \times \mathbf{A}) = (e\hbar/c)\boldsymbol{\Sigma} \times \mathbf{B}$. Hence, the simple relation

$$\boxed{H \frac{d\boldsymbol{\Sigma}}{dt} + \frac{d\boldsymbol{\Sigma}}{dt} H = -2ec\boldsymbol{\Sigma} \times \mathbf{B}} \tag{24.156}$$

is valid. In the nonrelativistic approximation $H \approx mc^2$, this equation becomes the equation of motion for the one-electron spin operator $\mathbf{S} = \frac{\hbar}{2} \boldsymbol{\Sigma}$,

$$\boxed{\frac{d\mathbf{S}}{dt} \approx -\frac{e}{mc} \mathbf{S} \times \mathbf{B}} \tag{24.157}$$

A straightforward interpretation of this equation may be given: The time rate of change of intrinsic angular momentum (spin) equals the torque produced by the applied magnetic field. If a magnetic moment $\boldsymbol{\mu}$ is associated with the spin, the torque is $\boldsymbol{\mu} \times \mathbf{B}$. Comparison with (24.157) shows that in this approximation the magnetic moment operator for an electron is

$$\boldsymbol{\mu} = -\frac{e}{mc} \mathbf{S} = -\frac{e}{2mc} g_s \mathbf{S} \tag{24.158}$$

with g_s defined in Eq. (17.101). Except for small radiative corrections, the value $g_s = 2$, derived here from relativistic quantum mechanics for a charged Dirac particle is in agreement with the experimental measurements.[12]

8. ***Dirac Theory in the Schrödinger Picture and the Nonrelativistic Limit.*** In this section we return to the *Schrödinger picture* for the one-particle Dirac equation:

$$\gamma^\mu\left(\frac{\partial}{\partial x^\mu} - \frac{ie}{\hbar c} A_\mu\right)\psi + i\kappa\psi = 0 \tag{24.159}$$

A convenient second-order differential equation for ψ is obtained by iterating this equation as follows:

$$\left[\gamma^\mu\left(\frac{\partial}{\partial x^\mu} - \frac{ie}{\hbar c} A_\mu\right) - i\kappa\right]\left[\gamma^\nu\left(\frac{\partial}{\partial x^\nu} - \frac{ie}{\hbar c} A_\nu\right) + i\kappa\right]\psi = 0 \tag{24.160}$$

whence

$$\left[\gamma^\mu\gamma^\nu\left(\frac{\partial}{\partial x^\mu} - \frac{ie}{\hbar c} A_\mu\right)\left(\frac{\partial}{\partial x^\nu} - \frac{ie}{\hbar c} A_\nu\right) + \kappa^2\right]\psi = 0$$

or, separating the terms with $\mu = \nu$ from those with $\mu \neq \nu$,

$$\left[-\left(\nabla + \frac{ie}{\nu c}\mathbf{A}\right)^2 + \frac{1}{c^2}\left(\frac{\partial}{\partial t} - \frac{ie}{\hbar}\phi\right)^2 + \frac{e}{\hbar c}\boldsymbol{\Sigma}\cdot\mathbf{B} - \frac{ie}{\hbar c}\boldsymbol{\alpha}\cdot\mathbf{E} + \kappa^2\right]\psi = 0 \tag{24.161}$$

In more elegant, but also more recondite, relativistic notation:

$$\left[\left(\frac{\partial}{\partial x_\mu} - \frac{ie}{\hbar c} A^\mu\right)\left(\frac{\partial}{\partial x^\mu} - \frac{ie}{\hbar c} A_\mu\right) + \frac{e}{2\hbar c}\Sigma^{\mu\nu}F_{\mu\nu} + \kappa^2\right]\psi = 0 \tag{24.162}$$

where

$$F_{\mu\nu} = \frac{\partial A_\nu}{\partial x^\mu} - \frac{\partial A_\mu}{\partial x^\mu}$$

is the electromagnetic field tensor.

Exercise 24.40. Derive Eqs. (24.161) and (24.162).

In the absence of an electromagnetic field, the second-order equation (24.162) reduces to the *Klein-Gordon equation*,

$$\left(\frac{\partial}{\partial x_\mu}\frac{\partial}{\partial x^\mu} + \kappa^2\right)\psi \equiv \left(-\nabla^2 + \frac{1}{c^2}\frac{\partial^2}{\partial t^2} + \kappa^2\right)\psi = 0 \tag{24.163}$$

which also governs the wave function for a relativistic particle with spin zero. In the presence of the external field, Eq. (24.162) differs from the relativistic Schrödinger equation for a scalar particle by the terms containing the Dirac matrices and coupling the wave function directly to the electromagnetic field **B** and **E**.

The second-order equation (24.160) has more solutions than the Dirac equation, from which it was obtained by iteration; it is therefore necessary to select among its solutions χ only those that also satisfy the Dirac equation. A convenient method, which also has physical significance, is to classify the four-component solutions of

[12]Quantum electrodynamics gives the value $g_s = 2(1 + \alpha/2\pi)$ to first order in the fine structure constant. See Sakurai (1967).

the second-order equation according to their chirality. Since the chirality γ^5 is Hermitian and anticommutes with all γ^μ, the solutions of Eq. (24.160) can be assumed to be simultaneously eigenspinors of γ^5 with eigenvalues $+1$ or -1, so that

$$\gamma^5\psi^{(\pm)} = \pm\chi^{(\pm)} \tag{24.164}$$

We call such solutions *chiral*. Since (24.160) can be written as

$$\frac{i}{\kappa}\gamma^\mu\left(\frac{\partial}{\partial x^\mu} - \frac{ie}{\hbar c}A_\mu\right)\frac{i}{\kappa}\gamma^\nu\left(\frac{\partial}{\partial x^\nu} - \frac{ie}{\hbar c}A_\nu\right)\chi = \chi$$

it follows that the chiral solutions are paired by the reciprocal relation

$$\chi^{(\mp)} = \frac{i}{\kappa}\gamma^\mu\left(\frac{\partial}{\partial x^\mu} - \frac{ie}{\hbar c}A_\mu\right)\chi^{(\pm)} \tag{24.165}$$

Each chiral pair generates a "self-paired" solution of the Dirac equation:

$$\psi = \chi^{(+)} + \chi^{(-)} \tag{24.166}$$

Exercise 24.41. Prove that $(1 \pm \gamma^5)/2$ is the chiral projection operator and that any solution of the Dirac equation can be uniquely decomposed into a superposition of two solutions of the second-order equation (24.160) with opposite chirality, but that in general the two chiral solutions are not paired as in (24.165).

Exercise 24.42. Show that chirality is conserved and that the Dirac theory exhibits *chiral symmetry* if the particle mass is zero. (See also Exercise 24.39.)

In order to interpret Eq. (24.161), we assume that the external field is time-independent and we consider a stationary state solution

$$\psi(\mathbf{r}, t) = e^{-(i/\hbar)Et}u(\mathbf{r})$$

Substitution into Eq. (24.161) gives

$$\left[\frac{1}{c^2}(E + e\phi)^2 - \left(\frac{\hbar}{i}\nabla + \frac{e}{c}\mathbf{A}\right)^2 - \frac{e\hbar}{c}(\boldsymbol{\Sigma}\cdot\mathbf{B} - i\boldsymbol{\alpha}\cdot\mathbf{E}) - m^2c^2\right]u = 0$$

This equation is still exact. For a nonrelativistic electron, for which $E \approx mc^2$, we approximate

$$\frac{1}{c^2}(E + e\phi)^2 - m^2c^2 \approx 2m(E + e\phi - mc^2)$$

Hence, we obtain

$$(E + e\phi - mc^2)u = \left[\frac{1}{2m}\left(\frac{\hbar}{i}\nabla + \frac{e}{c}\mathbf{A}\right)^2 + \frac{e\hbar}{2mc}(\boldsymbol{\Sigma}\cdot\mathbf{B} - i\boldsymbol{\alpha}\cdot\mathbf{E})\right]u \tag{24.167}$$

which is very similar to the nonrelativistic Schrödinger equation. In the absence of an electric field, this equation describes the motion of the electron in an external magnetic field and again shows that an intrinsic magnetic moment as given by (24.158) must be ascribed to the electron.

The physical appreciation of the Dirac theory is further aided by rewriting the conserved four-vector current density

$$j^\mu = -ec\bar{\psi}\gamma^\mu\psi \tag{24.168}$$

in terms of operators that have a nonrelativistic interpretation. To this end, we write the Dirac equations for ψ and $\bar{\psi}$ as

$$\psi = \frac{i}{\kappa}\gamma^\lambda\left(\frac{\partial}{\partial x^\lambda} - \frac{ie}{\hbar c}A_\lambda\right)\psi \tag{24.169}$$

$$\bar{\psi} = -\frac{i}{\kappa}\left(\frac{\partial}{\partial x^\lambda} + \frac{ie}{\hbar c}A_\lambda\right)\bar{\psi}\gamma^\lambda \tag{24.170}$$

Substituting (24.169) in one-half of the current density and (24.170) in the other half, we get

$$j^\mu = -\frac{eci}{2\kappa}\left(\bar{\psi}\gamma^\mu\gamma^\lambda\frac{\partial\psi}{\partial x^\lambda} - \frac{\partial\bar{\psi}}{\partial x^\lambda}\gamma^\lambda\gamma^\mu\psi\right)_{\lambda\neq\mu}$$
$$-\frac{eci}{2\kappa}g^{\mu\lambda}\left\{\bar{\psi}\left(\frac{\partial}{\partial x^\lambda} - \frac{ie}{\hbar c}A_\lambda\right)\psi - \left[\left(\frac{\partial}{\partial x^\lambda} + \frac{ie}{\hbar c}A_\lambda\right)\bar{\psi}\right]\psi\right\}$$

where the terms with $\mu \neq \lambda$ have been separated from those with $\mu = \lambda$, and the anticommutation relation (24.41) is used. With the definition (24.66) of $\Sigma^{\mu\nu}$, j^μ is finally transformed into a sum of *polarization* and *convection* terms,

$$j^\mu = j^\mu_{\text{pol}} + j^\mu_{\text{conv}} \tag{24.171}$$

where

$$j^\mu_{\text{pol}} = \frac{ec}{2\kappa}\frac{\partial}{\partial x^\lambda}(\bar{\psi}\,\Sigma^{\lambda\mu}\psi) \tag{24.172}$$

and

$$j^\mu_{\text{conv}} = -\frac{eci}{2\kappa}g^{\mu\lambda}\left\{\bar{\psi}\left(\frac{\partial}{\partial x^\lambda} - \frac{ie}{\hbar c}A_\lambda\right)\psi - \left[\left(\frac{\partial}{\partial x^\lambda} + \frac{ie}{\hbar c}A_\lambda\right)\bar{\psi}\right]\psi\right\} \tag{24.173}$$

This procedure is known as the *Gordon decomposition* of the current density. Note that the electromagnetic potential, which is absent from the expression (24.168) for the current, now appears explicitly in the convection current.

Exercise 24.43. Prove that both the polarization and convection currents are separately conserved, and show the relation of the convection current to the nonrelativistic current density (Exercise 4.17). Evaluate the polarization and convection currents for the free particle plane wave states.

Exercise 24.44. Show that, as in (3.7), a conserved *transition current density* can be defined to be

$$j^\mu_{12} = -ec\bar{\psi}_1\gamma^\mu\psi_2 \tag{24.174}$$

where both ψ_1 and ψ_2 are solutions of the same Dirac equation.

9. ***Central Forces and the Hydrogen Atom.*** In an electrostatic central field, the one-particle Dirac Hamiltonian for an election is

$$H = c\boldsymbol{\alpha}\cdot\frac{\hbar}{i}\nabla + \beta mc^2 - e\phi(r) \tag{24.175}$$

Since **L** and **S** do not separately commute with the free-particle Hamiltonian, they certainly will not commute with the Hamiltonian (24.175) either. However, the components of the total angular momentum,

$$\boxed{\mathbf{J} = \mathbf{r} \times \frac{\hbar}{i} \nabla + \frac{\hbar}{2} \boldsymbol{\Sigma}} \tag{24.176}$$

evidently do commute with H, and we may therefore seek to find simultaneous eigenspinors of H, $\mathbf{J}^2$, and J_z. Parity will join these as a useful constant of the motion.

At this point, it is convenient (though not unavoidable) to introduce the standard representation (23.32) of the Dirac matrices and to write all equations in two-component form. We introduce two two-component spinors φ_1 and φ_2 by the definition

$$\psi = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix} \tag{24.177}$$

The Dirac equation decomposes according to Exercise 24.11 into the coupled equations

$$(E - mc^2 + e\phi)\varphi_1 - \frac{\hbar}{i} c\boldsymbol{\sigma} \cdot \nabla \varphi_2 = 0 \tag{24.178}$$

$$(E + mc^2 + e\phi)\varphi_2 - \frac{\hbar}{i} c\boldsymbol{\sigma} \cdot \nabla \varphi_1 = 0 \tag{24.179}$$

where all $\boldsymbol{\sigma}$ are 2×2 Pauli matrices. The operators J_z and

$$\mathbf{J}^2 = \mathbf{L}^2 + \hbar \mathbf{L} \cdot \boldsymbol{\Sigma} + \frac{3}{4} \hbar^2$$

decompose similarly, and it is clear that we must seek to make both φ_1 and φ_2 two-component eigenspinors satisfying the conditions

$$\left(L_z + \frac{\hbar}{2} \sigma_z\right) \varphi_{1,2} = m\hbar \varphi_{1,2} \tag{24.180}$$

and

$$\left(\mathbf{L}^2 + \hbar \mathbf{L} \cdot \boldsymbol{\sigma} + \frac{3}{4} \hbar^2\right) \varphi_{1,2} = j(j + 1)\hbar^2 \varphi_{1,2} \tag{24.181}$$

Equation (17.64) contains the answer to this problem and shows that for a given value of j the spinors φ_1 and φ_2 must be proportional to $\mathcal{Y}^{jm}_{j\mp(1/2)}$. The two-component spinors $\mathcal{Y}$ are normalized as $\int \mathcal{Y}^\dagger \mathcal{Y} d\Omega = 1$, and they have the useful property

$$\boldsymbol{\sigma} \cdot \hat{\mathbf{r}}\, \mathcal{Y}^{jm}_{j \mp 1/2} = -\mathcal{Y}^{jm}_{j \pm 1/2} \tag{24.182}$$

Exercise 24.45. Prove Eq. (24.182) by using the fact that $\boldsymbol{\sigma} \cdot \hat{\mathbf{r}}$ is a pseudoscalar under rotation and that $\mathcal{Y}$ has a simple value for $\theta = 0$.

Exercise 24.46. Prove that

$$\boldsymbol{\sigma} \cdot \mathbf{L} \mathcal{Y}^{jm}_{j-1/2} = \left(j - \frac{1}{2}\right) \hbar \mathcal{Y}^{jm}_{j-1/2} \tag{24.183}$$

$$\boldsymbol{\sigma} \cdot \mathbf{L} \mathcal{Y}^{jm}_{j+1/2} = \left(-j - \frac{3}{2}\right) \hbar \mathcal{Y}^{jm}_{j+1/2} \tag{24.184}$$

Since the *parity operator* also commutes with the other available constants of the motion (energy and angular momentum), it may be chosen as further "good quantum number" and we may require that

$$\beta\begin{pmatrix}\varphi_1(-\mathbf{r})\\ \varphi_2(-\mathbf{r})\end{pmatrix} = \pm\begin{pmatrix}\varphi_1(\mathbf{r})\\ \varphi_2(\mathbf{r})\end{pmatrix} \qquad \begin{matrix}\text{(even)}\\ \text{(odd)}\end{matrix}$$

The parity of the eigenfunction clearly dictates how the spinors $\mathcal{Y}^{jm}_{j\pm 1/2}$ are associated with φ_1 and φ_2. It is seen that the two solutions must have the form

$$\psi = \begin{pmatrix} F(r)\mathcal{Y}^{jm}_{j-1/2}\\ -if(r)\mathcal{Y}^{jm}_{j+1/2}\end{pmatrix} \tag{24.185}$$

or

$$\psi = \begin{pmatrix} G(r)\mathcal{Y}^{jm}_{j+1/2}\\ -ig(r)\mathcal{Y}^{jm}_{j-1/2}\end{pmatrix} \tag{24.186}$$

Equation (24.185) has even or odd parity depending on the parity of $j - 1/2$. Equation (24.186) has even or odd parity depending on the parity of $j + 1/2$. The factor $-i$ have been introduced so that the radial equations will be real.

In order to derive the radial equations, we employ the following identity

$$\boldsymbol{\sigma}\cdot\mathbf{p} = \frac{1}{r^2}(\boldsymbol{\sigma}\cdot\mathbf{r})(\boldsymbol{\sigma}\cdot\mathbf{r})(\boldsymbol{\sigma}\cdot\mathbf{p}) = \boldsymbol{\sigma}\cdot\hat{\mathbf{r}}\left(\hat{\mathbf{r}}\cdot\mathbf{p} + \frac{i\boldsymbol{\sigma}\cdot\mathbf{L}}{r}\right) \tag{24.187}$$

which follows from (16.59). Here,

$$\hat{\mathbf{r}}\cdot\mathbf{p} = \frac{\hbar}{i}\frac{\partial}{\partial r}$$

If we substitute the last two relations into Eqs. (24.178) and (24.179), and take Eqs. (24.182), (24.183), and (24.184) into account, we obtain the coupled radial equations,

$$(E - mc^2 + e\phi)F - \hbar c\left(\frac{d}{dr} + \frac{j + 3/2}{r}\right)f = 0 \tag{24.188}$$

$$(E + mc^2 + e\phi)\,f + \hbar c\left(\frac{d}{dr} - \frac{j - 1/2}{r}\right)F = 0 \tag{24.189}$$

and

$$(E - mc^2 + e\phi)G - \hbar c\left(\frac{d}{dr} - \frac{j - 1/2}{r}\right)g = 0 \tag{24.190}$$

$$(E + mc^2 + e\phi)g + \hbar c\left(\frac{d}{dr} + \frac{j + 3/2}{r}\right)G = 0 \tag{24.191}$$

So far, it has not been necessary to introduce the explicit form of the potential, but at this point we assume that the electron moves in the Coulomb field of a point nucleus of charge Ze:

$$\phi(r) = \frac{Ze}{r} \tag{24.192}$$

We also define the dimensionless quantities,

$$\lambda = j + \frac{1}{2}, \qquad \frac{E}{mc^2} = \varepsilon, \qquad \frac{mcr}{\hbar} = x, \qquad \frac{e^2}{\hbar c} = \alpha \tag{24.193}$$

The coupled radial equations then become

$$\left(\varepsilon - 1 + \frac{Z\alpha}{x}\right)F - \left(\frac{d}{dx} + \frac{\lambda + 1}{x}\right)f = 0 \tag{24.194}$$

$$\left(\varepsilon + 1 + \frac{Z\alpha}{x}\right)f + \left(\frac{d}{dx} - \frac{\lambda - 1}{x}\right)F = 0 \tag{24.195}$$

and

$$\left(\varepsilon - 1 + \frac{Z\alpha}{x}\right)G - \left(\frac{d}{dx} - \frac{\lambda - 1}{x}\right)g = 0 \tag{24.196}$$

$$\left(\varepsilon + 1 + \frac{Z\alpha}{x}\right)g + \left(\frac{d}{dx} + \frac{\lambda + 1}{x}\right)G = 0 \tag{24.197}$$

Since these two sets of equations are obtained from one another by the transformation

$$F \to G, \qquad f \to g, \qquad \lambda \to -\lambda$$

it suffices to consider Eqs. (24.194) and (24.195).

The analysis of the radial equations proceeds as usual. Asymptotically, for $x \to \infty$ we find the behavior

$$F \cong f \cong e^{\pm\sqrt{1-\varepsilon^2}x}$$

For *bound states*, to which we confine our attention, we must require $|\varepsilon| \leq 1$ and choose the minus sign in the exponent. With the ansatz

$$F = e^{-\sqrt{1-\varepsilon^2}\,x} x^\gamma \sum_{\nu=0} a_\nu x^\nu \tag{24.198}$$

$$f = e^{-\sqrt{1-\varepsilon^2}\,x} x^\gamma \sum_{\rho=0} b_\rho x^\rho \tag{24.199}$$

we obtain by substitution, for $\nu > 0$,

$$(\varepsilon - 1)a_{\nu-1} + Z\alpha a_\nu + \sqrt{1-\varepsilon^2}\,b_{\nu-1} - (\lambda + 1 + \gamma + \nu)b_\nu = 0 \tag{24.200}$$

and

$$(\varepsilon + 1)b_{\nu-1} + Z\alpha b_\nu - \sqrt{1-\varepsilon^2}\,a_{\nu-1} + (-\lambda + 1 + \gamma + \nu)a_\nu = 0 \tag{24.201}$$

as well as

$$Z\alpha a_0 - (\lambda + 1 + \gamma)b_0 = 0, \qquad (-\lambda + 1 + \gamma)a_0 + Z\alpha b_0 = 0 \tag{24.202}$$

Equations (24.202) are compatible only if

$$(Z\alpha)^2 = \lambda^2 - (\gamma + 1)^2$$

Hence,

$$\gamma = -1 \pm\sqrt{\lambda^2 - (Z\alpha)^2} = -1 \pm\sqrt{(j + 1/2)^2 - (Z\alpha)^2}$$

The root with the minus sign must be excluded because the corresponding wave function would be too singular at the origin to be admissible. Hence,

$$\boxed{\gamma = -1 + \sqrt{(j + 1/2)^2 - (Z\alpha)^2}} \tag{24.203}$$

Provided that $Z < 1/\alpha \approx 137$, this will yield a real value for γ. For $j = 1/2$ we have $-1 < \gamma < 0$ and the wave function is mildly singular at the origin, but not enough to disturb its quadratic integrability. (See Section 12.4.)

The usual argument can now be made to show that both power series (24.198) and (24.199) must terminate at $x^{n'}$ (see Chapter 12).

Exercise 24.47. Carry out the study of the asymptotic behavior of (24.198) and (24.199), and show that the power series must terminate.

From the recursion relations (24.200) and (24.201), we then obtain for $\nu = n' + 1$ (with $a_{n'+1} = b_{n'+1} = 0$),

$$b_{n'} = \sqrt{\frac{1 - \varepsilon}{1 + \varepsilon}}\, a_{n'} \tag{24.204}$$

From Eqs. (24.200) and (24.201), we may simultaneously eliminate $a_{\nu-1}$ and $b_{\nu-1}$ to get

$$\begin{aligned} a_\nu[Z\alpha\sqrt{1+\varepsilon} + (\lambda - 1 - \gamma - \nu)\sqrt{1-\varepsilon}] \\ = b_\nu[Z\alpha\sqrt{1-\varepsilon} + (\lambda + 1 + \gamma + \nu)\sqrt{1+\varepsilon}] \end{aligned} \tag{24.205}$$

Letting $\nu = n'$ and comparing with Eqs. (24.204), we finally conclude that

$$\sqrt{\frac{1-\varepsilon}{1+\varepsilon}} = \frac{Z\alpha\sqrt{1+\varepsilon} + (\lambda - 1 - \gamma - n')\sqrt{1-\varepsilon}}{Z\alpha\sqrt{1-\varepsilon} + (\lambda + 1 + \gamma + n')\sqrt{1+\varepsilon}}$$

or

$$\sqrt{1-\varepsilon^2}\,(1 + \gamma + n') = Z\alpha\varepsilon$$

This condition translates into the formula for the energy:

$$\boxed{\frac{E}{mc^2} = \frac{1}{\sqrt{1 + \dfrac{(Z\alpha)^2}{[\sqrt{(j+1/2)^2 - (Z\alpha)^2} + n']^2}}}} \tag{24.206}$$

This is the famous *fine structure formula* for the hydrogen atom. The quantum numbers j and n' assume the values

$$j = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \ldots; \qquad n' = 0, 1, 2, \ldots$$

The principal quantum number n of the nonrelativistic theory of the hydrogen atom is related to n' and j by

$$n = j + 1/2 + n' \tag{24.207}$$

From (24.204) we have for $n' = 0$

$$\sqrt{1-\varepsilon}\, a_0 - \sqrt{1+\varepsilon}\, b_0 = 0$$

This relation between a_0 and b_0 is consistent with (24.202) only if $\lambda + 1 + \gamma > 0$, or $\lambda > -\sqrt{\lambda^2 - (Z\alpha)^2}$, hence $\lambda > 0$. The transformation $\lambda \to -\lambda$, which takes us from the states with $j = \ell + 1/2$ to those with $j = \ell - 1/2$, is therefore not permissible if $n' = 0$, and a solution of type (24.186) is not possible if $n = j + 1/2$. Hence, for a given value of the principal quantum number n, there is only one state

with $j = n - 1/2$, while there are two states of opposite parities for all $j < n - 1/2$. Since $\lambda = j + 1/2$ appears squared in the energy formula, pairs of states with the same j but opposite parities (e.g., $2S_{1/2}$ and $2P_{1/2}$) remain degenerate in the one-electron Dirac theory. Experiments have substantiated this formula and its radiative corrections (the Lamb shift that removes the $2S_{1/2} - 2P_{1/2}$ degeneracy, displacing the S state slightly above P) to very high accuracy, impressively vindicating the one-electron approximation of the full relativistic electron-positron theory. Figure 24.1 shows the fine structure of the energy levels with $n = 2$ and 3 of the hydrogen atom.

Exercise 24.48. Expand the relativistic energy of a hydrogenic atom in powers of $(Z\alpha)^2$ to obtain the Bohr-Balmer formula and the first correction to the nonrelativistic energies. For hydrogen ($Z = 1$), compute the energies of the ground state, the $2S_{1/2}$ metastable state, and the $2P$ doublet (see Figure 24.1).

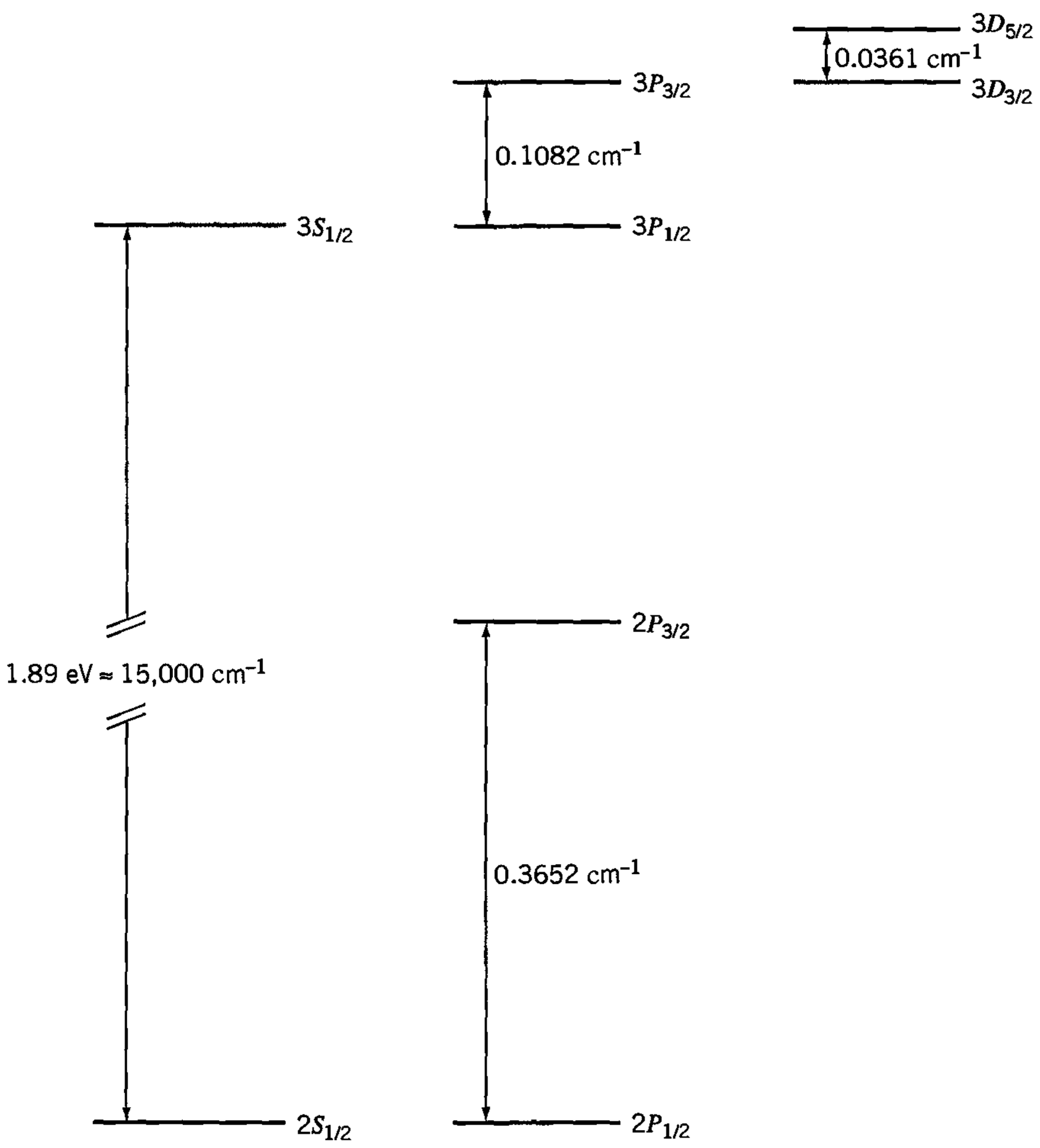


Figure 24.1. Detail of an energy-level diagram for the hydrogen atom. The manifolds of the $n = 2$ and 3 levels are shown, based on the Dirac theory, without radiative corrections (Lamb shift) or hyperfine splittings. The energy differences are given in units of cm^{-1} for the reciprocal wavelength. The Lamb shift displaces S levels upward in energy by about 10 percent of the fine structure splitting for the manifold.

Problems

1. If **A** and **B** are proportional to the unit 4×4 matrix, derive expansion formulas for the matrix products $(\boldsymbol{\alpha} \cdot \mathbf{A})(\boldsymbol{\alpha} \cdot \mathbf{B})$ and $(\boldsymbol{\alpha} \cdot \mathbf{A})(\boldsymbol{\Sigma} \cdot \mathbf{B})$ in terms of α and $\boldsymbol{\Sigma}$ matrices in analogy with formula (16.59).
2. If a field theory of massless spin one-half particles (neutrinos) is developed, so that the β matrix is absent, show that the conditions (24.30) and (24.31) are solved by 2×2 Pauli matrices, $\boldsymbol{\alpha} = \pm\boldsymbol{\sigma}$. Work out the details of the resulting *two-component* theory with particular attention to the helicity properties. Is this theory invariant under spatial reflection?
3. Develop the outlines of relativistic quantum field theory for neutral spinless bosons with mass. What modifications are indicated when the particles are charged?
4. Show that the vector operator

$$\mathbf{Q} = \beta\boldsymbol{\Sigma} + (1 - \beta)\boldsymbol{\Sigma} \cdot \hat{\mathbf{p}}\hat{\mathbf{p}}$$

satisfies the same commutation relations as $\boldsymbol{\Sigma}$ and that it commutes with the free Dirac particle Hamiltonian. Show that the eigenvalues of any component of **Q** are ± 1.

Apply the unitary transformation

$$\exp\left[i(\theta/2)\left(-p_y Q_x + p_x Q_y\right)/\sqrt{p_x^2 + p_y^2}\right]$$

to the spinors (24.92) and (24.93), and prove that the resulting spinors are eigenstates of H with sharp momentum and definite value of Q_z. Show that these states are the relativistic analogues of the nonrelativistic momentum eigenstates with "spin up" and "spin down."

5. Assume that the potential energy $-\mathbf{e}\phi(\mathbf{r})$ in the Dirac Hamiltonian (24.175) is a square well of depth V_0 and radius a. Determine the continuity condition for the Dirac wave function ψ at $r = a$, and derive a transcendental equation for the minimum value of V_0 which just binds a particle of mass m for a given value of a.
6. Solve the relativistic Schrödinger equation for a spinless particle of mass m and charge $-e$ in the presence of the Coulomb field of a point nucleus with charge Ze. Compare the fine structure of the energy levels with the corresponding results for the Dirac electron.
7. Consider a neutral spin one-half Dirac particle with mass and with an intrinsic magnetic moment, and assume the Hamiltonian

$$H = c\boldsymbol{\alpha} \cdot \frac{\hbar}{i}\nabla + \beta mc^2 + \lambda B\beta\Sigma_z$$

in the presence of a uniform constant magnetic field along the z axis. Determine the important constants of the motion, and derive the energy eigenvalues. Show that orbital and spin motions are coupled in the relativistic theory but decoupled in a nonrelativistic limit. The coefficient λ is a constant, proportional to the gyromagnetic ratio.
8. If a Dirac electron is moving in a uniform constant magnetic field pointing along the z axis, determine the energy eigenvalues and eigenspinors.

Appendix

The Appendix is a compilation of mathematical accessories, definitions, conventions, and mnemonics that are applicable in quantum mechanics. Instead of mathematical rigor, ease of use is the objective.

1. ***Fourier Analysis and Delta Functions.*** We first consider the generally complex-valued, periodic functions defined on the real x axis,

$$f(x + L) = f(x) \tag{A.1}$$

which can be expanded in terms of the Fourier series

$$f(x) = \frac{1}{\sqrt{L}} \sum_{n=-\infty}^{+\infty} c_n e^{2\pi nix/L} = \frac{1}{\sqrt{L}} \sum_{n=-\infty}^{+\infty} c_n e^{2\pi nix/L}\, \Delta n \tag{A.2}$$

Most functions of interest in quantum mechanics are or can be approximated by functions that are in this category. On the right-hand side, $\Delta n = 1$ is redundant, but by inserting it we are preparing for the transition from Fourier series to Fourier integrals. The Fourier coefficients are calculated from

$$c_n = \frac{1}{\sqrt{L}} \int_{-L/2}^{L/2} f(u) e^{-2\pi niu/L}\, du \tag{A.3}$$

where the integration interval $-L/2 \leq x \leq L/2$ has been chosen for convenience. (Show that any other interval of length L would give the same coefficients.) Substitution of (A.3) into (A.2) gives the identity

$$f(x) = \frac{1}{L} \sum_{n=-\infty}^{+\infty} \int_{-L/2}^{L/2} e^{(2\pi ni/L)(x-u)} f(u)\, du = \int_{-L/2}^{L/2} du\, f(u) \frac{1}{L} \sum_{n=-\infty}^{+\infty} e^{(2\pi ni/L)(x-u)} \tag{A.4}$$

if the exchange of integral and summation is permissible.

By taking the limit $L \to \infty$ and turning the Fourier series into an integral by the transformation

$$\frac{2\pi n}{L} \to k \qquad \frac{2\pi \Delta n}{L} \to dk \qquad \sqrt{\frac{L}{2\pi}}\, c_n \to g(k) \tag{A.5}$$

the reciprocal Fourier integral formulas are obtained from (A.2) and (A.3):

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} g(k) e^{ikx}\, dk \qquad g(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(u) e^{-iku}\, du \tag{A.6}$$

for functions f and g defined over the entire real axis. The identity (A.4) now becomes

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk \int_{-\infty}^{+\infty} e^{ik(x-u)} f(u)\, du = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(u)\, du \int_{-\infty}^{+\infty} e^{ik(u-x)}\, dk \tag{A.7}$$

Since for a fixed x we can change the value of f in the integrand almost everywhere (except near $u = x$) without affecting the value $f(x)$, (A.7) represents the delta function

$$\delta(x - u) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ik(x-u)}\, dk \tag{A.8}$$

with the property of being an even function of its argument,

$$\delta(x - u) = \delta(u - x) \tag{A.9}$$

and

$$f(x) = \int_{-\infty}^{+\infty} \delta(u - x) f(u)\, du \tag{A.10}$$

If condition (A.10) is applied to a simple function defined such that $f(x) = 1$ for $x_1 < x < x_2$ and $f(x) = 0$ outside the interval (x_1, x_2), we see that the delta function must satisfy the test

$$\int_{x_1}^{x_2} \delta(u - x)\, du = \begin{cases} 0 & \text{if } x \text{ lies outside } (x_1, x_2) \\ 1 & \text{if } x_1 < x < x_2 \end{cases} \tag{A.11}$$

This equation may be regarded as the definition of the delta function. It is effectively zero whenever its argument differs from zero, but it is singular when its argument vanishes, and its total area is unity.

The infinite integral in (A.8) does not exist by any of the conventional definitions of an integral. Yet it is convenient to admit this equality as meaningful with the proviso that the entities which are being equated must eventually be used in conjunction with a well-behaved function $f(x)$ under an integral sign. Physically, (A.8) may be interpreted as the superposition with equal amplitudes (and equal phases at $x = u$) of simple harmonic oscillations of all frequencies. The contributions to the Fourier integral completely cancel by destructive interference unless the argument of the delta function vanishes, i.e., $x = u$.

By choosing to make the delta function (A.8) the limit of a well-defined integral, we obtain (with $u = 0$) various useful representations of $\delta(x)$:

$$\begin{aligned} \delta(x) &= \frac{1}{2\pi} \lim_{K\to\infty} \int_{-K}^{+K} e^{ikx}\, dk = \frac{1}{2\pi} \lim_{\varepsilon\to 0^+} \int_{-\infty}^{+\infty} e^{ikx - \varepsilon^2 k^2}\, dk \\ &= \frac{1}{2\pi} \lim_{\varepsilon\to 0^+} \int_{-\infty}^{+\infty} e^{ikx - \varepsilon|k|}\, dk \end{aligned} \tag{A.12}$$

and many more can be invented. (Try some.) Small (''infinitesimal'') quantities like ε are assumed to be positive everywhere in this book, without special notice. The representations (A.12) are, explicitly

$$\delta(x) = \frac{1}{\pi} \lim_{N\to\infty} \frac{\sin Nx}{x} = \frac{1}{\sqrt{\pi}} \lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon} \exp\left(-\frac{x^2}{\varepsilon^2}\right) = \frac{1}{\pi} \lim_{\varepsilon\to 0^+} \frac{\varepsilon}{x^2 + \varepsilon^2} \tag{A.13}$$

Two further useful representations are:

$$\delta(x) = \frac{1}{2\pi} \lim_{\varepsilon\to 0^+} \int_{-\infty}^{+\infty} e^{ikx} \frac{\sin k\varepsilon}{k\varepsilon}\, dk = \lim_{\varepsilon\to 0^+} \begin{cases} 1/2\varepsilon & \text{if } -\varepsilon < x < \varepsilon \\ 0 & \text{if } |x| > \varepsilon \end{cases} \tag{A.14}$$

and

$$\delta(x) = \frac{1}{2}\frac{d^2}{dx^2}|x| \tag{A.15}$$

In quantum mechanics, we also need the incomplete Fourier integral,

$$\int_{-\infty}^{T} e^{i(\omega - i\varepsilon)t}\, dt = \frac{e^{(i\omega+\varepsilon)T}}{i\omega + \varepsilon} \tag{A.16}$$

If the limits $T \to \infty$ and $\varepsilon \to 0$ are taken, but such that $\varepsilon T \to 0$, i.e., the ε-limit precedes the T-limit, (A.16) is just 2π times (A.8), or

$$\lim_{\substack{\varepsilon\to 0 \\ T\to\infty \\ (\varepsilon T\to 0)}} \frac{e^{(i\omega+\varepsilon)T}}{i\omega + \varepsilon} = 2\pi\delta(\omega) \tag{A.17}$$

If, on the other hand, we choose $T = 0$ in (A.16), we obtain

$$\begin{aligned}\int_{-\infty}^{0} e^{i(\omega - i\varepsilon)t}\, dt = \int_{0}^{+\infty} e^{-(i\omega+\varepsilon)t}\, dt &= \frac{1}{i\omega + \varepsilon} \\ &= \frac{-i\omega + \varepsilon}{\omega^2 + \varepsilon^2} = \frac{\varepsilon}{\omega^2 + \varepsilon^2} - i\frac{\omega}{\omega^2 + \varepsilon^2}\end{aligned} \tag{A.18}$$

In the limit $\varepsilon \to 0$, (A.13) shows the first term on the right-hand side to be $\pi\delta(\omega)$, and the second term becomes $-i/\omega$ except if $\omega = 0$. If $f(\omega)$ is a well-behaved function, we have

$$\begin{aligned}\lim_{\varepsilon\to 0^+} \int_{-\infty}^{+\infty} f(\omega)\frac{\omega}{\omega^2 + \varepsilon^2}\, d\omega = \lim_{\varepsilon\to 0^+} \int_{-\infty}^{-\varepsilon} f(\omega)\frac{d\omega}{\omega} + \lim_{\varepsilon\to 0^+} \int_{\varepsilon}^{+\infty} f(\omega)\frac{d\omega}{\omega} \\ + \lim_{\varepsilon\to 0^+} \int_{-\varepsilon}^{\varepsilon} f(\omega)\frac{\omega d\omega}{\omega^2 + \varepsilon^2} = \mathsf{P}\int_{-\infty}^{+\infty} f(\omega)\frac{d\omega}{\omega}\end{aligned}$$

where P denotes the *Cauchy Principal value* of the integral. The integral from $-\varepsilon$ to ε vanishes because, since $f(\omega)$ is smooth, the integrand is an odd function of ω. Hence, (A.18) may, with the usual convention about its use in integration, be written as

$$\lim_{\varepsilon\to 0^+} \frac{1}{i\omega + \varepsilon} = \pi\delta(\omega) - i\mathsf{P}\frac{1}{\omega} \tag{A.19}$$

Both (A.17) and (A.19) are useful in formal scattering theory (Chapter 20). From (A.19) we deduce that

$$\lim_{\varepsilon\to 0^+} \frac{1}{2}\left[\frac{1}{\omega + i\varepsilon} + \frac{1}{\omega - i\varepsilon}\right] = \mathsf{P}\frac{1}{\omega} \tag{A.20}$$

and

$$\lim_{\varepsilon\to 0^+} \left[\frac{1}{\omega - i\varepsilon} - \frac{1}{\omega + i\varepsilon}\right] = 2\pi i\delta(\omega) \tag{A.21}$$

We also require the value of the Fourier transform of $1/(\omega - i\varepsilon)$ for finite values of $\varepsilon > 0$. We can obtain this by integrating (A.16) over all ω:

$$\int_{-\infty}^{+\infty} \frac{e^{i\omega T}}{i\omega + \varepsilon}\, d\omega = \int_{-\infty}^{T} dt \int_{-\infty}^{+\infty} e^{i\omega t}\, d\omega\, e^{\varepsilon(t-T)} = 2\pi\int_{-\infty}^{T} \delta(t)e^{\varepsilon(t-T)}\, dt$$

from which we conclude that, for positive ε,

$$\int_{-\infty}^{+\infty} \frac{e^{i\omega T}}{\omega - i\varepsilon}\, d\omega = \begin{cases} 2\pi i\, e^{-\varepsilon T} & \text{if} \quad T > 0 \\ 0 & \text{if} \quad T < 0 \end{cases} \tag{A.22}$$

This relation is important for exponential decay processes.

We can also use this last equation to construct the Fourier representation of the *Heaviside step function*, which is defined as

$$\eta(x) = \int_{-\infty}^{x} \delta(u)\, du = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases} \tag{A.23}$$

From (A.22) we see that

$$\frac{1}{2\pi i} \int_{-\infty}^{+\infty} \frac{e^{i\omega x}}{\omega - i\varepsilon}\, d\omega = e^{-\varepsilon x}\eta(x) \tag{A.24}$$

If we take the limit $\varepsilon \to 0$ and substitute (A.19) in the integrand, we obtain

$$\eta(x) = \frac{1}{2} + \frac{1}{2\pi i}\, \mathbf{P} \int_{-\infty}^{+\infty} \frac{e^{i\omega x}}{\omega}\, d\omega \tag{A.25}$$

We conclude this section by presenting several useful identities involving the delta function. The proofs are easy. First, we have the identity:

$$\delta(ax) = \frac{1}{|a|}\, \delta(x) \tag{A.26}$$

for any nonzero real constant a. Next, we see that for a well-behaved function $f(x)$,

$$f(x)\delta(x - a) = f(a)\delta(x - a) \tag{A.27}$$

A simple inference is that

$$x\delta(x) = 0 \tag{A.28}$$

Equation (A.26) can be generalized to give a formula for the delta function of a function $g(x)$. The delta function $\delta(g(x))$ vanishes except near the zeros of $g(x)$. If $g(x)$ is analytic near its zeros, x_i, the approximation $g(x) \approx g'(x_i)(x - x_i)$ may be used for $x \approx x_i$. From the definition (A.10) and from (A. 26) we infer the equivalence

$$\delta(g(x)) = \sum_i \frac{1}{|g'(x_i)|}\, \delta(x - x_i) \tag{A.29}$$

provided that $g'(x_i) \neq 0$. A special case of this identity is

$$\delta((x - a)(x - b)) = \frac{1}{|a - b|}\, [\delta(x - a) + \delta(x - b)] \tag{A.30}$$

For example,

$$\delta(x^2 - a^2) = \frac{1}{2|a|}\, [\delta(x - a) + \delta(x + a)] \qquad (a \neq 0) \tag{A.31}$$

and

$$\delta(\sqrt{x} - \sqrt{a}) = 2\sqrt{a}\, \delta(x - a) \qquad (a > 0) \tag{A.32}$$

The theory of Fourier integrals and the delta function in three dimensions is straightforward, if we generalize (A.8) by defining

$$\delta(\mathbf{r}) = \frac{1}{(2\pi)^3}\int d^3k\, e^{i\mathbf{k}\cdot\mathbf{r}} \tag{A.33}$$

2. ***Review of Probability Concepts.*** In this brief review of probability concepts, we assume familiarity with the basic rules of adding and multiplying probabilities.[1] The results of a real or imagined preparation, experiment, test, or trial on a system are unambiguously identified as *outcomes* $O_1, O_2, \ldots$ We assume that each possible outcome, O_i, is assigned a probability of occurrence p_i with

$$\sum_i p_i = 1$$

The set of probabilities p_i is called a (probability) *distribution*. In quantum physics, an experiment leading to a set of possible outcomes with probability p_i is generically referred to as a *measurement*, even though the outcome may not necessarily measure the value of a particular quantity.

In rolling an ordinary die, the six possible results, the number of dots on the top, may be identified as outcomes O_1 through O_6, their probabilities being $p_1 = p_2 = \ldots = p_6 = 1/6$, if the die is perfect, unbiased, and not loaded. If, on the other hand, the ''even'' (2, 4, 6) or ''odd'' (1, 3, 5) character of the die roll is chosen for two distinguishable outcomes, O_1 and O_2, the probabilities are 1/2 each.

A variable X which takes on the value of X_1 if outcome O_1 occurs, X_2 if O_2 occurs, etc., is called a *random variable*. A function of a random variable is itself also a random variable. If in rolling the die you are promised 2 pennies for each dot on the die, your winnings constitute a random variable with values $X_i = 2i$.

The *expectation value* (or expected or mean value), $\langle X\rangle$ [or $E(X)$ in the notation favored by mathematicians], of the random variable X for the given probability distribution is defined to be the weighted sum

$$\boxed{\langle X\rangle = \sum_i X_i p_i} \tag{A.34}$$

In the example, $\langle X\rangle = (2 + 4 + 6 + 8 + 10 + 12) \times 1/6 = 7$ pennies. Seven pennies is the expected gain per die roll, whence the term expectation value for $\langle X\rangle$. (The game will be a fair one, if the ante is 7 pennies.)

The *variance* $(\Delta X)^2$ of the random variable X is defined by

$$\boxed{(\Delta X)^2 = \langle (X - \langle X\rangle)^2\rangle = \sum_i (X_i - \langle X\rangle)^2\, p_i = \langle X^2\rangle - \langle X\rangle^2} \tag{A.35}$$

It is a measures of the deviation from the mean. In the example of the die, the random variable X with values $X_i = 2i$ $(i = 1, \ldots 6)$ gives the variance

$$\boxed{(\Delta X)^2 = 4 \times \frac{1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2}{6} - 7^2 = \frac{70}{6}}$$

and the root-mean-square or standard deviation is $\Delta X = 3.42$.

[1]For more on probability, see Chapter 6 in Bradbury (1984).

In the rigorous formulation of the Heisenberg uncertainty relations we also encounter the *covariance* of two random variables X and Y:

$$\boxed{\langle (X - \langle X\rangle)(Y - \langle Y\rangle)\rangle = \langle XY\rangle - \langle X\rangle\langle Y\rangle} \tag{A.36}$$

If X and Y are independent (uncorrelated), the average of the product XY equals the product of the averages $\langle X\rangle\langle Y\rangle$, and the covariance vanishes.

Exercise A.1. In the example of the perfect die, let $X = 0$ if the number of dots is less than its mean value and $X = 1$ if it exceeds its mean value, and $Y = 0$ or 1 depending on whether the number of dots is even or odd, respectively. Compute the covariance.

We can relate a given probability distribution to the statistical properties of a string of N similar trials which lead to outcomes, $O_1, O_2, \ldots O_n$ with probabilities $p_1, p_2, \ldots p_n$. The total number of distinguishable strings of N outcomes is n^N. The number of these strings in which outcome O_1 occurs N_1 times, outcome O_2 occurs N_2 times, and so forth, with $\sum_{N_i=1}^{n} N_i = N$, regardless of their order of occurrence in the string, is the multinomial coefficient

$$(N; N_1, N_2, \ldots N_n) = \frac{N!}{N_1!N_2! \ldots N_n!} \tag{A.37}$$

The probability that in the N trials outcome O_1 occurs N_1 times, outcome O_2 occurs N_2 times, and so forth, is given by the *multinomial distribution*,

$$P(N; N_1, N_2, \ldots N_n) = \frac{N!}{N_1!N_2! \ldots N_n!} p_1^{N_1} p_2^{N_2} \ldots p_n^{N_n} \tag{A.38}$$

The mean value $\langle N_i\rangle$ of the number of successes in producing the outcome O_i is

$$\langle N_i\rangle = Np_i \tag{A.39}$$

Exercise A.2. Show, by use of the multinomial expansion formula, that for given values of N and n, the sum of terms (A.37) is equal to n^N and that the probabilities (A.38) add up to unity. Prove (A.39) and show that the variance $(\Delta N_i)^2$ is

$$(\Delta N_i)^2 = \langle N_i^2\rangle - \langle N_i\rangle^2 = Np_i(1 - p_i) \tag{A.40}$$

Formulas (A.39) and (A.40) show that in the limit of large N, the multinomial probability distribution is sharply peaked at $N_i = Np_i$ $(i = 1, 2, \ldots, n)$. Hence, the *average* value (or *mean* value) of X obtained in the string of trials is, with high probability, equal to the expectation value $\langle X\rangle$,

$$\lim_{N\to\infty} \frac{\sum_i X_i N_i}{N} = \langle X\rangle \tag{A.41}$$

an expression of the *law of large numbers*. The two terms, average and expectation value, are therefore often used synonymously.

In quantum mechanics, as in classical physics, our information about the state of a system is often necessarily less than maximal and we must rely on statistical

methods to describe the state. The *entropy*, which is a measure of disorder or randomness in statistical thermodynamics, is thus expected to be of even greater importance in quantum mechanics, since the intrinsically probabilistic nature of quantum mechanics introduces novel features that have no counterpart in classical physics.[2]

The number (A.37) of distinguishable strings that correspond to the maximum of the multinomial probability distribution is

$$(N; \langle N_1\rangle, \langle N_2\rangle, \ldots \langle N_n\rangle) = \frac{N!}{(Np_1)!(Np_2)! \ldots (Np_n)!} \tag{A.42}$$

For large values of N, this is generally a very large number, the magnitude of which is a measure of the degree of randomness of the original probability distribution. Since, being a probability, $P(N; \langle N_1\rangle, \langle N_2\rangle, \ldots, \langle N_n\rangle)$ must be less than unity, we can derive an asymptotic estimate of (A.42) from (A.38):

$$\frac{N!}{(Np_1)!(Np_2)! \ldots (Np_n)!} \cong p_1^{-Np_1} p_2^{-Np_2} \ldots p_n^{-Np_n}$$

As in statistical mechanics, it is preferable to express this quantity in terms of its logarithm:

$$\log(N; \langle N_1\rangle, \langle N_2\rangle, \ldots \langle N_n\rangle) \approx -N \sum_{i=1}^{n} p_i \log p_i$$

We thus arrive at a quantity that is useful in characterizing the degree of randomness or disorder of a given probability distribution:

$$\boxed{H = -\sum_i p_i \ln p_i = -\ln 2 \sum_i p_i \log_2 p_i = -0.693 \sum_i p_i \log_2 p_i} \tag{A.43}$$

Known as the *Shannon entropy*, in analogy to the Boltzmann entropy, H can be interpreted as an average measure of the *information* that is missing and that is gained when it is determined which of the distinguishable outcomes satisfying the condition $\sum_i p_i = 1$ occurs. The unit of the entropy H in (A.43) is called the *nat*. If the entropy is expressed as $-\sum_i p_i \log_2 p_i$, the unit of information is the *bit*, which is equivalent to 0.693 nats. In the example of the die, the missing information is $2 \times \left(-\frac{1}{2}\log_2 \frac{1}{2}\right) = \log_2 2 = 1$ bit if the outcomes are "even" and "odd" rolls; it is $6 \times \left(-\frac{1}{6}\log_2 \frac{1}{6}\right) = \log_2 6 = 2.58$ bits if the outcomes are the six different possibilities of the roll of the die. Roughly, the value of the entropy in bits tells us the minimum number of binary yes–no questions about the outcome of the experiment that an uninformed observer must ask to be assured of learning which event has occurred. The value of H in bits is also a lower bound for the average length of the sequence of binary digits required to code a set of n messages, if the ith message occurs with probability p_i.

[2]Jones (1979) is a useful book on basic information theory. Quantum information theory, as well as quantum coding, is discussed in Peres (1995). See also Schumacher (1995).

Exercise A.3. Devise a strategy of asking yes–no questions that guarantees that one can ascertain the outcome of the roll of the die, with six distinct possibilities, in three attempts or less.

Exercise A.4. Compute the Shannon entropy in bits for the head-tail toss with an unbiased coin, and compare the result with the value of H for a slightly biased coin ($p_{\text{heads}} = 0.48$, $p_{\text{tails}} = 0.52$). How many trials with the biased coin does it take to ensure that the average number of "heads" differs from the average number of "tails" by more than two root-mean-square deviations?

A key property that characterizes the entropy, and that the Shannon definition (A.43) satisfies, is that if two statistically independent probability distributions p'_i and p''_j are combined by multiplication into one joint probability distribution, $p_{ij} = p'_i p''_j$, the corresponding entropies add:

$$\begin{aligned} H &= -\sum_{i,j} p_{ij} \ln p_{ij} = -\sum_{i,j} p'_i p''_j \ln p' p''_j \\ &= -\sum_i p'_i \ln p'_i - \sum_j p''_j \ln p''_j = H' + H'' \end{aligned} \tag{A.44}$$

The connection between the information entropy (A.43) and the thermodynamic entropy of a physical system in equilibrium can be glimpsed if we inquire about the probability distribution which, given the known physical constraints on the system, maximizes the missing information. Here we merely treat a simple idealized model. The application to a more physically realistic situation is outlined in Section 22.5.

Suppose that the system is constrained only by the expectation value $\langle X \rangle$ of a random variable X, which we assume to be known. We ask for the probabilities p_i which maximize the entropy H, subject to the normalization condition

$$\sum_i p_i = 1$$

and the condition that $\langle X \rangle$ must take on the prescribed value. Using the method of Lagrangian multipliers, we must determine the extremum of $H - \lambda \langle X \rangle$:

$$\delta H - \lambda \delta \langle X \rangle = -\sum_i \delta p_i \ln p_i - \sum_i \delta p_i - \lambda \sum_i X_i \delta p_i = 0 \tag{A.45}$$

Because of the normalization, the variations of the probabilities must be related by

$$\sum_i \delta p_i = 0$$

Thus, (A.45) gives the optimal probability distribution as

$$p_i = C e^{-\lambda X_i} \tag{A.46}$$

The constants C and λ are determined by the two constraint equations. The probability distribution (A.46) has an uncanny similarity with the Boltzmann distribution for a canonical ensemble in statistical thermodynamics, if X denotes the energy and if the Lagrangian multiplier λ is identified with the reciprocal temperature in appropriate units. The reciprocal of C is known as the *partition function*.

Exercise A.5. Show that in the absence of any constraint, the Shannon entropy H is maximized by the uniform probability distribution $p_1 = p_2 = \ldots = p_n = 1/n$, if n is the number of accessible outcomes. (This is a model for the microcanonical ensemble of statistical mechanics.)

Exercise A.6. Assuming that the expectation values of several random variables are known, generalize the derivation of the probability distribution that maximizes the entropy subject to these multiple constraints. (This is a model for the grand canonical ensemble of statistical mechanics.)

In this section, all formulas were written specifically for discrete distributions. In most cases of interest in this book, the extension to continuous distributions, with integrals replacing sums, is straightforward, provided that one employs an appropriate density of outcomes (or density of states). From a practical point of view, we can in many cases, approximate continuous probability distributions by ''discretized'' ones, in which small ranges of outcomes are assigned to discrete ''bins.'' Such a procedure is natural in many physical settings and analogous to experimental techniques with finite resolution, which invariably involve collecting data over narrow ranges of continuous variables. The continuous angular distribution of particles scattered from a target into detectors with finite aperture illustrate this point.

3. ***Curvilinear Coordinates.*** The symmetry of a problem often dictates the most advantageous choice of coordinates. Spherical polar coordinates are convenient for problems with rotational symmetry, and rectangular Cartesian coordinates are appropriate for problems with cubic symmetry. Parabolic coordinates are convenient for the Coulomb potential, even in the presence of a uniform electric field. Here we summarize the essential formulas for expressing the Schrödinger equation in arbitrary curvilinear coordinates.

The location of a particle in space is specified by its generalized coordinates:

$$q_i = q_i(\mathbf{r}) \qquad (i = 1, 2, 3) \tag{A.47}$$

The differential properties of the coordinates are conveniently characterized by the elements of the metric tensor g_{ik}, which are the coefficients in the quadratic form that represents the square of an infinitesimal displacement:

$$ds^2 = g_{ik}\, dq^i\, dq^k \tag{A.48}$$

The summation convention is assumed to be in force: One must sum over repeated indices, if one of them is a subscript and the other a superscript. We assume that the space is Euclidean, or flat, so that it is possible to express the metric as

$$ds^2 = (dx_1)^2 + (dx_2)^2 + (dx_3)^2 = dx^2 + dy^2 + dz^2$$

In general coordinates, the Laplacian operator has the form[3]

$$\nabla^2 = \frac{1}{\sqrt{g}}\, \partial_i \sqrt{g} g^{ik} \partial_k \tag{A.49}$$

where we have denoted the partial derivatives as

$$\partial_i = \frac{\partial}{\partial q_i} \tag{A.50}$$

[3]For a proof see Bradbury (1984), Chapter 4, Section 10.

The quantity of g is the determinant of the metric tensor:

$$g = \det g_{ik} \tag{A.51}$$

General coordinates are orthogonal if and only if the metric tensor is diagonal. For orthogonal coordinates,

$$g = g_{11}g_{22}g_{33}$$

and the Laplacian (A.49) reduces to

$$\nabla^2 = \frac{1}{\sqrt{g}}\left[\frac{\partial}{\partial q_1}\sqrt{\frac{g_{22}g_{33}}{g_{11}}}\frac{\partial}{\partial q_1} + \frac{\partial}{\partial q_2}\sqrt{\frac{g_{33}g_{11}}{g_{22}}}\frac{\partial}{\partial q_2} + \frac{\partial}{\partial q_3}\sqrt{\frac{g_{11}g_{22}}{g_{33}}}\frac{\partial}{\partial q_3}\right] \tag{A.52}$$

As an example, we derive the Laplacian for parabolic coordinates by a direct calculation and then compare the result with (A.52).

Parabolic coordinates are defined by

$$x = \sqrt{\xi\eta}\cos\varphi \qquad y = \sqrt{\xi\eta}\sin\varphi \qquad z = \frac{1}{2}(\xi - \eta) \tag{A.53}$$

Hence,

$$\mathbf{r} = \sqrt{\xi\eta}\cos\varphi\,\hat{\mathbf{x}} + \sqrt{\xi\eta}\sin\varphi\,\hat{\mathbf{y}} + \frac{1}{2}(\xi - \eta)\,\hat{\mathbf{z}} \tag{A.54}$$

and

$$\begin{aligned} d\mathbf{r} &= \frac{1}{2}\left(\sqrt{\frac{\eta}{\xi}}\cos\varphi\,\hat{\mathbf{x}} + \sqrt{\frac{\eta}{\xi}}\sin\varphi\,\hat{\mathbf{y}} + \hat{\mathbf{z}}\right)d\xi + \frac{1}{2}\left(\sqrt{\frac{\xi}{\eta}}\cos\varphi\,\hat{\mathbf{x}} + \sqrt{\frac{\xi}{\eta}}\sin\varphi\,\hat{\mathbf{y}} - \hat{\mathbf{z}}\right)d\eta \\ &\quad + (-\sqrt{\xi\eta}\sin\varphi\,\hat{\mathbf{x}} + \sqrt{\xi\eta}\cos\varphi\,\hat{\mathbf{y}})\,d\varphi \\ &= \frac{1}{2}\sqrt{\frac{\xi+\eta}{\xi}}\,\hat{\boldsymbol{\xi}}\,d\xi + \frac{1}{2}\sqrt{\frac{\xi+\eta}{\eta}}\,\hat{\boldsymbol{\eta}}\,d\eta + \sqrt{\xi\eta}\,\hat{\boldsymbol{\varphi}}\,d\varphi \end{aligned} \tag{A.55}$$

The last equation defines the orthogonal basis vectors that span the parabolic coordinate system. Since the gradient operator ∇ is defined by

$$d\mathbf{r}\cdot\nabla = d\xi\frac{\partial}{\partial\xi} + d\eta\frac{\partial}{\partial\eta} + d\varphi\frac{\partial}{\partial\varphi} \tag{A.56}$$

we obtain

$$\nabla = \hat{\boldsymbol{\xi}}\left(2\sqrt{\frac{\xi}{\xi+\eta}}\frac{\partial}{\partial\xi}\right) + \hat{\boldsymbol{\eta}}\left(2\sqrt{\frac{\eta}{\xi+\eta}}\frac{\partial}{\partial\eta}\right) + \hat{\boldsymbol{\varphi}}\left(\frac{1}{\sqrt{\xi\eta}}\frac{\partial}{\partial\varphi}\right) \tag{A.57}$$

After some algebraic manipulations working out the partial derivatives of the basis vectors with respect to the parabolic coordinates, we obtain for the Laplacian,

$$\nabla^2 = \nabla\cdot\nabla = \frac{4}{\xi+\eta}\frac{\partial}{\partial\xi}\left(\xi\frac{\partial}{\partial\xi}\right) + \frac{4}{\xi+\eta}\frac{\partial}{\partial\eta}\left(\eta\frac{\partial}{\partial\eta}\right) + \frac{1}{\xi\eta}\frac{\partial^2}{\partial\varphi^2} \tag{A.58}$$

On the other hand, in parabolic coordinates, the metric is found from (A.55) to be

$$ds^2 = d\mathbf{r} \cdot d\mathbf{r} = \frac{\xi + \eta}{4\xi}\, d\xi^2 + \frac{\xi + \eta}{4\eta}\, d\eta^2 + \xi\eta\, d\varphi^2 \tag{A.59}$$

from which the metric tensor can be obtained and the Laplacian (A.52) calculated. The result agrees with (A.58).

Exercise A.7. Derive the gradient and Laplacian in spherical polar coordinates, using the techniques of this section.

4. ***Units and Physical Constants.*** In line with common practice in introductory theoretical physics, in this book the Gaussian system is used for defining physical quantities. Table A.1 shows how symbols for electromagnetic quantities in equations and formulas (but not their numerical values) are converted from the Gaussian system to the Heaviside-Lorentz (or "rationalized" Gaussian) system, frequently used in particle physics, and the symbols that underlie the SI system of units.[4]

Table A.1

Quantity	*Gaussian*	*Heaviside-Lorentz*	*SI*
Charge	q or e	$\frac{q}{\sqrt{4\pi}}$ or $\frac{e}{\sqrt{4\pi}}$	$\frac{q}{\sqrt{4\pi\varepsilon_0}}$ or $\frac{e}{\sqrt{4\pi\varepsilon_0}}$
Charge and current density	ρ and $\mathbf{j}$	$\frac{\rho}{\sqrt{4\pi}}$ and $\frac{\mathbf{j}}{\sqrt{4\pi}}$	$\frac{\rho}{\sqrt{4\pi\varepsilon_0}}$ and $\frac{\mathbf{j}}{\sqrt{4\pi\varepsilon_0}}$
Electric field	$\mathbf{E}$	$\sqrt{4\pi}\,\mathbf{E}$	$\sqrt{4\pi\varepsilon_0}\,\mathbf{E}$
Electrostatic potential	ϕ	$\sqrt{4\pi}\,\phi$	$\sqrt{4\pi\varepsilon_0}\,\phi$
Vector potential	$\mathbf{A}$	$\sqrt{4\pi}\,\mathbf{A}$	$\sqrt{4\pi\varepsilon_0}\,\mathbf{A}$
Magnetic field	$\mathbf{B}$	$\sqrt{4\pi}\,\mathbf{B}$	$\sqrt{4\pi\varepsilon_0}\,\mathbf{B}$
Magnetic moment	$\boldsymbol{\mu}$	$\frac{\boldsymbol{\mu}}{\sqrt{4\pi}}$	$\sqrt{\frac{\mu_0}{4\pi}}\,\boldsymbol{\mu}$
Speed of light	c	c	$\frac{1}{\sqrt{\varepsilon_0\mu_0}}$
Fine structure constant α	$\frac{e^2}{\hbar c}$	$\frac{e^2}{4\pi\hbar c}$	$\frac{e^2}{4\pi\varepsilon_0\hbar c}$

In Table A.2, we list numerical values of important quantities and fundamental constants.[5] No rigid system of units is slavishly adhered to in this table, since the choice is always suggested by the context in which the quantities are used.

[4]In Table A.1 we adopt the format used in Jackson (1975), Appendix.
[5]The numbers in Table A.2 are adapted from Cohen and Taylor (1996).

Table A.2

c (speed of light)	$2.9979 \cdot 10^{8}$ m s^{-1}
h	$6.6261 \cdot 10^{-34}$ J s
$\hbar = h/2\pi$ (Planck's constant)	$1.0546 \cdot 10^{-27}$ erg s
	$6.5821 \cdot 10^{-16}$ eV s
e (electron charge)	$1.6022 \cdot 10^{-19}$ C
m_e (electron mass)	$9.1094 \cdot 10^{-31}$ kg
	0.51099 MeV/c^2
$\frac{\hbar}{m_e c}$ (Compton wavelength/2π)	$3.8616 \cdot 10^{-13}$ m
m_n/m_e (neutron-electron mass ratio)	1838.7
u (atomic mass unit)	$1.6605 \cdot 10^{-24}$ kg
$a = \frac{\hbar^2}{m_e e^2}$ (Bohr radius)	$0.52918 \cdot 10^{-10}$ m
$E_{1s} = \frac{e^2}{2a}$ (ground state energy of hydrogen atom)	13.61 eV
$\alpha = \frac{e^2}{\hbar c}$ (fine structure constant)	$7.2974 \cdot 10^{-3} = 1/137.035$
$\beta_0 = \frac{e\hbar}{2m_e c}$ (Bohr magneton)	$9.2740 \cdot 10^{-24}$ J T^{-1}
	$9.2740 \cdot 10^{-21}$ erg/gauss
	$5.788 \cdot 10^{-5}$ eV T^{-1}
k (Boltzmann's constant)	$1.3807 \cdot 10^{-23}$ J K^{-1}
	$8.6174 \cdot 10^{-5}$ eVK^{-1}
eV (electron volt)	$1.6022 \cdot 10^{-19}$ J

Often it is convenient to work with *natural* units, which are based on quantities that are prominent in a particular subfield of physics. For example, atomic and condensed-matter physics problems are conveniently formulated in terms of natural units defined by setting

$$\hbar = 1, \qquad m_e = 1, \qquad e = 1 \tag{A.60}$$

With this choice, the Bohr radius, $a = \hbar^2/m_e e^2$, becomes the unit of length, and $2E_{1s}$ (27.21 eV) the unit of energy, known as the *Hartree* unit. Since the fine structure constant is dimensionless, the speed of light in these natural atomic units (or a.u.) is $c \approx 137$ a.u. Since the proton mass is $m_P = 1836$ a.u., the kinetic energy of a proton with velocity $v = 1$ a.u. is $E = 918$ Hartree units, or about 25 keV.

In particle physics, the preferred natural units are quite different, since one often works in the relativistic regime. A frequent choice is

$$\hbar = 1 \qquad c = 1 \tag{A.61}$$

and a third unit, usually an energy, such as 1 GeV. The value of e is now determined by the fine structure constant to be $e = \sqrt{\alpha}$ in Gaussian units and $e = \sqrt{4\pi\alpha} \approx 0.3$ in Heaviside-Lorentz units.

References

Abramowitz, Milton, and Irene A. Stegun. *Handbook of Mathematical Functions*. Washington, D.C.: National Bureau of Standards, 1964.

Arfken, G. *Mathematical Methods for Physicists*. New York: Academic Press, 1985.

Ashcroft, N. W., and N. D. Mermin. *Solid State Physics*. New York: Holt, Rinehart, and Winston, 1976.

Ballentine, Leslie E. "Resource Letter IQM-2: Foundations of Quantum Mechanics Since the Bell Inequalities." *Am. J. Phys.* **55**, 785 (1987).

Ballentine, Leslie E. *Quantum Mechanics*. Englewood Cliffs, N.J.: Prentice–Hall, 1990.

Bargmann, V. "Note on Wigner's Theorem on Symmetry Operations." *J. of Math. Phys.* **5**, 862 (1964).

Baym, Gordon. *Lectures on Quantum Mechanics*. New York: Benjamin/Cummings, 1969 (revised 1981).

Bethe, Hans A., and Roman Jackiw. *Intermediate Quantum Mechanics*. New York: W. A. Benjamin, 1968, 2nd ed.

Bethe, H. A., and E. E. Salpeter. *Quantum Mechanics of One- and Two-Electron Atoms*. New York: Academic Press, 1957.

Bialynicki-Birula, Iwo, Marek Cieplak, and Jerzy Kaminski. *Theory of Quanta*. New York: Oxford University Press, 1992.

Biedenharn, L. C., and H. Van Dam, eds. *Quantum Theory of Angular Momentum*. New York: Academic Press, 1965.

Biedenharn, L. C., and J. D. Louck. "Angular Momentum in Quantum Physics" in *Encyclopedia of Mathematics and its Applications*, Vol. 8. Reading, Mass.: Addison-Wesley, 1981.

Bjorken, James D., and Sidney D. Drell. *Relativistic Quantum Mechanics*. New York: McGraw-Hill, 1964.

Bjorken, James D., and Sidney D. Drell. *Relativistic Quantum Fields*. New York: McGraw-Hill, 1965.

Blatt, J. M., and V. F. Weisskopf. *Theoretical Nuclear Physics*. New York: John Wiley, 1952.

Bohm, Arno. *Quantum Mechanics-Foundations and Applications*. New York: Springer-Verlag, 1994, 3rd ed., 2nd rev. printing.

Bohm, David. *Quantum Theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1951.

Bohr, Niels. *Atomic Physics and Human Knowledge*. New York: John Wiley, 1958.

Bradbury, T. C. *Mathematical Methods with Applications to Problems in Physical Sciences*. New York: John Wiley, 1984.

Brandt, Siegmund, and Hans Dieter Dahmen. *The Picture Book of Quantum Mechanics*. New York: John Wiley, 1985.

Bransden, Brian H. *Atomic Collision Theory*. New York: Benjamin/Cummings, 1983.

Brehm, John J., and William J. Mullin. *Introduction to the Structure of Matter*. New York: John Wiley, 1989.

Brink, D. M., and G. R. Satchler. *Angular Momentum*. Oxford: Clarendon Press, 1968, 2nd ed.

Cagnac, B., and J.-C. Pebay-Peyroula. *Modern Atomic Physics*. New York: John Wiley, 1971. Vol. 1: *Fundamental Principles*; Vol. 2: *Quantum Theory and Its Applications*.

Callen, Herbert B. *Thermodynamics and an Introduction to Thermostatistics*. New York: John Wiley, 1985, 2nd ed.

Christman, J. Richard. *Fundamentals of Solid State Physics*. New York: John Wiley, 1988.

Cohen, E. Richard, and Barry N. Taylor. "The Fundamental Physical Constants." *Physics Today*, Annual Buyers Guide, August 1996.

Cohen-Tannoudji, Claude, Bernard Diu, and Frank Laloë. *Quantum Mechanics*. New York: John Wiley, 1977, Vols. I and II.

Condon, E. U., and G. H. Shortley. *The Theory of Atomic Spectra*. London: Cambridge University Press, 1935, rev. printing 1953.

Courant, R., and D. Hilbert. *Methods of Mathematical Physics*. New York: Interscience, 1953.

Coxeter, H. S. M. *Introduction to Geometry*. New York: John Wiley, 1969.

Cushing, J. T. *Quantum Mechanics—Historical Contingency and the Copenhagen Hegemony*. Chicago: University of Chicago Press, 1994.

Dekker, Adrianus J. *Solid State Physics*. Englewood Cliffs, N.J., Prentice-Hall, 1957.

DeWitt, B. S., and R. N. Graham. "Resource Letter IQM-1: The Interpretation of Quantum Mechanics." *Am. J. Physi.* **39**, 724 (1971).

Dicke, R. H., and J. P. Wittke. *Introduction to Quantum Mechanics*. Reading, Mass.: Addison-Wesley, 1960.

Dirac, P. A. M. *The Principles of Quantum Mechanics*. Oxford: Clarendon Press, 1958, 4th ed.

Fano, Guido. *Mathematical Methods of Quantum Mechanics*. New York: McGraw-Hill, 1971.

Feagin, James M. *Quantum Methods with Mathematica*. New York: Springer-Verlag, 1993.

Fetter, Alexander L., and John Dirk Walecka. *Quantum Theory of Many-Particle Systems*. New York: McGraw-Hill, 1971.

Feynman, R. P., and A. R. Hibbs. *Quantum Mechanics and Path Integrals*. New York: McGraw-Hill, 1965.

Feynman, Richard P., Robert B. Leighton, and Matthew Sands. *The Feynman Lectures on Physics*, Volume III. Reading, Mass.: Addison-Wesley, 1965.

Gasiorowicz, Stephen. *Quantum Physics*. New York: John Wiley, 1996, 2nd ed.

Glauber, Roy J. "Optical Coherence and Photon Statistics" in *Quantum Optics and Electronics*. New York: Gordon and Breach, 1965.

Goldberger, M. L., and K. M. Watson. *Collision Theory*. New York: John Wiley, 1964.

Goldstein, H. *Classical Mechanics*. Reading, Mass.: Addison-Wesley, 1980, 2nd ed.

Gottfried, Kurt. *Quantum Mechanics*, Volume I. New York: W. A. Benjamin, 1966.

Gradshteyn, I. S., and I. M. Ryzhik. *Table of Integrals, Series, and Products*. New York: Academic Press, 1965.

Green, H. S. *Matrix Mechanics*. Groningen: Noordhoff, 1965.

Greiner, Walter. *Quantum Mechanics, an Introduction*. New York: Springer-Verlag, 1989.

Griener, Walter, and Berndt Müller. *Quantum Mechanics, Symmetries*. New York: Springer-Velag, 1989.

Griffiths, D. *Introduction to Quantum Mechanics*. Englewood Cliffs, N.J.: Prentice-Hall, 1995.

Gross, Franz. *Relativistic Quantum Mechanics and Field Theory*. New York: John Wiley, 1993.

Gutzwiller, Martin C. *Chaos in Classical and Quantum Mechanics*. New York: Springer-Verlag, 1990.

Haken, Hermann, and Hans Christoph Wolf. *The Physics of Atoms and Quanta*. New York: Springer-Verlag, 1993.

Halzen, Francis, and Alan D. Martin. *Quarks and Leptons*. New York: John Wiley, 1984.

Hassani, Sadri. *Foundations of Mathematical Physics*. Needham Heights, Mass.: Allyn and Bacon, 1991.

Heisenberg, Werner. *The Physical Principles of the Quantum Theory*. Chicago: University of Chicago Press, 1930. Dover reprint, 1949. (Translated by C. Eckart and C. Hoyt.)

Holstein, Barry. *Topics in Advanced Quantum Mechanics*. Reading, Mass.: Addison-Wesley, 1992.

Itzykson, Claude, and Jean-Bernard Zuber. *Quantum Field Theory*. New York: McGraw-Hill, 1980.

Jackson, J. D. *Classical Electrodynamics*. New York: John Wiley, 1975, 2nd ed.

Jammer, Max. *The Conceptual Development of Quantum Mechanics*. New York: McGraw-Hill, 1966.

Jones, D. S. *Elementary Information Theory*. Oxford: Clarendon Press, 1979.

Jordan, T. F. *Linear Operators for Quantum Mechanics*. New York: John Wiley, 1969. (Available from Robert E. Krieger Publishing Co., P.O. Box 9542, Melbourne, Fl. 32901.)

Jordan, Thomas F. *Quantum Mechanics in Simple Matrix Form*. New York: John Wiley, 1986.

Judd, Brian R. *Second Quantization and Atomic Spectroscopy*. Baltimore, Md.: Johns Hopkins Press, 1967.

Koltun, Daniel S., and Judah M. Eisenberg. *Quantum Mechanics of Many Degrees of Freedom*. New York: John Wiley, 1988.

Kramers, H. A. *Quantum Mechanics*. New York: Interscience, 1957.

Krane, Kenneth S. *Introductory Nuclear Physics*. New York: John Wiley, 1987.

Landau, L. D., and E. M. Lifshitz. *Quantum Mechanics*. Reading, Mass.: Addison-Wesley, 1958. (Translated by J. B. Sykes and J. S. Bell.)

Landau, Rubin H. *Quantum Mechanics II*. New York: John Wiley, 1990.

Lévy-Leblond, Jean-Marc, and Françoise Balibar. *Quantics (Rudiments of Quantum Physics)*. Amsterdam: North-Holland, 1990.

Liboff, Richard L. *Introductory Quantum Mechanics*. Reading, Mass.: Addison-Wesley, 1992, 2nd ed.

Loudon, Rodney. *The Quantum Theory of Light*. Oxford: Clarendon Press, 1983.

Magnus, Wilhelm, and Fritz Oberhettinger. *Formulas and Theorems for the Special Functions of Mathematical Physics*. New York: Chelsea, 1949.

Mandl, F., and G. Shaw. *Quantum Field Theory*. New York: John Wiley, 1984.

Mathews, J., and R. L. Walker. *Mathematical Methods of Physics*. New York: W. A. Benjamin, 1964.

Merzbacher, Eugen. "Single Valuedness of Wave Functions." *Am. J. Phys.* **30**, 237 (1962).

Merzbacher, Eugen. "Matrix Methods in Quantum Mechanics." *Am. J. Phys.* **36**, 814 (1968).

Merzbacher, Eugen. *Quantum Mechanics*. New York: John Wiley, 1970, 2nd ed.

Messiah, A. *Quantum Mechanics*. Amsterdam: North-Holland, 1961 and 1962. (Vol. I translated by G. Temmer, Vol. II translated by J. Potter.)

Milonni, Peter. *The Quantum Vacuum: An Introduction to Quantum Electrodynamics*. New York: Academic Press, 1995.

Mizushima, Masataka. *Quantum Mechanics of Atomic Spectra and Atomic Structure*. New York: W. A. Benjamin, 1970.

Morrison, Michael A. *Understanding Quantum Mechanics*. Englewood, N.J.: Prentice-Hall, 1990.

Morse, P. M., and H. Feshbach. *Methods of Theoretical Physics*, Volumes I and II. New York: McGraw-Hill, 1953.

Mott, N. F., and H. S. W. Massey. *The Theory of Atomic Collisions*. Oxford: Clarendon Press, 1965, 3rd ed.

Neumann, J. von. *Mathematical Foundations of Quantum Mechanics*. Berlin: Springer-Verlag, 1932. (Translated by R. T. Beyer. Princeton, N.J.: Princton University Press, 1955.)

Newton, R. G. *Scattering Theory of Waves and Particles*. New York: McGraw-Hill, 1982, 2nd ed.

Omnès, Roland. *The Interpretation of Quantum Mechanics*. Princeton, N.J.: Princeton University Press, 1994.

Park, David A. *Introduction to the Quantum Theory*. New York: McGraw-Hill, 1992, 3rd ed.

Parr, Robert G., and Weitao Yang. *Density-Functional Theory of Atoms and Molecules*. New York: Oxford University Press, 1989.

Pauli, Wolfgang. "Die allgemeinen Prinzipien der Wellenmechanik," in *Encyclopedia of Physics*, Vol. 5/1, pp. 1–168. New York: Springer-Verlag, 1958.

Pauling, Linus C., and B. E. Wilson, Jr. *Introduction to Quantum Mechanics with Applications to Chemistry*. New York: McGraw-Hill, 1935.

Peres, Asher. *Quantum Theory: Concepts and Methods*. Dordrecht: Kluwer Academic Publishers, 1995.

Perkins, Donald H. *Introduction to High Energy Physics*. Reading, Mass.: Addison-Wesley, 1982, 2nd ed.

Powell, J. L., and B. Crasemann. *Quantum Mechanics*. Reading, Mass.: Addison-Wesley, 1961.

Reed, M., and B. Simon. *Methods of Modern Mathematical Physics*, Volumes I–IV. New York: Academic Press, 1975–1980.

Reif, F. *Fundamentals of Statistical and Thermal Physics*. New York: McGraw-Hill, 1965.

Riesz, F., and B. Sz.-Nagy. *Functional Analysis*. New York: Ungar, 1955. (Reprinted by Dover.)

Robinett, Richard W. *Quantum Mechanics, Classical Results, Modern Systems, and Visualized Examples*. New York: Oxford University Press, 1997.

Rose, M. E. *Elementary Theory of Angular Momentum*. New York: John Wiley, 1957.

Rose, M. E. *Relativistic Electron Theory*. New York: John Wiley, 1961.

Sakurai, J. J. *Advanced Quantum Mechanics*. Reading, Mass.: Addison-Wesley, 1967.

Sakurai, J. J., with San Fu Tuan, Editor. *Modern Quantum Mechanics*. New York: Benjamin/Cummings, 1994, revised ed.

Schiff, Leonard. *Quantum Mechanics*. New York: McGraw-Hill, 1968, 3rd ed.

Schulman, L. S. *Techniques and Applications of Path Integration*. New York: John Wiley, 1981.

Schumacher, Benjamin. "Quantum coding." *Phys. Rev.* **A51**, 2738 (1995).

Schweber, Silvan S. *An Introduction to Relativistic Quantum Field Theory*. Evanston, Ill.: Row, Peterson and Co., 1961.

Shankar, R. *Principles of Quantum Mechanics*. New York: Plenum, 1980.

Shapere, Alfred, and Frank Wilczek, eds. *Geometric Phases in Physics*. Singapore: World Scientific, 1989.

Taylor, John R. *Scattering Theory*. New York: John Wiley, 1972.

Thompson, William J. *An Illustrated Guide to Rotational Symmetry for Physical Systems*. New York: John Wiley, 1994.

Thompson, William J. *Atlas for Computing Mathematical Functions*. New York: John Wiley, 1997.

Thouless, David J. *The Quantum Mechanics of Many-Body Systems*. New York: Academic Press, 1961.

Tinkham, Michael. *Group Theory and Quantum Mechanics*. New York: McGraw-Hill, 1964.

Townes, C. H., and A. L. Shawlow. *Microwave Spectroscopy*. New York: McGraw-Hill, 1955.

Townsend, John S. *A Modern Approach to Quantum Mechanics*. New York: McGraw-Hill, 1992.

van der Waerden, B. L. *Sources of Quantum Mechanics*. Amsterdam: North-Holland, 1967.

Weinberg, Steven. *The Quantum Theory of Fields*, Vol. I: *Foundations*, Vol. II: *Modern Applications*. London: Cambridge University Press, 1995, 1996.

Weinreich, Gabriel. *Solids: Elementary Theory for Advanced Students*. New York: John Wiley, 1965.

Werner, Samuel A., "Gravitational, Rotational and Topological Quantum Phase Shifts in Neutron Interferometry." *Class. Quantum Grav.* **11**, A207, (1994).

Wheeler, John Archibald, and Wojciech Hubert Zurek, eds. *Quantum Theory and Measurement*. Princeton, N.J.: Princeton University Press, 1983.

White, Harvey E. "Pictorial Representation of the Electron Cloud for Hydrogen-like Atoms." *Phys. Rev.* **37**, 1416 (1931).

Wigner, Eugene P. *Group Theory and Its Applications to the Quantum Mechanics of Atomic Spectra*. New York: Academic Press, 1959. (Translated by J. J. Griffin.)

Wilcox, R. M. "Exponential Operators and Parameter Differentiation in Quantum Physics." *Journal of Math. Phys.* **8**, 962 (1967).

Ziman, J. M. *Elements of Advanced Quantum Theory*. London: Cambridge University Press, 1969.

Index

Abelian group, 69, 445
Absorption of radiation, 491–501, 577–579
  cross section for, 494–501, 510, 514, 591
  by harmonic oscillator, 561
  rate of, 498
  selection rules for, 497
  sum rules for, 499
Action function, classical, 355
Action integral, *see* Phase integral
Active transformation, 76, 201, 411
Addition of angular momenta, 426–431, 555–556
Addition theorem for spherical harmonics, 251, 426
Additive one-particle operator, 544, 615
Additive two-particle operator, 545
Adiabatic approximation, 161
Adiabatic change, 151
Adiabatic invariant, 151
Adjoint Dirac field operator, 599
Adjoint of an operator, 52–53, 192
Aharonov-Bohm effect, 78
Airy equation, 117,123
Airy function, 117–119, 123, 139
Alkali atoms, spectra of, 469
Allowed bands, 166
Allowed transitions, 496
Almost degenerate levels, perturbation of, 145, 463
Alpha decay, 133
Amplitude, 8
Analyzing power, 403
Angular distribution, 281, 301
Angular momentum, 233–255, 414–439
  addition of, 426–431, 555–556
  commutation relations for, 234, 238, 384, 413
  conservation of, 233, 330–331, 375, 389–390, 414–415
  as constant of motion, 256, 389, 414–415, 624
  coupling of, 426–431, 449, 472–473, 555–556
  eigenvalue problem for, 238–248
  in Dirac theory, 604, 624
  Euler angles representation of, 449–450
  as generator of infinitesimal rotations, 236, 382, 389, 413, 449
  intrinsic, 372–377, 389. *See also* Spin
  and kinetic energy, 252–255
  in many-particle system, 555–556
  orbital, 233–237, 425, 443
  parity of eigenstates, 249, 440
  of photons, 569–570, 575–576
  of rigid bodies, 449–450
  rotations and, 236, 381–385, 413–416, 603–604
  superselection rule for, 414
  and time reversal, 443
  total, 389–390, 416, 426, 469–470, 555, 604, 624
    eigenstate of, 430
    for two identical particles, 555–556
Angular momentum operator, 238, 413
Angular momentum quantum number, 239, 245, 390, 422–423, 428–429, 555, 627
Anharmonic oscillator, 177
Anisotropic harmonic oscillator, 480
Annihilation operator, 538, 543–544, 567. *See also* Lowering operator
A-normalization, 61
Anticommutation relations for fermions, 540–543, 590
  for Dirac field, 598
Antilinear operator, 34, 188, 192, 412. *See also* Time reversal operator
Antisymmetry of fermion wave function, 547
Antiunitary operator, *see* Time reversal operator
Associated Laguerre polynomials, 271
Associated Legendre functions, 247
Asymmetry in scattering, right-left, 376–377, 403
Asymptotic expansion:
  of Airy function, 117–119
  of Coulomb wave function, 312
  of parabolic cylinder function, 157
  of plane wave function, 262
  of scattering eigenfunction, 287, 294, 298–300
  of spherical cylinder function, 259–260
  and WKB approximation, 120
Auger transition, *see* Radiationless transition
Autoionizing transition, *see* Radiationless transition
Average value of random variable, 635
Axial vector operator, 440
Axial (pseudo-)vector operator in Dirac theory, 605
Axis of quantization, 244, 423

Balmer formula, 2, 267, 269, 628
Bands, allowed and forbidden, in periodic potential, 166–178, 481
Band theory of solids, 166
Barrier penetration, 97–98, 125–133, 150, *see also* Tunneling
Basis functions, 139
Basis states, 186, 379, 537
Bell, John S., 9, 362
Bell's theorem, 18, 362
Berry's connection and geometric phase, 162
Bessel functions, spherical, 258
Bessel's equation, 258
Bessel's inequality, 58
Black-body radiation, 590
Bloch function, 71, 168, 295
Bohm, David, 29
Bohr, Niels, 1–10, 18
Bohr frequency condition, 1, 20, 396, 492
Bohr magneton, 374
Bohr radius of hydrogen, 267, 641
Boltzmann distribution, 91
Boltzmann statistics, 566
Boost, 75, 607
Born, Max, 7
Born approximation, 295–298, 314, 526, 534
Born interpretation, *see* Probability interpretation,
Born-Oppenheimer approximation, 161
Born series, 526
Bose-Einstein:
  commutation relations, 543, 546
  condensation, 11
  statistics, 543, 566
Bosons, 543
Boundary conditions, 263–265
  at fixed points, 66. *See also* Periodic boundary conditions
  at infinity, 43–45, 82, 104, 262, 265, 626
  at origin, 264, 299, 626–627
Bounded Hermitian operator, 212–214
Bound states, 83, 103–108, 121–125, 262–263, 265–277
Bra, 196
Bragg reflection, 176
Bra-ket, 205
Breit-Wigner resonance, 130, 306
Brillouin, Léon, 113
Brillouin's theorem, 562
Brillouin-Wigner perturbation theory, 467
Brillouin zone, 70–71, 167, 172. *See also* Dispersion function

Campbell-Baker-Hausdorff formula, 40
Canonical coordinates and momenta, 326
Canonical quantization, 326–332
Canonical transformation, 329
Casimir (van der Waals) forces, 574
Casimir operator, to label irreducible representations, 422

Casimir projection operator, 608
Cauchy principal value, 293, 513, 632
Causality, principle of, 315, 319
Cayley-Hamilton theorem, 212
Central forces, 256–275, 623–628
  scattering from, 298–302, 530–532
Centrifugal potential, 257
C-G coefficient, *see* Clebsch-Gordan coefficients
Change of basis, 199–202, 538–542
Chaotic state, 589
Characteristic equation, *see* Determinental equation; Secular equation
Characteristic value, 140. *See also* Eigenvalue
Charge conjugation, 408, 608–610
Chemical bond, 164
Chemical potential, 565
Chirality, 439, 620
Chiral solutions of Dirac equation, 622
Chiral symmetry, 620, 622
Classical approximation, 3, 123. *See also* Correspondence principle; WKB approximation
Classical dynamics, relation to quantum dynamics, 324
Classically accessible region, 116
Classical turning point, 116
Clebsch-Gordan coefficients, 427
  orthonormality relations, 429
  recursion relation, 428, 432
  selection rules, 428
  symmetry relations, 429, 436
  triangular condition, 428
Clebsch-Gordan series, 431–432
Closely coupled states, 486–487
Closure relation, 62–63, 67–68, 198, 529, 597
Coherence, of field, 583–586
Coherent state(s), 220, 225–231, 583
  displacement operator for, 225
  generated by current-field interaction, 581–582
  and Heisenberg uncertainty relation, 229
  inner product of, 226
  overcompleteness relation for, 227, 365
  relation to number operator eigenstates, 227
  representation by entire functions, 228
  rotation operator for, 226
  time development of, 335, 340–342, 353
Coherent superposition, 301
Coincidence rate, 588–589
Collapse of wave packet, 408
Commuting Hermitian operators, 214–217, 407
Commutation relations:
  for angular momentum, 234, 238, 384, 413
  for bosons, 540–543
  for canonically conjugate variables, 326, 332
  for coordinates and momenta, 38, 204, 325, 344
  for creation and annihilation operators, 540–543
  at different times, 332
  for electromagnetic field, 576
  for field operators, 546
  for photon operators, 569
  for tensor operators, 436
  for vector operators, 236, 384
Commutator, 37, 38–41, 218, 326, 332
Compatibility of observables, 180, 407
Complementarity, 4
Completeness:
  of Dirac matrices, 605–606
  of dynamical variables, 372
  of eigenfunctions, 46, 57, 206, 350
  of eigenvectors, 180, 198, 214–217, 529
  of Hermite polynomials, 88
  of spherical harmonics, 249
Completeness relation, 59, 206, 364
Complete set:
  of basis vectors, 209
  of commuting Hermitian operators, 180, 216
  of eigenfunctions, 59
  of functions, 142
  of observables, 180
  of vectors, 186, 217
Complex potential, 78
Complex vector space, 185
  two-dimensional, 377–381
Complex wave function, 13
Composition rule for amplitudes, 182, 315
Configuration, 559
Configuration space, 359, 547
Confluent hypergeometric equation, 270
Confluent hypergeometric functions, 156, 270, 311
Connection, for gauge field, 162, 447
Connection formulas, WKB approximation, 116–121
Conservation:
  of angular momentum, 233, 330–331, 375, 389–390, 414–415
  of charge, 600
  of current, 600, 623
  of energy, 38, 43, 321, 503, 509
  of linear momentum, 38, 330
  of parity, 441, 460
  of probability, 26–28, 42, 94–95, 100, 121, 318, 391, 514
Conservation laws, *see* Constant of the motion
Constant of the motion, 37, 319, 330, 415
Continuity equation, 26, 28, 36, 74, 599–600
Continuous spectrum, 44, 60–62, 94, 181, 202–206, 284, 546, 592
Continuum eigenfunctions, 60–62
  for Coulomb field, 310–312
Contraction:
  of two vectors, *see* Inner product
  of irreducible spherical tensor operators, 436
Convection current density, 623
Convergence in the mean, 58, 142
Coordinate operator, matrix element of:
  in coordinate representation, 204
  in momentum representation, 346
  in oscillator energy representation, 88, 224
Coordinate representation, 32, 204, 344–348
  wave function, 180, 345
Copenhagen interpretation, 18
Correlated state, 406, 552
Correlated wave function, 361
Correlation function, for field operators, 585–586
Correspondence principle, 3, 37, 324
Coulomb barrier, 127
Coulomb excitation, 487–491
Coulomb gauge for radiation field, 491, 572–573, 582
Coulomb interaction between identical particles, 553
Coulomb potential, 265–275, 625. *See also* Hydrogen atom
Coulomb scattering, 310–312. *See also* Rutherford scattering cross section
Counting rate, 280
Coupled harmonic oscillators, 371, 480, 568
Covalent bond, 164
Covariance, 635
Creation operator, 538. *See also* Raising operator
Crossing of energy eigenvalues, 465–466
Cross section, 278–286
  absorption, 494–501, 510, 514, 591
  Coulomb scattering, *see* Rutherford scattering cross section
  differential scattering, 281, 290, 301, 312, 401, 520, 525
  partial, 301, 304–306
  photoemission, 502
  resonance, 306
  total scattering, 281, 290, 301–302
Current density, 26, 553, 600, 610, 623. *See also* Probability current density
  gauge invariant, 74
Current distribution, interaction with photon field, 580–582
Curvilinear coordinates, 638–640

de Broglie, Louis, 2
de Broglie relation, 2, 12
de Broglie wavelength, 2, 115
Debye potential, 277, *see also* Screened Coulomb potential; Yukawa potential
Decaying state, 132–133, 307, 392, 514. *See also* Exponential decay
Degeneracy of energy eigenvalues, 44–45, 144, 207. *See also* Repeated eigenvalues

absence of, in linear harmonic oscillator, 83
connection with group representation, 419–420
for free particle, 65–66
for hydrogen atom, 267–270
for isotropic harmonic oscillator, 276
for periodic potential, 176
removal of, by perturbation, 144–146
Degenerate perturbation theory, 463–467
for periodic potential at band edges, 481
Delta function, 630–634
partial wave expansion of, 252
for solid angle, 252
Delta function normalization, 61
Delta function potential, 107–108, 206
Δ-variation, 476
δ-variation, 474–475
de Moivre formula, spin generalization of, 387
Density matrix:
spin one-half, 392–399
photons, 587–589
Density of states, 62, 65–67, 501, 504, 578
Density operator, 319, 322, 363, 370
for chaotic state, 589
for thermodynamic equilibrium, 564–565
Detailed balancing, 493
Determinantal equation, 140–141. *See also* Secular equation
Deuteron, energy and wave function, 275
Diagonalization of normal matrix, 209–211
in degenerate perturbation theory, 463
Diamagnetic susceptibility, 481
Diatomic molecule, model of, 112, 163–165
Dichotomic variable, 378
Differential cross section, *see* Cross section
Dipole moment, *see* Electric dipole moment, Magnetic moment
Dirac, Paul A. M., 196, 594
Dirac equation:
for electron, 469, 596–600
for free particle field, 606–608
Dirac field, 594
adjoint, 599
Dirac Hamiltonian, 596
Dirac matrices:
α and β matrices, 597
γ matrices, 598–599
physical interpretation, 617–620
standard representation, 597, 603
Dirac picture, *see* Interaction picture
Dirac spinor, 595
Direct integral, 479, 553, 558
Direct product:
of irreducible representations, 431–432
of matrices, 358, 431
of vector spaces, 358, 426, 430–431
Discrete spectrum, 43–44, 83, 181
Dispersion function, 166
extended-zone scheme, 172
reduced-zone scheme, 172. *See also* Brillouin zone
repeated-zone scheme, 167, 173
Displacement operator, 68–71, 165, 225
eigenvalue problem of, 70
Distorted plane wave, 300
Double oscillator, 149–159
Double scattering, 376–377, 403
Double-slit interference, *see* Two-slit interference
Double-valued representations of rotation, 387, 424
Driven harmonic oscillator, *see* Forced harmonic oscillator
$D(R)$-matrix for rotations, 423, *see also* Rotation matrix
Dual vector space, 196
Dynamical variable, 38, 53, 57

Ehrenfest's theorem, 36–37
Eigenfrequency, 5
Eigenfunction, 42,54
Eigenket, 198
Eigenstate, 54
Eigenvalue, 42, 54, 140, 198
Eigenvalue problem for normal operator, 207–214
Eigenvector, 198
Einstein, Albert, 2
Einstein *A* coefficient, 580
Einstein, Podolsky, Rosen (EPR), 361–362
Einstein principle of relativity, 600
Elastic scattering, 284–286, 518
of alpha particles, 284–286
of electrons by hydrogen atoms, 534
Electric dipole moment, 441, 459–463
matrix element of, 88, 489, 496, 579
induced, 461
permanent, 461
Electric dipole (E1) transition, 489, 496, 579
selection rules for, 436, 441, 497
Electric multipole expansion, 437, *see also* Multipole expansion of Coulomb interaction
Electric multipole (2ℓ) moment, 437, 488
parity and time reversal selection rules, 444
Electric quadrupole transition, 516
Electromagnetic field, quantization of, 569–573
Electron, relativistic theory of, 592–629
in Coulomb field, 625
in magnetic field, 620, 629
Electron-positron field, 592
charge of, 593–596
energy of, 593–596
momentum of, 593–595
Electrostatic polarization of atoms, 459–463
Elementary excitation, *see* Quantum; Quasiparticle
Emission probability, 577–579, 581–582
Energy bands, allowed and forbidden, in periodic potential, 166–178, 481
Energy eigenvalues, 42–44
of configuration p^2, 568
for delta function potential, 107–108, 208
for double oscillator, 153–155
for exponential well (S states), 275
for free particle:
nonrelativistic, 63
periodic boundary conditions, 64–66
relativistic, 592, 606
for harmonic oscillator, 83
for helium atom, 477–480, 505, 560
for hydrogen(ic) atom:
nonrelativistic, 267, 269
nuclear size effect, 277
relativistic, 627–628
for linear potential, 124
for particle in a box:
one dimensional, 106–107
three dimensional, 66
for positronium, 274–275
for scattering states, 522
for square well:
one dimensional, 105–106
three dimensional, 262–263
Energy gap, for periodic potential, 168, 172, 481
Energy level crossing, *see* Crossing of energy eigenvalues
Energy levels, *see* Energy eigenvalues
Energy normalization, 63–64
Energy operator, 35, 54
Energy representation, 206, 334
Energy shell, 531
Energy shift, 144, 452, 476, 522
Energy transfer, average:
in inelastic process, 486, 489–490
in absorption from light pulse, 494
Ensemble, *see* Statistical ensemble
Entangled state, 362, 406
Entropy, 636. *See also* Outcome entropy, von Neumann entropy, Shannon entropy
Equation of motion:
for density matrix, 395
for density operator, 319, 322, 369
for expectation value, 37, 319
in integral form, 338
for operator, 321–322,
in second quantization, 550–551
for spin state, 390–392
for state vector, 317, 482
for wave function, 37, 41, 348, 615–616
Equivalence transformation, 418
Equivalent representations, 418
Euclidean principle of relativity, 410
Euler angles, 242, 424, 449–450
Euler-Lagrange equation, 135
Euler's theorem for homogeneous functions, 48

Exchange, 154–155, 477–479, 553, 562
Exchange degeneracy, 477
Exchange integral, 479, 558
Excited states, 44
Exciton, 222
Exclusion principle, *see* Pauli exclusion principle
Expansion of wave packet, in terms of:
  coherent states, 227
  Dirac wave functions, 616
  eigenfunctions of an observable, 60–62
  momentum eigenfunctions, 15, 19, 30, 63, 286
  oscillator eigenfunctions, 89
  scattering eigenfunctions, 287, 294
  spherical harmonics, 249
Expansion postulate, 46, 59
Expectation value, 29, 198, 634
  of Dirac matrices, 619–620
  in Dirac theory, 615
  equation of motion for, 37, 319
  of function of position and momentum, 33, 35
  of observable, 198
  of operator in spin space, 393
  of position and momentum, in coordinate and momentum space, 29–32
Exponential decay, 133, 510–513
Exponential operator, 39–41
Exponential potential, 275
Extended Euclidean principle of relativity, 439

Fermi, Enrico, 507
Fermi-Dirac:
  anticommutation relations, 543, 547
  statistics, 543, 566
Fermi gas, 567
Fermions, 543
Feynman, Richard F., 185
Feynman path integral, 355–357
Field operators, 546, 551
  bilinear functions of, 605
Fine structure, of one-electron atom, 470, 627–628
Fine structure constant, 471–472, 641
Flavor, quantum number, 536
Floquet's theorem, 166
Fock space, 537
Foldy-Wouthuysen transformation, 618
Forbidden energy gap, 168, 172, 481
Forbidden transitions, 496–497
Forced harmonic oscillator, 335–342, 354, 486, 580
Form factor for scattering from finite nucleus, 534
Forward scattering amplitude, 302, 533
Fourier analysis, 15, 64, 630
Franck-Hertz experiment, 1
Free energy, generalized, 565
Free particle eigenfunctions:
  nonrelativistic, 44, 62–65
    normalization of, 63
    with sharp angular momentum, 257–262
  relativistic, 606–609
Free particle motion:
  one dimensional, 22–24
  propagator for, 351
  relativistic, 618–619
Functional integral, 357
Function of a normal operator, 212, 217

Galilean transformation, 5–6, 75–78
Gamow factor, 313
Gauge, 73
  field, 75, 445
  invariance, 71–75, 347
  natural, 75
  principle, 599
  symmetry, 75, 444–447
  theory, electroweak, 538
Gauge transformation:
  global, 73
  local, 73, 347, 444–447
Generalized free energy, 565
Generating function:
  for associated Laguerre polynomials, 271
  for Hermite polynomials, 85–86
  for Legendre polynomials, 246–247
Generator:
  of infinitesimal rotations, 236, 330–331, 382, 449
  of infinitesimal transformations, 328
  of infinitesimal translations, 70–71, 236, 330
g-factor:
  for atom, 439
  for electron, 620–621
  for nucleus, 449
Golden rule of time-dependent perturbation theory, 503–510
Gordon decomposition of Dirac current, 623
Good quantum number, 473
Goudsmit, Sam, 374
Grand canonical ensemble, 565
Grand canonical potential, 565
Grand partition function, 565
Green's function, 290, 349, 454
  advanced, 293, 337, 350
  for harmonic oscillator, 337
  incoming, 293
  outgoing, 293
  partial wave expansion of, 308
  in perturbation theory, 457–458
  retarded, 293, 337, 349–350
  in scattering theory, 290–295, 523–524
  standing wave, 293
  for wave equation, 349–350
Green's operator in scattering, 523–524
Ground state, 44, 84, 222
  variational determination of, 136, 213
Group, definition, 69, 416–417
Group representation, 416–421
Group velocity, 19, 175
Gyromagnetic ratio, 374, 398. *See also* *g*-factor

Hamiltonian, for particle in electromagnetic field, 72
Hamiltonian operator, 37, 317, 348
Hamilton-Jacobi equation, 23, 25, 114, 352
Hamilton's equations, 80, 324
Hamilton's principal function, 23–24, 354–355
Hamilton's principle, 355, 357
Hankel functions, spherical, 260
Hard sphere scattering phase shift, 303
Harmonic oscillator, 79–89, 205, 220–225
  coupled, 371, 480
    in Hartree-Fock theory, 568
  density operator for, 590
  eigenfunctions, 47, 83–89, 224
    recursion relations for, 81–82
  energy eigenvalues, 2, 83, 125
  in Heisenberg picture, 333
  in momentum representation, 34, 47
  propagator for, 352
  and reflection operator, 440
  in thermal equilibrium, 91
  three-dimensional isotropic, 276, 480
  and time development of wave packet, 49
  two-dimensional isotropic, 276, 567
  and uncertainties in x and p, 49
  WKB approximation for, 125
  zero-point energy for, 84
Hartree-Fock equations, 562
Hartree-Fock method, 560–564
Hartree units, 641
Heaviside step function, 93, 342, 633
Heisenberg, Werner, 18
Heisenberg (spin) Hamiltonian, 567
Heisenberg picture, 320, 550
  applied to harmonic oscillator, 333–335
  and canonical quantization, 321–322
  for Dirac field, 598
  in one-particle Dirac theory, 617–621
Heisenberg uncertainty principle, 18
Heisenberg uncertainty relations, 20–22, 217–220
  for angular momentum components, 240
  for energy and time, 21–22, 43
  for position and momentum, 14–18, 20, 219, 229, 231–232
  in second quantization, 553
Helicity, 449, 569, 576
Helium atom:
  energy levels, 477–480, 505
  stationary states, 477, 560
Hellmann-Feynman theorem, 175, 178, 465, 476
Hermite polynomials, 84–86
  completeness of, 88
  differential equation for, 84
  generating function for, 85
  integral representation for, 88
  normalization of, 87
  orthogonality of, 87
  recurrence relation for, 84, 224–225

Hermitian adjoint operator, 197. *See also* Adjoint of an operator
Hermitian conjugate (adjoint) of a matrix, 100, 192, 380
Hermitian matrix, 193, 380
Hermitian operator(s), 51–56, 192. *See also* Normal operator(s)
  eigenvalue problem for, 54, 212–214
  as observables, 53, 179–180
Hermitian scalar product, *see* Inner product
Hidden variables, 9, 18
Hilbert space, 185
Hindered rotation, 158, 481
Hole state in shell model, 567
Hole theory and positrons, 618
Holonomy, 447. *See also* Berry's phase
Hydrogen(ic) atom, 265–275, 623–628
  degeneracy in, 267, 468–469, 628
  effect of electric field on, 459–460, 467–469,
  eigenfunctions of, 270–275
    recursion relations for, 266
  emission of light from, 580
  energy levels of, 267, 269, 627–628
  fine structure of, 627
  lifetime of excited state, 580
  linear Stark effect, 467–469
  in momentum space, 502
  parity in Dirac theory of, 624
  reduced mass effect in spectrum, 274
  relativistic correction to kinetic energy, 481
  and rotational symmetry in four dimensions, 268–270
  and WKB method, 275

Ideal experiment (measurement), 406, 408, 515
Ideal gas, in quantum statistics, 565
Idempotent operator (matrix), 69, 189, 394
Identical particles, 535
  quantum dynamics of, 549–552
  and symmetry of wave function, 547
Identity operator, 189
Impact parameter, 282, 488–489
Impulsive change of Hamiltonian, 342
Impulsive measuring interaction, 408
Incoherent sum, partial wave cross sections, 301
Incoming spherical waves in asymptotic scattering eigenfunctions, 294, 502, 524
Incoming wave, 100
Incoming wave Green's functions, 293, 524
Incompatibility of simultaneous values of noncommuting observables, 53, 407
Indenumerable set of basis vectors, 202
Independent particle approximation, 559, 560–561, 564
Indeterminacy principle, *see* Heisenberg uncertainty principle
Indistinguishability of identical particles, 535–538
Induced electric dipole moment, *see* Electric dipole moment
Infinitely deep well, 106–107, 275
Infinitely high potential barrier, 95
Infinitesimal displacement, 234
Infinitesimal rotations, 235, 330, 382
  representation of, 423
Infinitesimal transformations, 328
Infinitesimal translations, 70, 236, 330
Inflection point of wave function, 94
Information, 636
  in quantum mechanics, 363–370, 403–408
Infrared vibration, of oscillator, 159
Inhomogeneous linear equation, 453–455
Inner product:
  bra-ket notation for, 196
  of functions, 59
  of vectors, 187
"In" states in scattering, 518
Integral equation:
  for radial wave function, 309
  for scattering state, 293, 521–525
  for stationary state wave function, 291
  for time-dependent wave function, 549
  for time development operator, 338
Interacting fields, 577
Interaction between states or energy levels, 146, 167, 178
Interaction picture, 323, 483
Interference, 8, 12–13, 184–185, 357, 583–589
Intermediate states, 509
Internal conversion, *see* Radiationless transition
Interpretation of quantum mechanics:
  Copenhagen, 18
  ontological, 9, 29
  realistic, 9
  statistical, 25–29, 408
Interval rule, 470
Intrinsic angular momentum, 372–377, 389. *See also* Spin
Invariance:
  under canonical transformations, 329–330
  under charge conjugation, 609–610
  under *CP* transformations, 409
  under gauge transformations, 71–75, 347, 445
  under Lorentz transformations, 600–605
  under reflections, 81, 101, 441, 460, 605
  under rotations, 233, 269–270, 330–331, 383, 390, 414 , 530–532
  under time reversal, 46, 100, 167, 442, 612
  under translations, 165, 330
Invariant subspace, 419
Inverse of an operator, 69, 194
Inversion of coordinates, 249, 439, 605
Irreducible representations ("irreps"), 418, 421
  of rotation group, 423
  of translation group, 166
Irreducible spherical tensor operator, 434
  commutation relations for, 436
  selection rules for, 435–436
  time reversal properties of, 443–444
Isobaric spin, *see* Isospin
Isometric mapping, 411
Isometric operator W, in scattering, 530
Isospin, 445, 536

Joining conditions for wave function, 46, 95, 104, 157
Jones vector, 395

Kernel for Schrödinger equation, 349. *See also* Green's function
Ket, 196
Kinetic energy operator, 35
  and orbital angular momentum, 252–255
Klein-Gordon equation, 621
k-Normalization, 63
Koopman's theorem, 563
Kramers, Hendrik A., 113
Kramers degeneracy, 442, 612
Kronig-Penney potential, 168–169, 481
Kummer function, 156

Ladder method:
  for angular momentum, 239
  for harmonic oscillator, 221
Lagrangian multiplier, 136
Laguerre polynomials, associated, 271
  and confluent hypergeometric functions, 270–271
Lamb shift, 628
Landé *g*-factor, *see* *g*-factor
Landé's interval rule, 470
Laplacian, in curvilinear coordinates, 639
Larmor precession, 398
Laser, as two-level system, 500
Laue condition, 295
Law of large numbers, 635
Legendre's differential equation, 244
Legendre polynomial expansion *see* Partial wave expansion
Legendre polynomials, 245
  completeness of, 249
  generating function for, 246
  normalization of, 246
  orthogonality of, 246
  recurrence fromula for, 247
  recursion relation for, 345
Levi-Civita (antisymmetric) tensor symbol, 252
Lie group, semisimple, 421–422
Lifetime of decaying state, 133, 307–308, 513–514, 580
Light pulse, and its absorption, 492–494

Light quantum, *see* Photons
Linear displacement operator, *see* Translation operator
Linear harmonic oscillator, *see* Harmonic Oscillator
Linear independence, 186
  of eigenfunctions, 55
Linear momentum, *see* Momentum
Linear operator, 34, 188, 412
Linear potential, 123
  energy eigenvalues for, 124
  ground state wave function for, 139
  variational estimate for ground state of, 138
  and WKB approximation, 139
Linear Stark effect, 467–469
Linear vector space, *see* Vector space
Line broadening, 500
Line shape, 514
Lippmann-Schwinger equation, 522
Local interaction, of identical particles, 549
Logarithmic derivative of wave function, 105, 302, 304
Longitudinal polarization of particle with spin, 400
Lorentz boost, 607
Lorentz equation, 618
Lorentz group, 600
Lorentz transformation, 601
  infinitesimal, 602
  proper orthochronous, 601
Lowering operator, 221. *See also* Annihilation operator
  eigenvalue problem for 225
L-S coupling, 559
Luminosity, 279–280

Madelung flow, 28
Magnetic moment, 372–375, 438
  of atom, 438–439
  of electron, 374, 388, 620, 622
  of nucleus, 449
Magnetic quantum number, 244
Magnetic resonance, 399
Maser, as two-level system, 500
Matching conditions, *see* Joining conditions for wave function
Matrix element(s):
  of operator, 191, 198
    in coordinate representation, 204, 345
    in oscillator energy representation, 88, 223–224
Matrix mechanics, 142
Matrix methods, for transmission and reflection in one dimension, 97–99, 108–109
Matter waves, 2
Maxwell-Boltzmann statistics, 566
Maxwell equations, 573
Measurement of observables, 53, 57, 364, 370, 403–408
  ideal, of first kind, 408
Mehler's formula for Hermite polynomials, 89, 353
Minimum uncertainty (product) state (wave packet), 220, 229–230, 232
  time development of, 333,351
Mixing entropy, *see* Shannon entropy
Mixture of states, 365, 399
Mode(s), of elastic medium or field, 4, 569, 584
Momentum eigenfunction, 62–65
  partial wave expansion of, 261
Momentum:
  canonical, 72
  expectation value of, 32, 36, 90
  kinetic, gauge invariant, 74
  local, 115
  of photon field, 574–575
  radial, 255
Momentum operator, 35, 62, 71, 204
  matrix element of, 205
Momentum representation, 30–33
  and equation of motion, 30–31, 347–348
  for harmonic oscillator, 34, 47, 329
  for hydrogen atom, 502
  wave function in, 180, 345
Momentum space, *see* Momentum representation
Momentum transfer, 296
Multinomial distribution, 635
Multiple scattering, 286
Multiplet, of spectral lines, 420
Multipole expansion, of Coulomb interaction, 308, 488, 507, 568

Nats, 367
Natural units, 641
Negative energy states in Dirac electron theory, 616
Negative frequency part of field, 572, 594
Neumann, John von, 52
Neumann functions, spherical, 259
Neutral kaon, decay of, 408–409
Neutrino, 629
Nodal line, defining Euler angles, 424
Nodes:
  as adiabatic invariants, 151
  of oscillator eigenfunctions, 87
  of hydrogen atom eigenfunctions, 274
  of square well eigenfunctions, 106
  of WKB bound state wave function, 122
Noncrossing of energy levels, 465
Nonorthogonal basis functions, 146–149
Nonorthogonal projection operators, for generalized measurement, 364–365
Nonrelativistic limit of Dirac theory, 622
No-particle state, 222, 537
Norm, of state vector, 59, 187
Normalization, 27–28, 57, 187
  of associated Laguerre functions, 270
  of associated Legendre functions, 247
  of coherent states, 225
  of continuum eigenfunctions, 61, 203
  of Coulomb eigenfunctions, 313
  of free particle eigenfunctions, 62–65
  of hydrogen eigenfunctions, 270
  of identical particle states, 556
  of Legendre polynomials, 246
  of momentum space wave functions, 31
  of oscillator eigenfunctions, 87
  of perturbation eigenvectors, 456–457
  of radial eigenfunctions, 263
    in continuum, 300
  of scattering states, 527
  of spherical harmonics, 249
  of spinors, 393
Normal operator, 195
  eigenvalue problem of, 207–211
Normal ordering of operators, 228, 558
Null vector, 187
Number of particles operator, 83, 222, *see also* Occupation number operator

$O(n)$, orthogonal group, 421
Observables, 59, 180
  commuting and compatible, 214–217, 407
    complete set of, 180, 216
  simultaneously measurable, 180, 214–217
Occupation number operator, 537, 542
Old quantum theory, 241
One-electron atoms, spectra of, 469–471
One-electron state(s), relativistic, 613–614
One-form, 196
One-particle operator, additive, 544–545, 615
Opacity of barrier, 127
Operators, 34–38, 188–195
  algebra of, 38–41
Optical potential, 27
Optical theorem, 103, 112, 302, 532–533
Orbital angular momentum, 233–255, 425–426, 443
  eigenvalues:
    of component of, 242–244
    of magnitude of, 244–245
Orbital angular momentum quantum number, 245
Ordering:
  of noncommuting operators, 33, 325
  normal, 228, 558
  time, 338, 484
Orthogonality:
  of continuum eigenfunctions, 61
  of eigenfunctions of Hermitian operators, 55
  of eigenvectors of normal operators, 208–209
  of scattering states, 527
  of spinors, 379
  of state vectors, 187
  of stationary states, 43
Orthohelium, 480, 560
Orthonormality, 56, 187
Orthonormal set, basis vectors, 55, 187, 201, 537
Oscillator, *see* Harmonic Oscillator
Oscillator strength, 488
Outcome entropy, 368, 404–405
Outer product, *see* Direct product

Outgoing spherical waves in asymptotic scattering eigenfunctions, 287, 294, 502, 523
Outgoing wave, 100
Outgoing wave Green's function, 293, 523
"Out" states in scattering, 518
Overcomplete set of coherent states, 227, 365
Overlap integral, 147, 153

Pair, electron-positron, annihilation of, 616
Pair density operator, 567
Pair distribution operator, 545
Pair state, 556
Parabolic coordinates, 310, 462, 639
Parabolic cylinder functions, 156–157
Parahelium, 480, 560
Parity, 81, 440
  and angular momentum, 249
  conservation of, 441, 460
  in Dirac theory, 605, 610–611
  and electric dipole moment, 441, 460
  nonconservation of, 441
  operator, 249, 441, 605
  selection rules, 441
  in spin space, 440
Parseval's equality, 59
Partial wave cross section, 301
Partial wave expansion:
  of delta function, 252
  of Green's function, 308
  of plane wave, 261
  of scattering amplitude, 301, 531
  of *S* matrix, 531
Particle-antiparticle transformation, 408, 608–610
Particle-antiparticle oscillation, 409
Particle density operator, 553, *see also* Probability density operator
Particle in a box, 66–67
Partition function, 637
Passive transformation, 76–77, 201, 602
Pauli exclusion principle, 543
Pauli spin matrices, 386, 603
PCT theorem, 613
Penetrability, 128
Penetration of potential barrier, *see* Barrier penetration
Periodic boundary conditions, 45, 64–66, 107
Periodic potential, 156–176
  eigenvalue problem for, 168–173
  perturbation theory for, 481
Perturbation, 128
Perturbation expansion, 452, 475
  arbitrary constants in, 456–457
  to first order, 452–453, 455–459
  to second order, 456–459, 461–462
Perturbation theory, 142–146, 451–459
  for degenerate levels, 144–146, 463–465
  for *n*-electron atom, 558–560
Phase integral, 2, 122

Phase shift, 110, 298–309, 631
  Born approximation for, 307
  integral formula for, 309
  in transmission through a barrier, 110
Phase space, in WKB approximation, 122
Phonons, 3, 222
Photoelectric effect, 501–502, 515
Photoemission, 515–516
Photon correlations, 586–589
Photon field operator(s), 572
Photons, 3, 222, 569
  absorption of, 492–493, 577–579
  detection of, 583
  emission of, 577–580
  orbital angular momentum of, 575
  spin of, 569, 575–576
Picture, of quantum dynamics, 319–323,
  Heisenberg, 320
  interaction (Dirac), 323, 483
  Schrödinger, 316–320
Planck's constant, 1, 348, 641
Planck's black-body radiation formula, 590
Plane wave, 13–14, 43
  expansion in spherical harmonics, 261
p-Normalization, 63
Poincaré vector, 395
Poisson bracket, 326
Poisson distribution, 227, 341, 582
Polarizability:
  of atom, 461
  of hydrogen atom, 462
  of isotropic oscillator, 461
Polarization:
  of electron, 376–377
  of light, 576
Polarization current density, 623
Polarization vector, 376, 394
  and density matrix, 392–399, 403–404
  equation of motion for, 396
  precession of, 396–397
  and scattering, 376–377, 399–403
  for statistical ensemble, 403
Positive definite operator, 193
Positive frequency part of field, 572, 594
Positron, 592
  vacuum, 614
  wave function, 614–615
Positronium, decay of, 449
Positrons, sea of, 618
Potential:
  Coulomb, 265
  delta function, 107
  double oscillator, 149–150
  double well, 11
  exponential, 275,
  harmonic oscillator, 79
  hindered rotation, 158
  Kronig-Penney, 168
  linear, 123
  periodic, 165
  rectangular barrier, 97
  sectionally constant, 92
  spherically symmetric (central), 256
  spherical square well, 262
  square well, 103

Potential barrier, 97
Potential energy surface, 163
Potential step, 92
Poynting vector, 494
Principal quantum number, 267, 311, 627
Principle of complementarity, 4
Principle of relativity, 75
Principle of superposition, 12–14, 57–58
  and time development, 316
Probability:
  basic theory of, 634–638
  in coordinate and momentum space, 29–34
  conservation of, *see* Conservation of probability
  current density, 26–27
  in Dirac theory, 600, 610, 616, 623
  as expectation value of operator, 49
  gauge invariant form of, 74
  represented by Wigner distribution, 49, 370
  density, 26–27, 29–30, 203
  in Dirac theory, 616
  as expectation value of operator, 49
  in momentum space, 32–34
  represented by Wigner distribution, 49
  interpretation, 7, 9, 25–29, 57
  sources and sinks of, 78
  in spin theory 380, 403
Probability amplitude(s), 8, 59, 179, 195
  closure relation for, 183
  composition rule for, 182
  interference of, 182
  as inner product, 195
  orthonormality of, 183
  reciprocal property of, 182
  time development of, 315
Probability distribution, of radial coordinate in hydrogen(ic) atom, 274
Projection operator, 189, 217, 364, 393, 404
  rank of, 217
Propagator, 349
  for free particle, 351
  for harmonic oscillator, 352
Pseudoscalar operator in Dirac theory, 605
Pure state, 366

Quadratic integrability, 27
Quadratic Stark effect, 460
Quadrupole approximation, 516
Quadrupole interaction, 450
Quantization postulates, rules, 323–326
Quantum (quanta), 3, 222
Quantum condition, 2, 122
Quantum correlations, 228,262
Quantum defect, 268
Quantum chromodynamics (QCD), 538
Quantum electrodynamics (QED), 538, 577
Quantum field operator, 546
Quantum field theory, 551
Quantum fluctuations, 228

Quantum measurement theory, 363–365, 370, 408
Quantum number(s), 84, 473
  group theoretical meaning, 422
Quantum of action, 1
Quantum potential, 29, 354
Quantum theory of radiation, 501
Quarks, 536
Quasiclassical states, 228, *see also* Coherent states
Quasiparticle, 222
Quasiparticle transformation, 231

R(3), rotation group in three dimensions, 421
Radial Dirac equation, 625
Radial eigenfunction, 257
  boundary condition for, 263
Radial Schrödinger equation, 257, 263–265
Radiation, *see* Absorption and Emission of radiation
Radiation field, quantum theory of, 569–576
Radiationless transition(s), 504–505, 507–508
Raising operator, 221. *See also* Creation operator
Randomness, 366–367
Random variable, 638
Rank:
  of group, 421
  of projection operator, 217, 364
Rate of transition, 503–510, 520–521
Rayleigh-Ritz trial function, 139–142
Rayleigh-Schrödinger perturbation theory, 451–459
  and variational method, 473–476
Reciprocal basis, 147
Reciprocal lattice, 71, 167–168
Reciprocal lattice vector, 314
Reciprocity relation, 532
Rectangular potential barrier, 97
Rectangular well, *see* Square well
Reduced matrix element, 435
Reduction:
  of direct product representation, 431–432, 557
  of group representation, 418
  of state by measurement, 408
Reflection, 439
  of coordinates, 81
  and rotation, 440–441
  of incident wave, 96
Reflection coefficient, 96
Reflection operator, 440
Regeneration, of amplitudes, 407
Relative motion, 149, 274, 359–360
Relative probabilities, 28
Relativistic invariance of Dirac equation, 600–606
Relativistic Schrödinger equation for scalar particle, 621, 629
Relativistic wave equation for electron, 621
Repeated eigenvalues, 56, 207, 214
Representation,of groups, 417–421
  in quantum mechanics, 191,199. *See also* Coordinate representation; Energy representation; Momentum representation
  of rotations, 417, 421–426
    in spin space, 382–385, 388
  of state, by entire functions, 228
Repulsion of perturbed energy levels, 462
Resolvent operator, 525
Resonance, in spin precession, 397
  magnetic, 399
Resonance(s), 110
  profile of, 514
  in scattering, 289, 304–308
  spacing of, 130
  in transmission, 109–111
  and wave packets, 130–133, 289, 307
  width of, 130, 133, 304–306, 514
  in WKB approximation, 130
Riesz representation theorem, 188
Rigid rotator, 480
Rotation matrix, 383–384, 387, 423–426
  symmetry relations for, 425, 443
Rotation operator, 381–382, 413
Rotations, 234–236, 381–385, 417
Runge-Lenz vector, 268
Russell-Saunders (*L-S*) coupling, 559
Rutherford scattering cross section, 284, 297, 312–313

Saturation of absorption line, 500
Scalar operator, 236–237
Scalar operator in Dirac theory, 605
Scalar product, *see* Inner product
Scattering, 278–313
  in Coulomb field, 310–313
  of particles with spin, 399–403
  by square well, 108–11
  of wave packets, 286
Scattering amplitude, 289, 295
  in Born approximation, 296
  partial wave expansion of, 301
  and scattering matrix, 531
  for spin one-half particle, 399
  and transition matrix, 524
Scattering coefficient, 111, 533
Scattering cross section, *see* Cross section
Scattering equation, 525–527
Scattering matrix, 400, 519, 527, *see also S* matrix
  invariance of, 400, 530–532
  one-dimensional analogue of, 99–103
Scattering operator, 340, 528
  relation to time development operator, 529
  unitarity of, 529
Scattering phase shift, *see* Phase shift
Schmidt orthogonalization method, 55–56, 207
Schmidt values for magnetic moment of nucleus, 449
Schrödinger, Erwin, 5
Schrödinger equation, 42
  time-dependent, 25
  for relative motion, 359–360
  for two particles, in configuration space, 359
Schrödinger picture, 316–320, 617
Schrödinger representation, 345
Schrödinger's cat, 362
Schur's lemma, 421
Schwarz inequality, 193
Screened Coulomb potential, 277, 297
Screening constant for helium atom, 478–479
Second-order equation in Dirac theory, 621
Second -order perturbation theory, time-dependent, 508–509
Second quantization, 551
Sectionally constant potential, 92–112
Secular equation, 140, 209, 464, 473
Selection rule, 90
  for CG coefficients, 428
  for electric dipole transition, 497
  for electric multipole moments, 437, 441, 444
  for irreducible tensor operators, 435–436
  relation to symmetry, 466
Self-adjoint operator, 52, 192
Self-consistent solution, 552, 563
Self-reciprocal basis, 147
Semiclassical approximation, 24, 113
Semiclassical state, 228, *see also* Coherent state
Separable Hilbert space, 185
Separable scattering potential, 534
Separable two-particle wave function, 359, 361
Separation of variables, 257, 270
Shannon (mixing) entropy, 367, 403, 636
Shells, atomic, 559
Similarity transformation, 200
Simple eigenvalue, 56
Simple scattering, 518
Simultaneous measurements, 180, 214–217
Singlet state, 431
Single-valued wave function, 45, 243
Slater determinant, 564
S matrix, 100–103, 530–532
  eigenvalue of, 302, 532
  poles of, 105
  unitarity of, 529
SO(*n*), special orthogonal matrices, *n* dimensions, 421
S operator, *see* Scattering operator
Space quantization, 373
Spectral decomposition, 217
Spectroscopic stability, principle of, 499
Spectrum, 54, 181
  of Schrödinger equation, 44
Spherical cylinder functions, Bessel, Hankel, Neumann functions, 259–260
Spherical harmonics, 248–252
  and harmonic functions, 254
  in momentum space, 443
  reflection properties of, 249
  and rotation matrices, 425–426
Spherical polar coordinates, 242
Spin, 372, 390
  of photon, 575–576
  operators, 385–390
  and statistics, 543, 556
  quantum dynamics of, 390–392
  total, 430–431

Spin filter, 408
Spin flip amplitude, 401
Spin matrices, in Dirac theory, 603. *See also* Pauli spin matrices
Spin one-half bosons, 556–558
Spin-orbit interaction, 389, 399, 416, 469–473, 480
Spinors, 379,595
Spin polarization, 392–399. *See also* Polarization
Spinor wave function, 378, 614
Spins, addition of, 430–431
  localized, 567
Spin variable, 378
Splitting of degenerate energy levels, 154–155, 178, 468, 472–474, 480
Spontaneous emission, 501, 579–580
Spontaneous symmetry breaking, 151–152
Spreading of wave packet, 20–21, 24, 49, 333, 351
Square well:
  in one dimension, 92–93, 103
    eigenvalues and eigenfunctions of, 103–108
    transmission through, 108–111
  in three dimensions, 262–263
Squeezed states, 230–231, 343
S state, 245, 265
  as ground state, 263
Standing wave Green's function, 293, 524
Stark effect, 460, 462
  linear, of hydrogen, 467–468
State, 28, 185
  pure, mixed, and unpolarized, 366, 399
State vector, 185, 388
  and wave function, 203
Stationary state, 41–47, 334–335
Statistical ensemble, density matrix for, 366, 399
Statistical thermodynamics, 369–370, 564–567
Statistics of particles, 554
Step function, *see* Heaviside step function
Stern-Gerlach experiment, 373–374, 406–408
Stieltjes integral, 181
Stimulated emission, 493, 499, 578
Stochastic process, 510
Stokes parameters, 395, 404
Sturm-Liouville equation, 59, 121, 261
$SU(2)$ group, 387
$SU(n)$, special unitary group, n dimensions, 421
Sudden approximation, 342
Sum rule:
  for electric dipole cross section, 489–490
  generalization of, 516
  for oscillator strengths, 489
  Thomas-Reiche-Kuhn, 489, 516
Superposition of states, *see* Principle of superposition
Superposition of stationary states, 44
Superselection rule, 414, 612
Symmetric top, 450
Symmetry:
  chiral, 620, 622
  four dimensional rotation, 269–270
  local gauge, 444–447
  reflection, 101
  rotational, 390
  of Schrödinger equation, 102
  of S matrix, 101, 105, 400
  time reversal, 100–101
  translational, 165
Symmetry group, 417
Symmetry operation, 411

Tensor operator, 432–437
Tensor operator in Dirac theory, 605
Tensor of polarizability, 461
Tensor product, *see* Direct product
Thermal equilibrium, 369, 564
Thomas-Reiche-Kuhn sum rule, 489, 516
Thomas (precession) term, 470
Tight-binding approximation, 167
Time delay in scattering resonance, 110, 307
Time-dependent perturbation theory, 485–487
Time-dependent Schrödinger equation, 22, 25, 41, 44, 46
Time development:
  of x, p, Δx, and Δp, 49, 332–333, 351
  of density operator (matrix), 319, 322, 369–370, 395
  of operators, 321, 332
  of physical system, 41–44, 315–319
  of polarization vector, 396–398
  of spin state, 390–392
  of state vector, 317, 482
Time development (evolution) operator, 41, 316, 484
Time-independent Schrödinger equation, *see* Schrödinger equation
Time-independent wave function, 42
Time-ordered product, 338, 484
Time reversal, 100, 441–444
  in Dirac theory, 611–612
  in scattering, 532
Total angular momentum, *see* Angular momentum
Transfer matrix, 169
Transformation coefficients, 199, 201, 205, 346, 538–539
Transition amplitude, 316, 323, 484
Transition current density, 26, 623
Transition matrix (element), 519–521
Transition probability per unit time, *see* Rate of transition
Translation operator, 69, 165,
Transmission coefficient, 96, 109, 126, 533
Transmission through barrier (WKB), 125–133
Transpose of an operator, 192
Transposition of matrix, 595
Triangular condition for adding angular momenta, 428
Triplet state, 431
Tunneling, 97–98, 125–133, 155, 167
Turing's paradox, *see* Zeno's paradox
Two-component theory of relativistic spin one-half particle, 629
Two-level system, 391
Two-particle matrix element, 545–546
  diagonal form, 551
Two-particle operator, 545, 555
Two-particle state, 555–556
Two-particle system, relative motion of, 359–360
Two-photon emission, 591
Two-slit interference, 8–9, 12, 183–185, 584–546

Uhlenbeck, George E., 374
Uncertainties, 218
Uncertainty principle, *18*
Uncertainty relation, *see* Heisenberg uncertainty relation
Unimodular matrix, 385, 387
Unitary matrix, 100, 195, 382,
Unitary operator, 68, 194,
  eigenvalues of, 210
Unitary symmetry, principle of, 539
Unitary transformation, 201
  and states of identical particles, 538–539
Unitary unimodular group in two dimensions, $SU(2)$, 424
Units, 640–641
Unit vector, 187
Universal covering group, 424
Unstable particles, 44

Vacuum expectation value, for electron-positron field, 596
Vacuum state, 222, 537
Variance, 16, 49, 634
  of observable, 218
Variational method, 135–139, 212–214, 474
  accuracy of, 481
  applied to helium atom, 478
  for n identical fermions, 560–562
  and perturbation theory, 473–476
Variational trial function, 137–140, 176–178, 276–277, 560
Vector addition coefficients, *see* Clebsch-Gordan coefficients
Vector model, in old quantum theory, 241, 438
Vector operator, 236, 383, 388, 433–434, 438
  commutation relations for, 236, 384
  Wigner-Eckart theorem for, 438
Vector operator in Dirac theory, 605
Vector potential as quantum field, 572
Velocity-dependent interaction, 335
Velocity operator in Dirac theory, 617
Virial theorem, 47–48, 177, 476
Virtual transition(s), 509
von Neumann entropy, 368, 564

Wave equation, 5, 25, 46, 347–348
  in momentum space, 46, 180, 348
Wave function, 5, 28, 180, 345
  complex valuedness of, 13
  in configuration space, 345, 547

Wave function (*Continued*)
meaning of, 4
in momentum space, 47, 345
for photon, 571
quantization of, 551
and Wigner distribution, 49–50
Wave mechanics, 142, 205
Wave packet, 14–18, 24
collapse of, 408
in oscillator potential, 89–90
scattering of, 286–290
splitting of, 96
spreding of, 20–22, 24, 49, 333, 351
in WKB approximation, 130–133
Wentzel, Gregor, 113
Width of resonance, 130, 133, 306, 514
Wigner coefficients, *see* Clebsch-Gordan coefficients
Wigner distribution, 49–50, 370–371
Wigner-Eckart theorem, 386, 435
applications of, 437–439
and time reversal, 444
Winding number, 414
WKB approximation, 113–134
applied to radial equation,
and bound states, 121–125
connection formulas for, 116
conservation of probability in, 112
and Coulomb potential (hydrogenic atom), 275
and double well potential, 134
and periodic potential, 178
Wronskian, 45, 121, 259

Yang-Mills field equations, 447
Yukawa potential, 277, 297

Zeeman effect, 473–474
Zeno's paradox, 514–515
Zero point energy, 84. 232, 574
Zitterbewegung, 619